



## Mini-review

# Meaningful respirometric measurements of UCP1-mediated thermogenesis



Yongguo Li, Tobias Fromme, Martin Klingenspor\*

Chair of Molecular Nutritional Medicine, Technical University Munich, Else Kröner-Fresenius Center for Nutritional Medicine, Gregor-Mendel-Str. 2, 85354 Freising, Germany

## ARTICLE INFO

## Article history:

Received 31 October 2016

Accepted 9 December 2016

Available online 14 December 2016

## Keywords:

Brown adipocyte

Brite/beige adipocyte

Mitochondria

Uncoupling protein 1

Uncoupled respiration

Fatty acids

Purine nucleotides

## ABSTRACT

Regarding the enormous interest in brown and brite/beige adipose tissue in the context of metabolic disease, reliable quantification of thermogenesis in these adipocytes is a central issue. This requires an assay specific for uncoupling protein 1 (UCP1) mediated thermogenesis in adherent intact cells. In a recent study we identified a major pitfall associated with established procedures generally applied for this purpose. Meaningful respirometry of UCP1-mediated thermogenesis imperatively requires activation of UCP1 and control over free fatty acid levels. By comparison of respiration profiles of wild-type (WT) and UCP1 knock-out (KO) cells we reproducibly quantified the thermogenic capacity enabled by UCP1 in both brown and brite adipocytes. Employing this protocol, we demonstrated that (1) brite adipocytes display a similar thermogenic capacity as classical brown adipocytes, (2) variations in brite adipogenesis known for inbred mouse strains are associated with differential capacities for thermogenesis in these cells, and (3) adipose triglyceride lipase (ATGL) activity is required for UCP1 activation in intact cells. We here further refined our cell-based respirometry assay by implementation of two strategies to inhibit UCP1 in WT cells. First, we employed the purine nucleotide guanosine diphosphate (GDP) to directly quantify the fraction of thermogenesis enabled by UCP1 activity. Second, applying siRNA mediated knockdown of UCP1 and ATGL we demonstrated the feasibility of this technology to study the functional relevance of candidate genes for thermogenesis in brown and brite adipocytes.

© 2016 Elsevier B.V. and Société Française de Biochimie et Biologie Moléculaire (SFBBM). All rights reserved.

## 1. Introduction

It is well established that the mitochondrial uncoupling protein 1 (UCP1) is responsible for nonshivering thermogenesis in brown adipose tissue (BAT) [1]. When cold sensation is relayed in the hypothalamus, catecholamines are released from the sympathetic nervous system, which activates adrenergic receptors of brown adipocytes and in turn stimulates the cAMP-dependent protein kinase PKA, leading to the activation of triacylglycerol lipases and thereby increased lipolysis. Adipose triglyceride lipase (ATGL) and hormone-sensitive lipase (HSL) are the major enzymes contributing to triacylglycerol breakdown [2]. Free fatty acids (FFAs) liberated by lipolysis serve at least four functions: 1) activate UCP1 by overcoming the constitutive inhibitory effects of cytosolic purine

nucleotides (i.e. guanosine diphosphate (GDP), guanosine triphosphate (GTP), adenosine diphosphate (ADP) and adenosine triphosphate (ATP)) on UCP1; 2) serve as a substrate for mitochondrial  $\beta$ -oxidation to fuel UCP1-mediated thermogenesis; 3) store energy by reesterification into triglycerides; 4) export into the bloodstream raising the plasma level of FFA. Upon activation, UCP1 short circuits the electrochemical proton gradient that drives ATP synthesis and thereby stimulates electron transport and enhances respiration. Heat is generated from the combustion of available substrates and is distributed to the rest of the body through blood circulation [3,4].

## 2. Free fatty acids and UCP1

Historically, even before the discovery of UCP1, the close relationship between FFA and the uncoupled state of brown fat mitochondria was well recognized. For example, from the early studies of isolated brown adipocyte mitochondria, it is clear that freshly isolated brown fat mitochondria contain significant amounts of

\* Corresponding author. Technical University Munich, Gregor-Mendel-Str. 2, 85354 Freising, Germany.

E-mail address: [mk@tum.de](mailto:mk@tum.de) (M. Klingenspor).

FFAs which are partially responsible for the uncoupled state [5,6]; while in intact brown fat cells, lipolytic stimuli markedly stimulate oxygen consumption [7]. Notably, the uncoupled state of brown fat mitochondria can be reversed by incubation with albumin to remove endogenous FFA and by addition of purine nucleotides [8], indicating regulatory roles of FFA and purine nucleotides on uncoupled respiration. Although mitochondria from any tissue can be uncoupled by sufficiently high fatty acid concentrations, brown fat mitochondria are extremely sensitive to FFA, suggesting a physiological role as regulators of uncoupling, from which the concept of ‘fatty acid uncoupling’ was developed and later confirmed and widely accepted [9]. This concept states that FFA concentrations in the nanomolar range activate UCP1-mediated uncoupling, not excluding that higher FFA levels in the micromolar range can also induce additional ‘fatty acid uncoupling’ as proposed by Wojtczak and Schonfeld [10] for any cell.

The mechanism by which FFA activate UCP1 is a matter of debate. The controversy is mainly centered on whether fatty acids are the substrate transported by UCP1 or if they are activators/cofactors for proton transport. Recently, Fedorenko and colleagues suggested that UCP1 operates in a long-chain fatty acid shuttling mode. In their model UCP1 is a symporter of fatty acid anions and protons that simultaneously transports one fatty acid anion and one proton per transport cycle. Due to strong hydrophobic interactions, however, the long-chain fatty acid anion cannot dissociate from UCP1 and is retained to initiate another proton translocation cycle [11]. Anyway, free fatty acid activation of UCP1 fits precisely within the physiological context, with their generation via  $\beta$ -adrenergic stimulation in the cell and utilization as substrate and UCP1-activator. Indeed, all manipulations that induce lipolysis in brown adipocytes will activate UCP1 and induce thermogenesis, and no UCP1 dependent thermogenesis can be evoked without simultaneously evoking lipolysis. In this scenario, lipolytic agents which are independent of adrenergic receptors are potential novel activators of brown adipose tissue and may therapeutically bypass the negative effect of sympathomimetics on the cardiovascular system.

### 3. Purine nucleotides and UCP1

In contrast to fatty acids, purine nucleotides are negative regulators of UCP1 activity. Purine nucleotides have been associated with UCP1 inhibition since the earliest studies on the protein [8]. Actually, the nucleotide sensitivity of UCP1 has been a key tool to investigate its properties. By taking advantage of nucleotide binding, Ricquier and Kader in 1976 found that a 32 kDa protein of unknown function was increased in brown fat mitochondria from cold-adapted rats [12]. Nevertheless, the nature of purine nucleotide interaction with fatty acids remains unresolved. Questions pertain to whether fatty acids displace nucleotides in order to activate UCP1, as well as how fatty acids can physiologically overcome inhibition of UCP1 by purine nucleotides (See review [4,13]). On the other hand, the noradrenaline-stimulation of respiration in isolated brown adipocytes is accompanied by a decrease in cytosolic purine nucleotides [14], and an upregulation of a purine degrading enzyme [15], suggesting a mechanism in which the downregulation of purine nucleotide concentration facilitates UCP1 activation. This hypothesis is unproven and has remained so for decades.

### 4. Albumin and uncoupled respiration

In respirometry studies of isolated brown adipocyte mitochondria, albumin must be used as fatty acid acceptor during isolation and measurement to observe coupled respiration. Essentially fatty

acid free bovine serum albumin (BSA) has generally been used for this purpose. In accordance, in early publications on oxygen consumption measurements of intact brown adipocytes following norepinephrine stimulation, 4% BSA (0.6 mM, similar to the concentration found in blood) was included into the respiration medium [7,16–18]. Under this condition, most of the fatty acids released by norepinephrine-stimulated lipolysis are exported from brown adipocytes, while the level of cell-associated free fatty acids does not change dramatically or even remains constant. However, when the extracellular free fatty acid buffering capacity is limited, e.g. in a fatty acid to albumin ratio well above 3:1, the level of cell-associated free fatty acids increases dramatically [19].

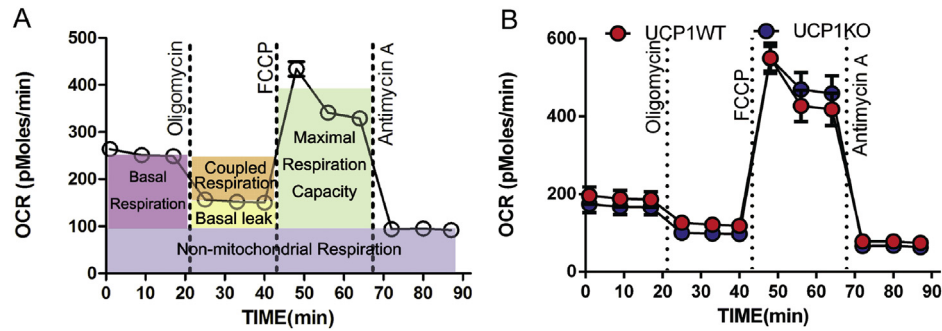
Shortly after the discovery of UCP1, based on the rationales that (1) activated brown fat releases fatty acids and (2) *in vivo* these released fatty acids are bound by albumin in blood, in 1979, Nedergaard and Lindberg investigated the effect of albumin on the metabolism of isolated brown fat cells. The addition of albumin increased norepinephrine-induced fatty acid release and induced a more stable norepinephrine-stimulated respiration rate [20]. This study provided direct evidence that the presence of albumin has beneficial effects on brown adipocyte metabolism, but the underlying mechanisms were not known. Consistent with previous findings, this study also demonstrated that a larger fraction of the fatty acids released upon norepinephrine stimulation is exported while less than 20% is combusted, indicating that fatty acid export is an intrinsic property for brown fat cells. Nevertheless, the mechanism(s) of FFA export from adipocytes, either through passive diffusion (flip-flop) or a protein-mediated process (fatty acid exporter), are not finally settled.

By using the albumin-buffering setup, in 2000, Matthias et al. reported a marked difference in oxygen consumption between isolated brown adipocytes from wild-type (WT) and UCP1 knockout (KO) mice, both following norepinephrine stimulation as well as addition of oleate [21]. This study also showed that there was no difference in basal respiration between genotypes, demonstrating UCP1 was not active without stimulation. Thus, in the basal state, the basal rate of lipolysis is not sufficient to override the inhibition of UCP1 thereby favoring coupled ATP-generating respiration. These observations were further confirmed by Shabalina et al. when studying the bioenergetics of WT and UCP1 KO brown-fat mitochondria [22]. In addition, investigations with trypsinized primary cultures of brown and brown-like cells (termed beige or brite adipocytes) by Petrovic et al., in 2008 and 2010, respectively, showed a correlation between UCP1 expression (as induced by rosiglitazone) and norepinephrine-stimulated oxygen consumption [23,24]. In all these studies, UCP1-mediated thermogenesis was observed when albumin was present in the respiration medium using a Warburg apparatus or Clark-type oxygen electrode systems.

### 5. Quantification of UCP1 mediated thermogenesis using microplate-based respirometry

Regulation of UCP1 within adherent intact brown-fat cells has been examined less frequently than its regulation in isolated mitochondria. Nowadays, microplate-based respirometry has become a mainstream method for measuring thermogenesis in cultured brown and brite cells. This technique allows for the study of mitochondrial bioenergetics in a higher-throughput fashion in intact, adherent cells [25]. The bioenergetic profile widely assayed in most studies published, so far, are illustrated in Fig. 1A: Basal respiration is determined first, after which oligomycin (5  $\mu$ M), an inhibitor of ATP synthase, is added to distinguish oxygen consumption used for ATP synthesis (coupled respiration) from basal proton leak (basal uncoupled respiration). The basal proton leak is





**Fig. 1.** The widely used mitochondrial bioenergetic profiling protocol fails to reveal any differences in differential respiration states between UCP1 WT and KO brown adipocytes. (A) Schematic representation of the mitochondrial bioenergetic profiling. Bioenergetic parameters including basal respiration, basal coupled respiration, basal proton leak respiration and maximal oxygen consumption as well as non-mitochondrial respiration can be measured following the addition of specific chemical effectors. (B) UCP1 WT and KO brown adipocytes have identical cellular bioenergetic profiles when using the mitochondrial bioenergetic profiling method, demonstrating UCP1 is inherently not leaky without stimulation. Part of this figure is reprinted with permission from ref. [27].

unregulated, compared to the inducible leak which is catalysed by specific mitochondrial inner membrane proteins and can be inhibited and activated (i.e. UCP1 mediated leak respiration). The precise mechanism of the basal proton leak is not fully understood. Proton leakage at the interface of mitochondrial inner membrane proteins and the phospholipid bilayer may contribute to the basal proton leak. Actually, up to two-thirds of the basal leak is attributable to the adenine nucleotide translocator (ANT), a transmembrane protein responsible for shuttling ATP and ADP across the mitochondrial inner membrane [26]. Next, an uncoupling agent such as carbonyl cyanide 4- (trifluoromethoxy) phenylhydrazone (FCCP, 1  $\mu$ M) is used to assess maximal respiratory capacity. Finally, antimycin A (5  $\mu$ M) is added to block the electron transport chain, leaving only non-mitochondrial oxygen consumption. Apparently, this method largely ignores the fact that UCP1 is inherently not leaky without stimulation. Consequently, using the procedure described above, there is no difference in any of these four different respiration states between UCP1 WT and KO brown adipocytes (Fig. 1B). Thus, any differential respiration between cell types using this protocol is not due to UCP1 but rather reflects differences in mitochondrial content [27].

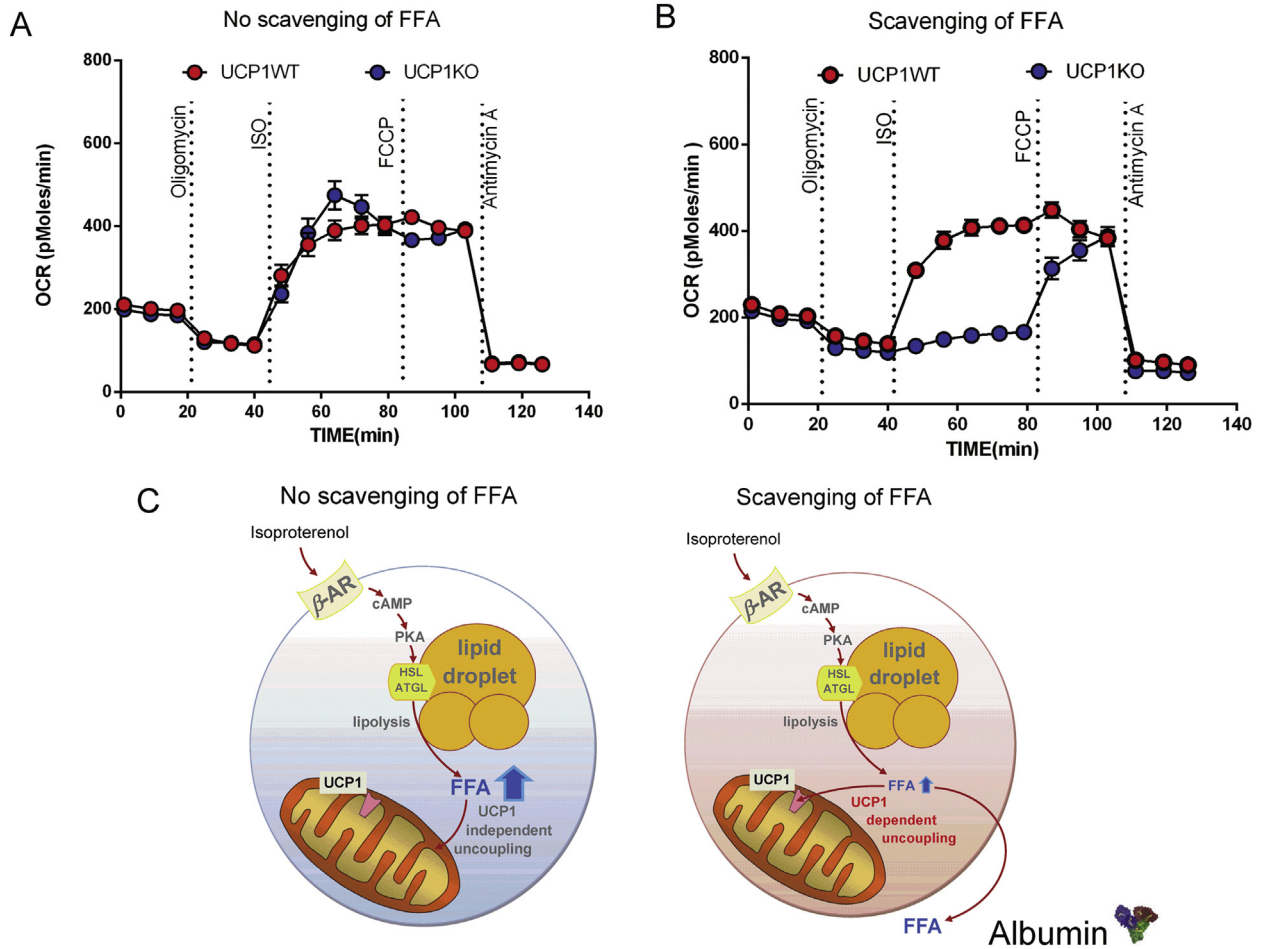
In brown adipocytes, FFAs liberated by lipolysis serve as activators of UCP1. We therefore modified the protocol as illustrated in Fig. 2A. After measuring coupled respiration and basal leak respiration, UCP1 was activated by FFAs released via  $\beta$ -adrenergic stimulation of lipolysis with isoproterenol (ISO, 0.5  $\mu$ M), and the increase in UCP1-mediated leak respiration was measured. Of note, albumin is not included in the respiration buffer, and not recommended, as defined by the provider of the microplate respirometry apparatus (Seahorse Biosciences), which neglects previously established respiration media for brown fat mitochondria and cells in suspension. In retrospect, it is not surprising that we observed identical ISO-induced leak respiration in UCP1 WT and KO primary cells (Fig. 2A). From our understanding, when UCP1 is activated by stimulation of lipolysis, it is essential to take control over intracellular FFA levels to avoid unspecific uncoupling effects of FFAs. Otherwise, an excessive rise of intracellular FFA levels released during lipolysis masks UCP1-mediated leak respiration (Fig. 2C) [27]. Accordingly, when applying a suitable concentration of essentially fatty acid-free BSA (2% BSA) to scavenge extracellular FFAs released by lipolysis in the respiration medium, we observed a prominent difference in uncoupled respiration between WT and KO cells, which represents UCP1-mediated uncoupled respiration (Fig. 2B and C).

Therefore, we conclude that meaningful respirometric measurements of cultured brown and brown-like (beige or brite)

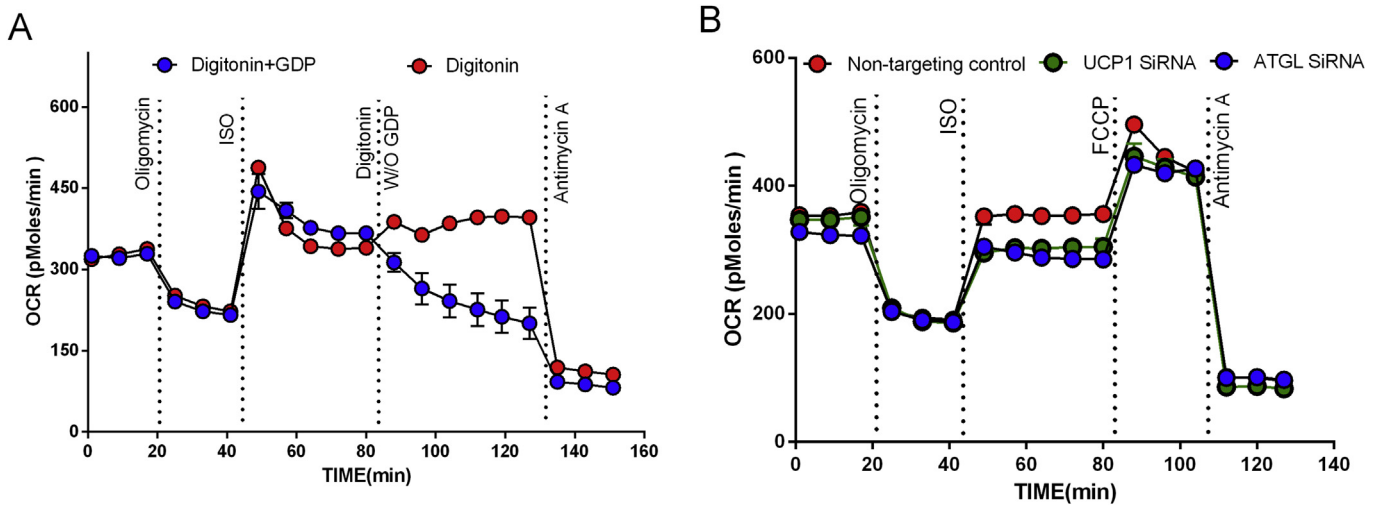
adipocytes imperatively require (1) activation of UCP1 and (2) control over free fatty acid levels. Employing this protocol, we for the first time demonstrated that brite adipocytes display a similar thermogenic capacity as classical brown adipocytes and strain differences in brite adipogenesis are associated with differential capacity for uncoupled respiration [27,28]. This protocol has been successfully applied in a few recent studies [29–31]. Recently, we found that the degree of adipocyte differentiation must also be considered, especially in the study of human adipocytes. When cell differentiation is sparse or patchy, the amount of free fatty acids released by a low number of lipid-filled adipocytes is quickly diluted in the respiration buffer and unspecific uncoupling artifacts of FFAs are avoided.

It is generally believed that it is through adrenergic stimulation of lipolysis and the resulting release of fatty acids within the cell that UCP1 is activated. Evidences supporting this notion include the following observations: 1) lipolytic stimulus markedly stimulated oxygen consumption in intact brown fat cells, which coincided with a rise in cytosolic fatty acid concentration; 2) exogenous addition of fatty acids activated UCP1 in isolated brown fat mitochondria. However, whether endogenously mobilized free fatty acids are required for UCP1 function has not been critically tested until recently. Employing our protocol and taking advantage of the available inhibitors targeting lipases involved in lipolysis, we demonstrated that inhibition of ATGL and HSL results in an almost complete blockage of UCP1-mediated uncoupled respiration [27]. To our knowledge, this is the first data set providing direct proof that intracellular mobilization of fatty acids is essential for UCP1 activation upon adrenergic stimulation.

Our study not only highlights the extreme sensitivity of UCP1 to fatty acids, but also reveals the necessity of using UCP1 knockout cells to isolate the UCP1-mediated component of uncoupled respiration. However, in the attempt to study novel molecular components of the thermogenic machinery in brown and brite adipocytes the demand to include UCP1 knockout cells as controls may be inconvenient and not always feasible. To circumvent this problem, we sought to validate the susceptibility of UCP1-mediated uncoupled respiration to inhibition by the purine nucleotide GDP. The cell membrane, however, is not permeable for GDP. Therefore, we modified our established 2% BSA-buffering setup as follows: after ISO-stimulation of UCP1-mediated uncoupled respiration, we simultaneously injected digitonin and GDP to permeabilize the cells and inhibit UCP1 activity. As expected, UCP1-mediated uncoupled respiration was gradually inhibited by GDP (20 mM), while digitonin (10  $\mu$ M) alone had no effect on respiration (Fig. 3A). In addition, we also found that digitonin at a concentration of 1  $\mu$ M



**Fig. 2.** Comparison of cellular bioenergetic profiles of UCP1 WT and KO brown adipocytes under non-buffering (A) and free fatty acids buffering (B) conditions, i.e. in the absence or presence of bovine serum albumin, and the proposed working model (C). Under non-buffering condition, free fatty acids accumulate to high concentrations and mask UCP1-mediated leak respiration in cells, as ISO-induced leak respiration was completely independent of UCP1 (A;C left). The unspecific uncoupling effects of free fatty acids can be avoided by adding essentially fatty acid-free BSA to the respiration medium to scavenge extracellular free fatty acids released by lipolysis. The difference in uncoupled respiration between WT and KO cells can be considered UCP1-mediated leak respiration (B;C right). Figure is reprinted with permission from ref. [27].



**Fig. 3.** Validation of the specificity of UCP1 mediated uncoupled respiration by GDP mediated inhibition (A) and siRNA mediated knockdown of UCP1 and ATGL (B). After measuring the ISO-induced leak respiration, brown adipocytes were permeabilized and treated with GDP via injection of digitonin plus or minus GDP. UCP1-mediated leak respiration is gradually and completely inhibited by GDP, while digitonin alone has no effect (A). (B) Functional consequences of UCP1 and ATGL knockdown in brown adipocytes. Reverse transfection of Ucp1 and Atgl siRNA into differentiated brown adipocytes were performed, respectively. Three days after transfection, cellular bioenergetic analysis was conducted. The ISO-induced leak respiration was decreased in siRNA-transfected cells, validating the efficiency of knockdown and the specificity of our assay for UCP1 mediated uncoupled respiration.

did not permeabilize the cells since GDP had no effect on respiration, while 6  $\mu$ M–10  $\mu$ M of digitonin was effective (data not shown). This protocol thus allows us to directly quantify the UCP1-mediated component of uncoupled respiration, without the requirement to use UCP1-KO cells. This refined approach also validated the specificity of our method to quantify UCP1-mediated uncoupled respiration. Uncoupled respiration insensitive to GDP must operate through UCP1-independent mechanisms. The possible pleiotropic effects of GDP on brown adipocyte metabolic rate can be further tested using UCP1-KO cells.

To further validate the specificity of UCP1-mediated uncoupled respiration and establish an assay for gene function in thermogenic adipocytes, we performed reverse transfection of differentiated brite adipocytes with Ucp1 and Atgl siRNA directly in cell culture microplates (Seahorse XF96) as recently demonstrated by Isidor et al. [30], respectively. Three days after transfection, we performed the cellular bioenergetic analysis using our established buffering system. The ISO-induced respiration was decreased in both Ucp1 siRNA and ATGL siRNA-transfected cells, demonstrating the efficiency of knockdown (50%) and the specificity of our assay for UCP1-mediated uncoupled respiration.

Based on our results, it is tempting to speculate that brown adipocytes do not strictly require UCP1 to perform uncoupled respiration but alternatively increase fatty acid flux, even in vivo. However, this seems unlikely, since UCP1 knockout mice are cold intolerant [32]. One possible explanation for this discrepancy between in vitro and in vivo could be that the albumin in blood serves as a fatty acid sink for brown adipocytes in vivo. Another possible explanation is that cultured adipocytes in vitro lose their ability to control the intracellular level of free fatty acid. Indeed, some fatty acid binding proteins (FABPs) such as FABP3 are not expressed in cultured brown and brite adipocytes, but highly expressed in vivo [33]. Whether the lack of fatty acid binding proteins renders cultured adipocytes less efficient in controlling intracellular FFA levels warrants further investigation.

## 6. Conclusions

In summary, UCP1 is regulated physiologically by fatty acids and purine nucleotides. Based on the observations that UCP1 is inherently not leaky in intact brown and brite adipocytes and that FFA mask UCP1 mediated uncoupled respiration, meaningful respirometric measurements of this component using microplate-based respirometry imperatively requires (1) activation of UCP1 and (2) control over free fatty acid levels. The specificity of UCP1-mediated thermogenesis can be further validated by the specific inhibitory effect of GDP on UCP1 activity, which allows quantification of the UCP1-mediated component of uncoupled respiration without using UCP1-KO cells. Gene silencing by siRNA mediated knockdown of UCP1 and ATGL provides proof-of-concept to study the functional relevance of candidate genes in brown and brite adipocytes applying our respirometry protocol.

## Acknowledgements

Our research was supported by grants from the Else Kröner-Fresenius Stiftung (EKFS), and the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) (DFG/KL 973/12-1).

## References

- [1] B. Cannon, J. Nedergaard, Brown adipose tissue: function and physiological significance, *Physiol. Rev.* 84 (2004) 277–359.
- [2] M. Schweiger, R. Schreiber, G. Haemmerle, A. Lass, C. Fledelius, P. Jacobsen, H. Tornqvist, R. Zechner, R. Zimmermann, Adipose triglyceride lipase and hormone-sensitive lipase are the major enzymes in adipose tissue triacylglycerol catabolism, *J. Biol. Chem.* 281 (2006) 40236–40241.
- [3] Y. Li, D. Lasar, T. Fromme, M. Klingenspor, White, brite, and brown adipocytes: the evolution and function of a heater organ in mammals, *Can. J. Zool.* 92 (2013) 615–626.
- [4] M. Klingenspor, T. Fromme, Brown adipose tissue, in: M.E. Symonds (Ed.), *Adipose Tissue Biology*, Springer, New York Dordrecht Heidelberg London, 2012, pp. 39–79.
- [5] Z. Drahotová, E. Honová, P. Hahn, The effect of ATP and carnitine on the endogenous respiration of mitochondria from brown adipose tissue, *Experientia* 24 (1968) 431–432.
- [6] K.J. Hittelman, O. Lindberg, B. Cannon, Oxidative phosphorylation and compartmentation of fatty acid metabolism in brown fat mitochondria, *Eur. J. Biochem. FEBS* 11 (1969) 183–192.
- [7] S.B. Prusiner, B. Cannon, O. Lindberg, Oxidative metabolism in cells isolated from brown adipose tissue. 1. Catecholamine and fatty acid stimulation of respiration, *Eur. J. Biochem. FEBS* 6 (1968) 15–22.
- [8] J. Rafael, H. Ludolph, H. Hohorst, Mitochondria from brown adipose tissue: uncoupling of respiratory chain phosphorylation by long fatty acids and recoupling by guanosine triphosphate, *Hoppe Seyler's Z. Physiol. Chem.* 350 (1969) 1121–1131.
- [9] D.G. Nicholls, R.M. Locke, Thermogenic mechanisms in brown fat, *Physiol. Rev.* 64 (1984) 1–64.
- [10] L. Wojtczak, P. Schonfeld, Effect of fatty acids on energy coupling processes in mitochondria, *Biochim. Biophys. Acta* 1183 (1993) 41–57.
- [11] A. Fedorenko, P.V. Lishko, Y. Kirichok, Mechanism of fatty acid-dependent UCP1 uncoupling in brown fat mitochondria, *Cell* 151 (2012) 400–413.
- [12] D. Ricquier, J.C. Kader, Mitochondrial protein alteration in active brown fat: a sodium dodecyl sulfate-polyacrylamide gel electrophoretic study, *Biochem. Biophys. Res. Commun.* 73 (1976) 577–583.
- [13] V. Azzu, M.D. Brand, The on-off switches of the mitochondrial uncoupling proteins, *Trends Biochem. Sci.* 35 (2010) 298–307.
- [14] D. Nicholls, R. Locke, Cellular mechanisms of heat dissipation, in: L. Girardier, M.J. Stock (Eds.), *Mammalian Thermogenesis*, Springer Netherlands, Dordrecht, 1983, pp. 8–49.
- [15] D. Salvatore, T. Bartha, P.R. Larsen, The guanosine monophosphate reductase gene is conserved in rats and its expression increases rapidly in brown adipose tissue during cold exposure, *J. Biol. Chem.* 273 (1998) 31092–31096.
- [16] J.N. Fain, N. Reed, R. Saperstein, The isolation and metabolism of brown fat cells, *J. Biol. Chem.* 242 (1967) 1887–1894.
- [17] N. Reed, J.N. Fain, Stimulation of respiration in brown fat cells by epinephrine, dibutyl-3',5'-adenosine monophosphate, and m-chloro(carbonyl cyanide) phenylhydrazide, *J. Biol. Chem.* 243 (1968) 2843–2848.
- [18] S.B. Prusiner, B. Cannon, T.M. Ching, O. Lindberg, Oxidative metabolism in cells isolated from brown adipose tissue. 2. Catecholamine regulated respiratory control, *Eur. J. Biochem. FEBS* 7 (1968) 51–57.
- [19] S.W. Cushman, J.J. Heindel, B. Jeanrenaud, Cell-associated nonesterified fatty acid levels and their alteration during lipolysis in the isolated mouse adipose cell, *J. Lipid Res.* 14 (1973) 632–642.
- [20] J. Nedergaard, O. Lindberg, Norepinephrine-stimulated fatty acid release and oxygen consumption in isolated hamster brown-fat cells. Influence of buffers, albumin, insulin and mitochondrial inhibitors, *Eur. J. Biochem. FEBS* 95 (1979) 139–145.
- [21] A. Matthias, K.B. Ohlson, J.M. Fredriksson, A. Jacobsson, J. Nedergaard, B. Cannon, Thermogenic responses in brown fat cells are fully UCP1-dependent. UCP2 or UCP3 do not substitute for UCP1 in adrenergically or fatty acid-induced thermogenesis, *J. Biol. Chem.* 275 (2000) 25073–25081.
- [22] I.G. Shabalina, M. Ost, N. Petrovic, M. Vrbáček, J. Nedergaard, B. Cannon, Uncoupling protein-1 is not leaky, *Biochim. Biophys. Acta BBA Bioenergetics* 1797 (2010) 773–784.
- [23] N. Petrovic, I.G. Shabalina, J.A. Timmons, B. Cannon, J. Nedergaard, Thermogenically competent nonadrenergic recruitment in brown preadipocytes by a PPAR $\gamma$  agonist, *Am. J. Physiol. Endocrinol. Metab.* 295 (2008) E287–E296.
- [24] N. Petrovic, T.B. Walden, I.G. Shabalina, J.A. Timmons, B. Cannon, J. Nedergaard, Chronic peroxisome proliferator-activated receptor  $\gamma$  (PPAR $\gamma$ ) activation of epididymally derived white adipocyte cultures reveals a population of thermogenically competent, UCP1-containing adipocytes molecularly distinct from classic brown adipocytes, *J. Biol. Chem.* 285 (2010) 7153–7164.
- [25] D.A. Ferrick, A. Neilson, C. Beeson, Advances in measuring cellular bioenergetics using extracellular flux, *Drug Discov. Today* 13 (2008) 268–274.
- [26] M. Jastroch, A.S. Divakaruni, S. Mookerjee, J.R. Treberg, M.D. Brand, Mitochondrial proton and electron leaks, *Essays Biochem.* 47 (2010) 53–67.
- [27] Y. Li, T. Fromme, S. Schweizer, T. Schottl, M. Klingenspor, Taking control over intracellular fatty acid levels is essential for the analysis of thermogenic function in cultured primary brown and brite/beige adipocytes, *EMBO Rep.* (2014).
- [28] Y. Li, F. Bolze, T. Fromme, M. Klingenspor, Intrinsic differences in BRITE adipogenesis of primary adipocytes from two different mouse strains, *Biochim. Biophys. Acta* 1841 (2014) 1345–1352.
- [29] J.R. Alvarez-Dominguez, Z. Bai, D. Xu, B. Yuan, K.A. Lo, M.J. Yoon, Y.C. Lim, M. Knoll, N. Slavov, S. Chen, P. Chen, H.F. Lodish, L. Sun, De novo reconstruction of adipose tissue transcriptomes reveals long non-coding RNA regulators of brown adipocyte development, *Cell Metab.* 21 (2015) 764–776.
- [30] M.S. Isidor, S. Winther, A.L. Basse, M.C. Petersen, B. Cannon, J. Nedergaard,

- J.B. Hansen, An siRNA-based method for efficient silencing of gene expression in mature brown adipocytes, *Adipocyte* 5 (2016) 175–185.
- [31] E.T. Chouchani, L. Kazak, M.P. Jedrychowski, G.Z. Lu, B.K. Erickson, J. Szpyt, K.A. Pierce, D. Laznik-Bogoslavski, R. Vetrivelan, C.B. Clish, A.J. Robinson, S.P. Gygi, B.M. Spiegelman, Mitochondrial ROS regulate thermogenic energy expenditure and sulfenylation of UCP1, *Nature* 532 (2016) 112–116.
- [32] S. Enerback, A. Jacobsson, E.M. Simpson, C. Guerra, H. Yamashita, M.E. Harper, L.P. Kozak, Mice lacking mitochondrial uncoupling protein are cold-sensitive but not obese, *Nature* 387 (1997) 90–94.
- [33] L. Vergnes, R. Chin, S.G. Young, K. Reue, Heart-type fatty acid-binding protein is essential for efficient brown adipose tissue fatty acid oxidation and cold tolerance, *J. Biol. Chem.* 286 (2011) 380–390.





## Review

# Targeted mitochondrial uncoupling beyond UCP1 – The fine line between death and metabolic health



Mario Ost <sup>a,\*</sup>, Susanne Keipert <sup>b</sup>, Susanne Klaus <sup>a</sup>

<sup>a</sup> Research Group Physiology of Energy Metabolism, German Institute of Human Nutrition Potsdam-Rehbruecke, 14558, Germany

<sup>b</sup> Helmholtz Diabetes Center, Helmholtz Zentrum München, Neuherberg, 85764, Germany

## ARTICLE INFO

## Article history:

Received 23 September 2016

Received in revised form

7 November 2016

Accepted 13 November 2016

Available online 2 December 2016

## Keywords:

Energy metabolism

Longevity

Mitochondria

Obesity

Protonophore

Uncoupling protein 1

## ABSTRACT

In the early 1930s, the chemical uncoupling agent 2,4-dinitrophenol (DNP) was promoted for the very first time as a powerful and effective weight loss pill but quickly withdrawn from the market due to its lack of tissue-selectivity with resulting dangerous side effects, including hyperthermia and death. Today, novel mitochondria- or tissue-targeted chemical uncouplers with higher safety and therapeutic values are under investigation in order to tackle obesity, diabetes and fatty liver disease. Moreover, in the past 20 years, transgenic mouse models were generated to understand the molecular and metabolic consequences of targeted uncoupling, expressing functional uncoupling protein 1 (UCP1) ectopically in white adipose tissue or skeletal muscle. Similar to the action of chemical mitochondrial uncouplers, UCP1 protein dissipates the proton gradient across the inner mitochondrial membrane, thus allowing maximum activity of the respiratory chain and compensatory increase in oxygen consumption, uncoupled from ATP synthesis. Consequently, targeted mitochondrial uncoupling in adipose tissue and skeletal muscle of UCP1-transgenic mice increased substrate metabolism and ameliorates obesity, hypertriglyceridemia and insulin resistance. Further, muscle-specific decrease in mitochondrial efficiency promotes a cell-autonomous and cell-non-autonomous adaptive metabolic remodeling with increased oxidative stress tolerance. This review provides an overview of novel chemical uncouplers as well as the metabolic consequences and adaptive processes of targeted mitochondrial uncoupling on metabolic health and survival.

© 2016 Elsevier B.V. and Société Française de Biochimie et Biologie Moléculaire (SFBBM). All rights reserved.

## Contents

1. Introduction .....	78
2. Metabolic impact of chemical mitochondrial uncoupling agents .....	78
2.1. DNP (2,4-dinitrophenol) .....	79
2.2. Next generation chemical uncouplers .....	79
2.3. Liver-targeted chemical uncouplers .....	80
3. Targeted expression of UCP1 to white fat depots .....	80
3.1. Metabolic consequences of white adipose tissue-targeted mitochondrial uncoupling .....	81
3.2. Induction of endogenous UCP1 in white fat depots .....	81
4. Targeted expression of UCP1 to skeletal muscle .....	81
4.1. Metabolic consequences of muscle-targeted mitochondrial uncoupling .....	81
4.2. Muscle-targeted mitochondrial uncoupling promotes adaptive metabolic remodeling .....	81

**Abbreviations:** AMPK, AMP-activated protein kinase; aP2, adipocyte protein 2; ATP, adenosine triphosphate; BAT, brown adipose tissue; CRMP, controlled-released mitochondrial protonophore; DNP, 2,4-dinitrophenol; DNPME, DNP-methylethyl; FGF21, fibroblast growth factor 21; NEN, niclosamide ethanolamine salt; OXPHOS, mitochondrial oxidative phosphorylation; RMR, resting metabolic rate; ROS, reactive oxidant species; T2D, type 2 diabetes; TPP, triphenylphosphonium; UCP1, uncoupling protein 1; WAT, white adipose tissue.

\* Corresponding author. Research Group Physiology of Energy Metabolism, German Institute of Human Nutrition Potsdam-Rehbrücke, Nuthetal, 14558, Germany.

E-mail address: [Mario.ost@dife.de](mailto:Mario.ost@dife.de) (M. Ost).

<http://dx.doi.org/10.1016/j.biochi.2016.11.013>

0300-9084/© 2016 Elsevier B.V. and Société Française de Biochimie et Biologie Moléculaire (SFBBM). All rights reserved.

4.3. Muscle-targeted mitochondrial uncoupling delay age-related disease .....	82
5. Conclusion and therapeutic perspectives .....	82
Conflict of interests .....	82
Acknowledgements .....	82
References .....	82

## 1. Introduction

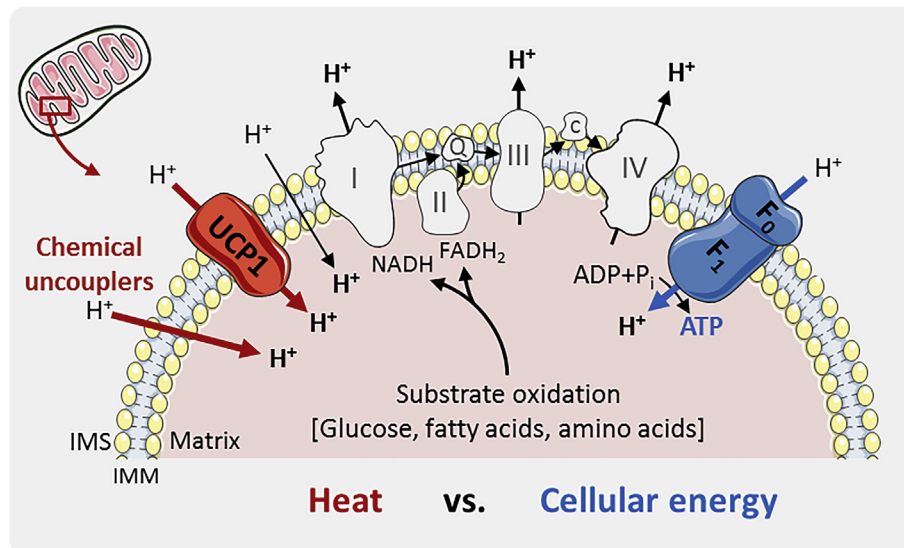
Metabolic diseases such as obesity, hypertriglyceridemia and type 2 diabetes (T2D) have reached an epidemic level globally [1,2]. Metabolic health is closely associated with body weight and whole-body energy balance which can be regulated by the amount of energy intake or energy expenditure. Targeting processes that lead to a reduction in mitochondrial coupling/efficiency could be a promising therapeutic strategy to combat obesity and its comorbidities. The uncoupling protein 1 (UCP1) is the first identified and most studied uncoupling protein, discovered almost 40 years ago [3,4]. Despite the existence of other UCPs such as UCP2 [5] and UCP3 [6,7], only UCP1 seems to mediate energy dissipation as heat for adaptive thermogenesis via functional mitochondrial uncoupling *in vivo* [8,9]. UCP1 is predominantly expressed in brown adipose tissue (BAT) mitochondria and dissipates upon activation the proton gradient across the inner mitochondrial membrane [10], thus uncoupling electron transfer system from ATP synthesis and accelerating mitochondrial oxidative phosphorylation (OXPHOS) in order to maintain ATP homeostasis [11]. Basically, mitochondrial uncoupling refers to a loss of coupling between the mitochondrial inner membrane electrochemical proton gradient and the synthesis of ATP (Fig. 1), thereby releasing energy as heat [12]. Within BAT depots, this metabolic process is called non-shivering thermogenesis which is UCP1-dependent and strongly increased by cold exposure [13,14]. The manipulation of UCP1 activity is an excellent

approach to influence energy expenditure and a natural defense against obesity. However, it is localized in BAT, a tissue which, when not activated by cold induction, represents only a small part of the human body [15–18].

Pharmacological agents that increase metabolic rate by increasing uncoupling of mitochondrial OXPHOS were intensively studied in the past. However, systemic chemical mitochondrial uncoupling agents, such as 2,4-dinitrophenol (DNP) or carbonyl cyanide p-(trifluoromethoxy) phenylhydrazone (FCCP) lack selectivity and have a narrow therapeutic window, largely due to their severe side effects and toxic doses [19]. Instead, tissue-specific and thereby targeted mitochondrial uncoupling has been investigated during the last decades as a powerful strategy to regulate whole-body energy homeostasis and metabolic health for the treatment of obesity and associated metabolic disorders. This review provides an overview of the metabolic consequences and adaptive processes in response to a targeted treatment with chemical uncoupling agents or ectopic overexpression of functional UCP1, as a model of tissue-targeted mitochondrial uncoupling.

## 2. Metabolic impact of chemical mitochondrial uncoupling agents

Apart from dietary and pharmacological interventions affecting satiety or intestinal absorption efficiency to decrease energy intake, increasing energy output through pharmacological uncoupling has



**Fig. 1. Mitochondrial oxidative phosphorylation and uncoupling.** Cellular energetics is efficiently controlled by rate of the electron transfer system (ETS) and oxidative phosphorylation, referred to as the mitochondrial respiratory chain (RC). Substrate muscle metabolism of glucose (glycolysis), fatty acids ( $\beta$ -Oxidation) and amino acids is closely coupled with ATP formation through mitochondrial RC. The primary reducing equivalents of the ETS are nicotinamide adenine dinucleotide (NADH), and flavin adenine dinucleotide (FADH<sub>2</sub>) which are mainly generated by the tricarboxylic acid cycle or during  $\beta$ -Oxidation of fatty acids/acetyl-CoA. Electrons received from NADH or FADH<sub>2</sub> are passed through the series of OXPHOS complexes in the RC, ultimately reducing oxygen to water. This electron flow particular through complex I, III and IV results in pumping of protons from the matrix into the intermembrane space (IMS), generating a membrane potential ( $\Delta\psi_m$ ) and proton motive force that in turn is used to generate ATP from ADP and inorganic phosphate (P<sub>i</sub>) via the ATP synthase (F<sub>1</sub>/F<sub>0</sub>). Predominantly expressed in brown adipose tissue, uncoupling protein 1 (UCP1) dissipates the proton gradient across the inner mitochondrial membrane (IMM), thus uncoupling ETS from ATP synthesis and accelerating mitochondrial RC activity in order to maintain energy homeostasis [11]. During mitochondrial uncoupling the energy amount stored in the proton gradient is released as heat. See text for details on chemical uncouplers. This figure was created using Servier Medical Art (<http://www.servier.com>).

been proposed as a weight-loss therapy [20]. Below we review well-known and novel chemical uncoupling agents and discuss their relevance for metabolic health and treatment of human metabolic disease.

### 2.1. DNP (2,4-dinitrophenol)

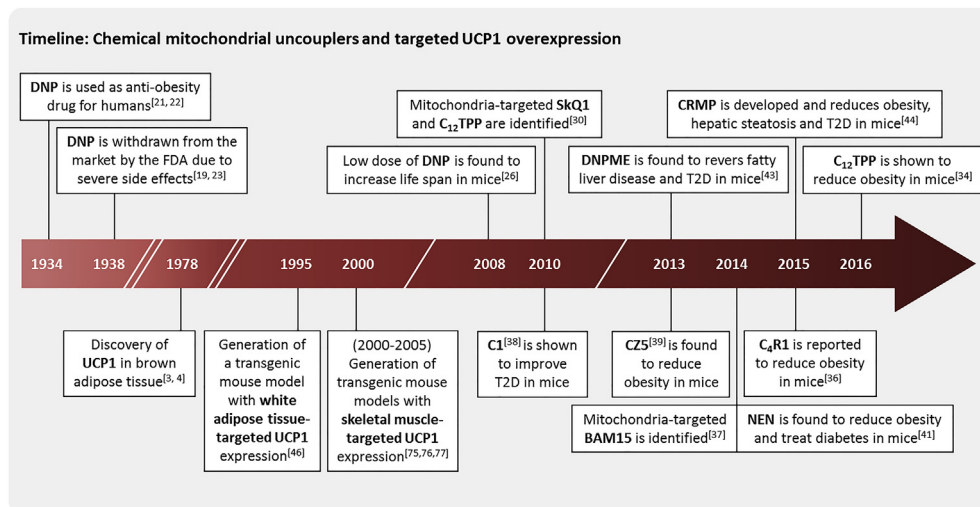
The first and best-studied example is the artificial uncoupler DNP, a lipid-soluble weak acid which acts as a chemical protonophore and allows protons to leak across the inner mitochondrial membrane [21], mimicking the uncoupling effect of activated UCPs. In the 1930s, DNP was widely used to treat obesity [22]. Nevertheless, because at high doses nonspecific uncoupling in all tissues causes dangerous side effects including hyperthermia and death [23], DNP was withdrawn from the market by the US Food and Drug Administration (FDA) in 1938. Case reports demonstrated that an acute administration of 20–50 mg per kilogram of body weight in humans can be lethal [24]. Nevertheless, in 2015 in the UK, a substantial increase was reported in clinical presentations with toxicity and associated high mortality caused by exposure to DNP [25]. In contrast, recent studies demonstrate that long-term treatment with low doses of DNP protects against diet-induced obesity, improves insulin sensitivity and increases lifespan in mice [26,27]. Because of the strong effects on body weight in humans and survival in mice, the mechanism of action of DNP remains under investigation as a potential approach for the treatment of obesity and associated metabolic disorders (Fig. 2). In 2006, Murphy and colleagues developed a mitochondrial-targeted form of DNP, by coupling it to the lipophilic triphenylphosphonium (TPP) cation, which accumulates within mitochondria driven by the membrane potential [28]. They found that MitoDNP was extensively taken up by mitochondria, however no increase in uncoupling could be observed. Six years later Chalmers et al. developed a compound called MitoPhotoDNP, a mitochondria-targeted photoactivated protonophore [29]. Comparable with MitoDNP, it is targeted to mitochondrion by TPP, but releases DNP only in response to directed irradiation with UV light. Indeed, MitoPhotoDNP led to the selective uncoupling of individual

and/or several mitochondria within a cell when used in conjunction with fluorescence imaging. Thus, MitoPhotoDNP represents a promising tool to elucidate the effects of mitochondrial uncoupling on cellular metabolism *in vitro* which is of great importance in the development of less toxic protonophore.

### 2.2. Next generation chemical uncouplers

Several novel mitochondrial- or tissue-targeted chemical uncouplers with a higher safety and therapeutic value were developed and investigated in the last couple of years (Fig. 2). In 2010, the group of Vladimir P. Skulachev synthesized penetrating cation/fatty acid anion pairs as mitochondria-targeted protonophore. While initially searching for mitochondria-targeted antioxidants, they discovered that the synthesized plastoquinone derivatives SkQ1 (10-(6-plastoquinonyl) decyltriphenylphosphonium) and C<sub>12</sub>TPP (dodecyltriphenylphosphonium) potentiated the fatty acid-induced uncoupling of respiration and OXPHOS in isolated rat-liver mitochondria [30]. SkQ1 was further investigated as mitochondria-targeted antioxidant for potential treatment of various age-related diseases, such as Alzheimer's disease [31] or retinopathy [32,33]. Furthermore, a recent study demonstrated that the mitochondrial-targeted C<sub>12</sub>TPP effectively increased oxygen consumption in isolated brown-fat mitochondria, independent of UCP1, and abolished diet-induced obesity in mice by reducing food intake, increasing the resting metabolic rate and overall fatty acid oxidation [34].

In addition, Rhodamine 19 butyl ester (C<sub>4</sub>R1), a short-chain alkyl derivative of Rhodamine 19, was found to decrease the membrane potential and stimulate respiration of isolated liver mitochondria as well as reduce oxidative stress induced by brain ischemia and reperfusion in rats [35]. In addition, similar to the *in vivo* effects of C<sub>12</sub>TPP, the penetrating cation C<sub>4</sub>R1 effectively reduced body weight and fat mass of obese mice fed a high-fat diet [36]. Thus, both C<sub>12</sub>TPP and C<sub>4</sub>R1 are considered as novel promising candidates for mild mitochondrial uncoupling anti-obesity drugs. With the overall aim to identify mitochondrial-targeted uncouplers that lack



**Fig. 2. Timeline of chemical uncoupling agents and targeted UCP1 overexpression.** This illustrates the increasing interest in mitochondrial uncoupling for the treatment of obesity and associated metabolic disorders during the past 20 years. Of note, in parallel the re-evaluation of the role of brown adipose tissue (BAT) for the treatment of obesity took place, supported by observations using positron emission tomography-computed tomography (PET/CT) scanning that revealed the presence of BAT in adult humans [15–18]. Moreover, in the last decade research focused intensively on the induction of endogenous UCP1 in white adipose tissue depots, also called the “browning” of white fat, and the effects on metabolic health [56,62]. Abbreviations: BAM15, N<sup>5</sup>,N<sup>6</sup>-bis(2-Fluorophenyl)-oxadiazolo-pyrazine-5,6-diamine [37]; CRMP, controlled-released mitochondrial protonophore [44]; C<sub>4</sub>R1, Rhodamine 19 butyl ester [36]; C1, nomenclature not defined [38]; CZ5, nomenclature not defined [39]; C<sub>12</sub>TPP, dodecyltriphenylphosphonium [34]; FDA, US Food and Drug Administration; DNP, 2,4-dinitrophenol; DNPME, DNP-methylethyl [43]; NEN, niclosamide ethanolamine salt [41]; SkQ1, 10-(6-plastoquinonyl) decyltriphenylphosphonium [30]; UCP1, uncoupling protein 1. See text for further details.

off-target activity at the plasma membrane, the lipophilic weak acid named BAM15 was recently identified by a small molecule library and bioenergetics screening approach [37]. Interestingly, the authors could show that although BAM15 is less cytotoxic because it does not depolarize the plasma membrane, it still effectively uncouples OXPHOS in L6 myoblast mitochondria *in vitro*. Furthermore, first *in vivo* experiments demonstrated that BAM15 protects mice from acute renal ischemic-reperfusion injury, whereas effects on metabolic diseases have not been addressed so far.

However, the above-mentioned uncoupling agents are not selective for particular mitochondria and the challenge still is to find a way to deliver those molecules to mitochondria within individual tissues or cells. Two additional novel mitochondrial uncouplers, named C1 and CZ5, were recently uncovered performing a high-throughput screening assay for modulators of mitochondrial membrane potential [38,39]. The small-molecule compound C1 increased fat oxidation and activity of key cellular energy sensor AMP-activated protein kinase (AMPK) [40] acutely 2 h after intraperitoneal injection specifically in liver of lean mice as well as reduced hyperglycemia and plasma fatty acids in diabetic *db/db* mice after long-term oral administration for 4 weeks [38]. Remarkably, the compound CZ5 was described to act as a cell type-specific uncoupler that only targets skeletal muscle and adipose tissue, but not liver [39]. Chronic orally administrated CZ5 for 25 days ameliorated diet-induced obesity via both increased energy expenditure and suppressed food intake. Besides, CZ5 treatment improved glucose and lipid metabolism *in vivo*, accompanied by an activated AMPK phosphorylation in targeted white fat depot and skeletal muscle.

Finally, a recent study by Tao et al. could show that treatment with niclosamide ethanolamine salt (NEN) uncouples mammalian mitochondria at upper nanomolar concentrations and increases energy expenditure and lipid metabolism in mice [41]. Interestingly, NEN represents a salt form of niclosamide (5-chloro-salicyl-(2-chloro-4-nitro) anilide), an anthelmintic drug approved by the FDA for treating intestinal infections of tapeworms [42]. Remarkably, oral NEN efficaciously prevented and treated hepatic steatosis and insulin resistance in high-fat-fed mice and improved glycemic control accompanied by delayed disease progression in *db/db* mice. Moreover, NEN activated AMPK in a dose- and time-dependent

manner in liver, but not in muscle and adipose tissue of treated mice. Thus, the authors concluded that liver is a direct target of NEN treatment and AMPK-activation as key mechanism on promoting lipid oxidation.

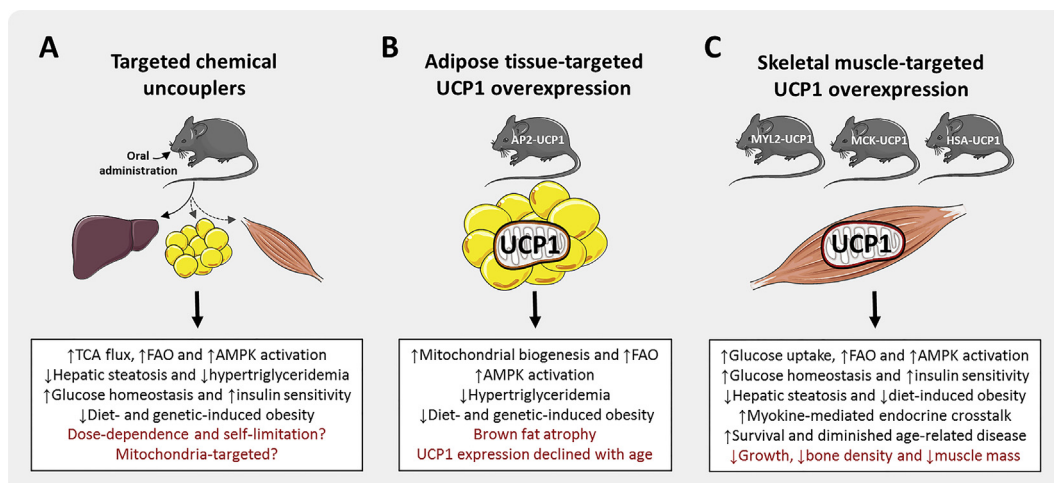
### 2.3. Liver-targeted chemical uncouplers

A study has recently demonstrated that a functionally liver-targeted derivative of DNP, DNP-methyl ether (DNPME), reversed a high-fat diet-induced hypertriglyceridemia, hepatic steatosis, and whole-body insulin resistance in rats without inducing hyperthermia or associated hepatic or systemic toxicities [43]. Here, the authors hypothesized that derivatives of DNP, such as DNPME, would be preferentially metabolized by the hepatic cytochrome P450 system to yield the active protonophore DNP in hepatocytes as primary target cells. Interestingly, they reported that the 50% lethal dose (LD<sub>50</sub>) of DNPME is almost 10-fold higher than that of classic DNP. In a follow-up study, the same group developed a controlled-release oral formulation of DNP, called CRMP (controlled-release mitochondrial protonophore) in order to further improve the safety and efficacy of DNP, consequently increasing the therapeutic window of this agent [44]. Using a polymer hydroxypropylcellulose/ethylcellulose coating system they generated a novel DNP version with lower peak plasma levels and sustained-release pharmacokinetics. In rat models of diet-induced or genetic obesity, daily CRMP administration reversed hepatic steatosis, insulin resistance, T2D, steatohepatitis, and liver fibrosis without detectable toxicity [44].

Altogether, tissue-targeted chemical mitochondrial uncoupling agents provide an elegant strategy to combat obesity and associated disorders, although the issue of dose-dependence and self-limitation remains an important open question (Fig. 3A). Thus, regulation and activation of endogenous proteins with uncoupling action, such as UCP1 or the adenine nucleotide translocase (ANT) [45], may provide an alternative strategy for dose-independent and self-limiting tissue-restricted mitochondrial uncoupling.

### 3. Targeted expression of UCP1 to white fat depots

White adipose tissue (WAT) plays an important role as an



**Fig. 3. Overview on metabolic effects of chemical mitochondrial uncouplers and targeted UCP1 overexpression.** (A) Tissue-targeted chemical mitochondrial uncoupling agents provide an additional elegant strategy to combat metabolic syndrome and associated disorders. However, the issues of target tissue, dose-dependence and self-limitation remain important open questions. (B)+(C) Studies using transgenic mice with targeted UCP1 overexpression uncovered key molecular mechanisms how mitochondrial uncoupling affects energy metabolism and metabolic health *in vivo*. Abbreviations: AMPK, AMP-activated protein kinase; AP2, adipocyte protein 2; FAO, fatty acid oxidation; HSA, human skeletal actin; MCK, muscle creatine kinase; MYL2, myosin light chain 2, TCA, tricarboxylic acid cycle; UCP1, uncoupling protein 1. See text for further details. This figure was created using Servier Medical Art (<http://www.servier.com>).



endocrine regulator and is involved in whole body glucose as well as energy homeostasis. However, its major role is the control of systemic fatty acids levels and the storage of metabolic energy, thus the opposite of energy burning thermogenic brown adipose tissue.

### 3.1. Metabolic consequences of white adipose tissue-targeted mitochondrial uncoupling

In 1995, Kopecky and coworkers established a transgenic mouse model (AP2-UCP1 mice) expressing UCP1 in WAT [46–48]. In these mice, the fat-specific AP2 promoter was used to drive expression of UCP1, resulting in enhanced protein expression of UCP1 in both WAT and BAT. Notably, whereas the total protein amount of transgenic UCP1 in WAT of adult mice did not exceed 2% of native UCP1 found in bona fide BAT depots, this was still sufficient to uncouple OXPHOS in WAT adipocytes [49]. Moreover, AP2-UCP1 mice showed resistance to diet induced obesity, which is consistent with the hypothesis, that thermogenesis from elevated expression of UCP1 reduced adiposity (Fig. 3B). In contrast, the transgene particular led to BAT atrophy which impaired thermogenic mechanisms for protecting body temperature [50]. In fact, it was shown that ectopic WAT-targeted UCP1 overexpression particularly activated AMPK and mitochondrial biogenesis in unilocular white adipocytes [51,52]. Transgenic UCP1 also reduced lipogenesis and modulated lipolysis and hormonal control of lipid metabolism [53,54] but resulted in only a marginal stimulation of RMR [50]. Moreover, UCP1 expression in WAT of AP2-UCP1 mice decreases with age [52], suggesting a posttranscriptional control of the ectopic UCP1 expression or an elimination of UCP1-containing adipocytes with time. Finally, Yamada et al. could demonstrate that ectopic expression of very low levels of UCP1 in epididymal WAT through injection of a UCP1 adenovirus vector reversed both insulin and leptin resistance, improved glucose tolerance and decreased food intake in both diet-induced and genetically obese mouse models [55].

### 3.2. Induction of endogenous UCP1 in white fat depots

Nowadays research focus intensively on the induction of endogenous UCP1 in WAT depots, also called the “browning” of white fat, which is characterized by the infiltration or trans-differentiation of so called beige/brite fat cells within white adipose tissue depots [56]. It was already reported 30 years ago that cold stimulus induces multilocular, UCP1-positive fat cells within certain WAT depots [57,58]. Today we know that endogenous UCP1 can be induced in WAT not only by cold exposure, but also in response to different pharmacological and nutritional stimuli [59] as well as by secreted endogenous factors [60,61]. In line with the results of ectopic expression of UCP1 in WAT, the induction of WAT browning has been associated with improved metabolic health [62]. However, while it was demonstrated that UCP1 in brite/beige adipose tissue mitochondria is indeed thermogenically functional [63], we could recently demonstrate a clear dissociation of WAT browning from obesity resistance and improved glycemic control and insulin sensitivity [61]. Thus, the precise physiological relevance of WAT remodeling, including UCP1 expression, still remains hotly debated [64].

## 4. Targeted expression of UCP1 to skeletal muscle

Recent studies implicated that BAT and muscle cells, but not WAT cells, differentiate from a common precursor [65], suggesting that BAT cells are more similar to muscle cells than to white adipocytes. Moreover, skeletal muscle represents a plastic and highly metabolic active organ that constitutes up to 40% of total body mass

in mammals and is the major contributor to RMR and total energy expenditure [66–68]. Therefore, skeletal muscle represents a predominant site of glucose disposal and plays a crucial role in glycemic control [69] as well as uptake and utilization of plasma lipoprotein-derived and free fatty acid [70,71]. In contrast, chronic metabolic disorders such as obesity and T2D are closely associated to impaired muscle mitochondrial function [72,73] and an involvement of skeletal muscle in mitochondrial dysfunction-associated diseases is frequent [74]. Thus, proper muscle mitochondrial performance is tightly connected to metabolic health and the question came up whether skeletal muscle represents another suitable tissue for targeted mitochondrial uncoupling?

### 4.1. Metabolic consequences of muscle-targeted mitochondrial uncoupling

Almost 16 years ago, the first transgenic mouse model with ectopic muscle-targeted UCP1 overexpression (mUCP1-Tg) driven by the rat myosin light chain 2 (MYL2) promoter was described by Li et al. demonstrating that low muscle UCP1 expression doubled muscle oxygen consumption without affecting thermoregulation [75]. Moreover, transgenic mice with muscle-targeted UCP1 overexpression under control of the mouse muscle creatine kinase (MCK) [76] or the human skeletal actin (HSA) [77] promoter were generated. Again, while mUCP1-Tg mice indeed show an increased mitochondrial uncoupling and a reduced muscle mitochondrial OXPHOS capacity [78], body temperature was normal or rather decreased with declining ambient temperature [77]. mUCP1-Tg mice display increased energy expenditure [77] and a whole range of metabolic improvements such as an increased metabolic flexibility [79], muscle glucose uptake and fatty acid oxidation [80,81], as well as an increased insulin sensitivity accompanied by decreased insulin levels, especially under high-fat diet-feeding. Interestingly, this increased insulin sensitivity is independent of diet and body fat accumulation suggesting a dissociation of obesity and insulin resistance in mUCP1-Tg mice [79,82,83]. Surprisingly, not only muscle is affected in mUCP1-TG mice. Furthermore they show an increased glucose uptake [84], an augmented lipid metabolism [79] as well as an induction of beige/brite adipocytes in WAT depots [85], suggesting an endocrine role of skeletal muscle uncoupling. Overall, improved glucose tolerance accompanied by an increased insulin sensitivity and muscle glucose uptake seems to be the most robust metabolic phenotype of mUCP1-Tg models (Fig. 3C).

### 4.2. Muscle-targeted mitochondrial uncoupling promotes adaptive metabolic remodeling

Glycolytic and oxidative metabolic processes are rapidly activated to maintain cellular energy homeostasis in skeletal muscle. In line with the above-mentioned liver- and WAT-targeted mitochondrial uncoupling approach it was shown that muscle-restricted mitochondrial uncoupling also led to an increased AMPK activity [84,86]. Notably, AMPK $\alpha$ 2 activity, but not AMPK $\alpha$ 1 was highly induced in muscle of mUCP1-Tg mice, and loss of active AMPK $\alpha$ 2 promoted a severe reduction of overall muscle integrity together with a highly diminished physical activity tolerance and impaired mitochondrial biogenesis [80]. This revives the significance of AMPK for regulating cellular plasticity in response to chronic decreased mitochondrial energy efficiency. Interestingly, enhancing AMPK activity in brown adipocytes also increased BAT activity [87]. Thus, targeting AMPK as a key mediator of a cell-autonomous adaptive response holds therapeutic potential for the treatment of obesity and associated metabolic disorders.

Furthermore, muscle-targeted mitochondrial uncoupling promotes cell-non-autonomous effects and endocrine crosstalk via

secretion of myokines. Treatment of mouse C2C12 muscle cells *in vitro* with the chemical uncoupler FCCP resulted in a strong induction of integrated stress response (ISR) as well as fibroblast growth factor 21 (FGF21) as myokine [85,88]. Notably, muscle mitochondrial uncoupling induces FGF21, which was previously described as enhancer of WAT browning [89], in skeletal muscle of mUCP1-Tg mice *in vivo* [85]. Indeed, using mUCP1-Tg/FGF21-knockout mice, we recently demonstrated that cell-non-autonomous WAT browning and metabolic remodeling is fully FGF21 dependent [61]. Remarkably, the cell-autonomous muscle metabolic adaptation and obesity resistance was independent of FGF21 action as a myokine. With regard to metabolic health and aging it was shown that transgenic overexpression of FGF21 in liver markedly extends lifespan in mice probably in an auto-/paracrine manner by blunting the growth hormone/insulin-like growth factor-1 signaling pathway [90]. Furthermore, a recent study could demonstrate that serum FGF21 levels in humans are related to BAT activity [91]. However, to date, the ultimate contribution of FGF21 to a cell-autonomous and cell-non-autonomous response on effects of targeted mitochondrial uncoupling and survival remains to be elucidated.

#### 4.3. Muscle-targeted mitochondrial uncoupling delay age-related disease

Remarkably, two independent studies found that skeletal muscle-targeted respiratory uncoupling promotes survival [82,86] and diminishes age-related diseases, such as atherosclerosis, vascular disease and blood pressure as well as diabetes [86,92]. Interestingly, markers of lipid-oxidative stress levels were highly induced in muscle of mUCP1-Tg mice independent of the diet-feeding regime [93]. In addition, increased activity of endogenous antioxidant defense enzymes such as catalase or superoxide dismutase (total SOD) provides evidence for an elevated rather than reduced formation of reactive oxidant species (ROS) in muscle of mUCP1-Tg mice. Mitochondria are a major site of cellular ROS production [94] which by itself represents an important mediator of mitochondrial stress signaling to promote cellular adaptation [95]. In line with that, glutathione metabolism was induced followed muscle-targeted ectopic mitochondrial uncoupling [96,97] which fits with the induced serine and glycine biosynthetic pathway [96], as both amino acids are important precursors for glutathione biosynthesis [98].

Initially, a “rate of living” hypothesis has been proposed by Pearl in the late 1920s, predicting that increased energy metabolism would increase the ROS production and thus reduce life span [99]. While 40 years later Peter Mitchell’s chemi-osmotic hypothesis provided the basis for understanding the actual process of OXPHOS and energy/ATP synthesis [100], it took another 40 years to reveal that an inefficiency of the respiratory chain through mitochondrial uncoupling might increase longevity by reducing ROS generation, despite increased energy expenditure [101]. This so called “uncoupling to survive” hypothesis was supported by the study of Speakman et al., showing that mice with highest metabolic intensity and mitochondrial uncoupling live longer than littermates with lower metabolic rate [102]. However, whether mild mitochondrial uncoupling under physiological conditions indeed plays a role as alleviator of oxidative damage remains unclear [103]. Although mitochondrial function and increased oxidative stress are usually associated with aging [94], a ROS-induced cellular stress adaption through an increased endogenous antioxidant defense system is in line with the concept of “mitohormesis”, suggesting a link between mild oxidative stress and enhanced cellular function [104]. To date, mitohormesis has been best studied in *Caenorhabditis elegans* (*C. elegans*) as a model for neurologic and metabolic diseases

[105,106]. In addition, a *Drosophila* model of mild muscle mitochondrial distress showed preserved mitochondrial and muscle function during aging and a prolonged lifespan [107]. Thus, the significance of the “uncoupling to survive” hypothesis related to longevity by uncoupling-mediated reduced ROS formation [101] should be re-evaluated taking into account the concept of “mitohormesis” based on the *in vivo* mitochondrial adaptation in response to muscle-targeted ectopic mitochondrial uncoupling.

Altogether, despite using different promoters, all mUCP1-Tg mice display an overall improved metabolic phenotype of increased insulin sensitivity, reduced obesity and increased survival (Fig. 3C). However, the molecular mechanisms behind the phenotype are quite complex and not fully understood so far.

## 5. Conclusion and therapeutic perspectives

While discussing potential therapeutic targets for obesity, T2D and fatty liver disease strategies of life style change such as dietary restriction and regular exercise programs should always be kept in mind for improvement of metabolic health status. Nevertheless, if proved to be safe and effective, tissue-targeted chemical mitochondrial uncoupling agents still provide an additional therapeutic strategy to combat metabolic syndrome and associated disorders.

Studies using transgenic mice with targeted UCP1 overexpression uncovered key molecular mechanisms how mitochondrial uncoupling affects energy metabolism and metabolic health *in vivo*. Thereby, it was proven that AMPK plays a crucial role as a housekeeper for mitochondrial function to maintain energy homeostasis and cellular integrity. Together with mitochondrial uncoupling-induced endocrine crosstalk via secretion of cytokines, such as FGF21, this potentially could open up new avenues of investigations that may help to understand how a specific target tissue is sufficient to reprogram and tune the metabolic health of the whole organism. It is worth mentioning that the first human mitochondrial disease discovered around 50 years ago, Luft’s disease, leads to a muscle atrophy due to increased uncoupled mitochondrial oxidative phosphorylation and energy depletion within skeletal muscle which is also affecting whole-body energy metabolism [108]. Thereby, Luft and colleagues described the first example how a single dysfunctional organelle within one specific tissue can affect the whole organism. However, our knowledge regarding the integrated signaling network of cell plasticity remains rudimentary and further studies are required to enlarge our understanding.

Overall, there is a still increasing interest in mitochondrial uncoupling during the past 20 years [109], and we are just half-way on our journey to discover a safe “polypill” that treats metabolic disorders such as obesity and T2D. Whether the light at the end of the tunnel will be a targeted chemical mitochondrial uncoupler or a train remains to be unknown.

## Conflict of interests

The authors declare no conflict of interest.

## Acknowledgements

Skeletal muscle related research of our group is supported by the Leibniz Society (SAW-2013-FBN-3).

## References

- [1] W.P. James, WHO recognition of the global obesity epidemic, *Int. J. Obes.* 32 (Suppl 7) (2008) S120–S126.

- [2] L. Chen, D.J. Magliano, P.Z. Zimmet, The worldwide epidemiology of type 2 diabetes mellitus—present and future perspectives, *Nat. Rev. Endocrinol.* 8 (2012) 228–236.
- [3] G.M. Heaton, R.J. Wagenvoort, A. Kemp Jr., D.G. Nicholls, Brown-adipose-tissue mitochondria: photoaffinity labelling of the regulatory site of energy dissipation, *Eur. J. Biochem. FEBS* 82 (1978) 515–521.
- [4] D. Ricquier, J.C. Kader, Mitochondrial protein alteration in active brown fat: a soidum dodecyl sulfate-polyacrylamide gel electrophoretic study, *Biochem. Biophys. Res. Commun.* 73 (1976) 577–583.
- [5] C. Fleury, M. Neverova, S. Collins, S. Raimbault, O. Champigny, C. Levi-Meyrueis, F. Bouillaud, M.F. Seldin, R.S. Surwit, D. Ricquier, C.H. Warden, Uncoupling protein-2: a novel gene linked to obesity and hyperinsulinemia, *Nat. Genet.* 15 (1997) 269–272.
- [6] O. Boss, S. Samec, A. Paoloni-Giacobino, C. Rossier, A. Dulloo, J. Seydoux, P. Muzzin, J.P. Giacobino, Uncoupling protein-3: a new member of the mitochondrial carrier family with tissue-specific expression, *FEBS Lett.* 408 (1997) 39–42.
- [7] A. Vidal-Puig, G. Solanes, D. Grujic, J.S. Flier, B.B. Lowell, UCP3: an uncoupling protein homologue expressed preferentially and abundantly in skeletal muscle and brown adipose tissue, *Biochem. Biophys. Res. Commun.* 235 (1997) 79–82.
- [8] V. Golozubova, E. Hohtola, A. Matthias, A. Jacobsson, B. Cannon, J. Nedergaard, Only UCP1 can mediate adaptive nonshivering thermogenesis in the cold, *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* 15 (2001) 2048–2050.
- [9] J. Nedergaard, V. Golozubova, A. Matthias, A. Asadi, A. Jacobsson, B. Cannon, UCP1: the only protein able to mediate adaptive non-shivering thermogenesis and metabolic inefficiency, *Biochim. Biophys. Acta* 1504 (2001) 82–106.
- [10] I.G. Shabalina, M. Ost, N. Petrovic, M. Vrbacky, J. Nedergaard, B. Cannon, Uncoupling protein-1 is not leaky, *Biochim. Biophys. Acta* 1797 (2010) 773–784.
- [11] D.G. Nicholls, R.M. Locke, Thermogenic mechanisms in brown fat, *Physiol. Rev.* 64 (1984) 1–64.
- [12] M. Jastroch, A.S. Divakaruni, S. Mookerjee, J.R. Treberg, M.D. Brand, Mitochondrial proton and electron leaks, *Essays Biochem.* 47 (2010) 53–67.
- [13] M. Klingenspor, Cold-induced recruitment of brown adipose tissue thermogenesis, *Exp. Physiol.* 88 (2003) 141–148.
- [14] S. Klaus, A. Seivert, S. Boeuf, Effect of the beta(3)-adrenergic agonist C316,243 on functional differentiation of white and brown adipocytes in primary cell culture, *Biochim. Biophys. Acta* 1539 (2001) 85–92.
- [15] W.D. van Marken Lichtenbelt, J.W. Vanhommerig, N.M. Smulders, J.M. Drossaerts, G.J. Kemerink, N.D. Bouvy, P. Schrauwen, G.J. Teule, Cold-activated brown adipose tissue in healthy men, *N. Engl. J. Med.* 360 (2009) 1500–1508.
- [16] K.A. Virtanen, M.E. Lidell, J. Orava, M. Heglind, R. Westergren, T. Niemi, M. Taittonen, J. Laine, N.J. Savisto, S. Enerback, P. Nuutila, Functional brown adipose tissue in healthy adults, *N. Engl. J. Med.* 360 (2009) 1518–1525.
- [17] A.M. Cypess, S. Lehman, G. Williams, I. Tal, D. Rodman, A.B. Goldfine, F.C. Kuo, E.L. Palmer, Y.H. Tseng, A. Doria, G.M. Kolodny, C.R. Kahn, Identification and importance of brown adipose tissue in adult humans, *N. Engl. J. Med.* 360 (2009) 1509–1517.
- [18] J. Nedergaard, T. Bengtsson, B. Cannon, Unexpected evidence for active brown adipose tissue in adult humans, *Am. J. Physiol. Endocrinol. Metab.* 293 (2007) E444–E452.
- [19] J. Parascandola, Dinitrophenol and bioenergetics: an historical perspective, *Mol. Cell. Biochem.* 5 (1974) 69–77.
- [20] J.A. Harper, K. Dickinson, M.D. Brand, Mitochondrial uncoupling as a target for drug development for the treatment of obesity, *Obes. Rev. Off. J. Int. Assoc. Study Obes.* 2 (2001) 255–265.
- [21] D.M. Dunlop, The use of 2,4-dinitrophenol as a metabolic stimulant, *Br. Med. J.* 1 (1934) 524–527.
- [22] M.L. Tainter, W.C. Cutting, A.B. Stockton, Use of dinitrophenol in nutritional disorders: a critical survey of clinical results, *Am. J. Public Health Nat. Health* 24 (1934) 1045–1053.
- [23] J. Grundlingh, P.I. Dargan, M. El-Zanfaly, D.M. Wood, 2,4-dinitrophenol (DNP): a weight loss agent with significant acute toxicity and risk of death, *J. Med. Toxicol.* 7 (2011) 205–212.
- [24] A.L. Hsiao, K.A. Santucci, P. Seo-Mayer, M.R. Mariappan, M.E. Hodsdon, K.J. Banasiak, C.R. Baum, Pediatric fatality following ingestion of dinitrophenol: postmortem identification of a “dietary supplement”, *Clin. Toxicol. Phila* 43 (2005) 281–285.
- [25] A. Kamour, N. George, D. Gwynnette, G. Cooper, D. Lupton, M. Eddleston, J.P. Thompson, J.A. Vale, H.K. Thanacoody, S. Hill, S.H. Thomas, Increasing frequency of severe clinical toxicity after use of 2,4-dinitrophenol in the UK: a report from the National Poisons Information Service, *Emerg. Med. J.* 32 (2015) 383–386.
- [26] C.C. Caldeira da Silva, F.M. Cerqueira, L.F. Barbosa, M.H. Medeiros, A.J. Kowaltowski, Mild mitochondrial uncoupling in mice affects energy metabolism, redox balance and longevity, *Aging Cell* 7 (2008) 552–560.
- [27] M. Goldhof, C. Xiao, T. Chanturiya, W. Jou, O. Gavriloiva, M.L. Reitman, The chemical uncoupler 2,4-dinitrophenol (DNP) protects against diet-induced obesity and improves energy homeostasis in mice at thermoneutrality, *J. Biol. Chem.* 289 (2014) 19341–19350.
- [28] F.H. Blaikie, S.E. Brown, L.M. Samuelsson, M.D. Brand, R.A. Smith, M.P. Murphy, Targeting dinitrophenol to mitochondria: limitations to the development of a self-limiting mitochondrial protonophore, *Biosci. Rep.* 26 (2006) 231–243.
- [29] S. Chalmers, S.T. Caldwell, C. Quin, T.A. Prime, A.M. James, A.G. Cairns, M.P. Murphy, J.G. McCarron, R.C. Hartley, Selective uncoupling of individual mitochondria within a cell using a mitochondria-targeted photoactivated protonophore, *J. Am. Chem. Soc.* 134 (2012) 758–761.
- [30] F.F. Severin, Severina II, Y.N. Antonenko, T.I. Rokitskaya, D.A. Cherepanov, E.N. Mokhova, M.Y. Vyssokikh, A.V. Pustovidko, O.V. Markova, L.S. Yaguzhinsky, G.A. Korshunova, N.V. Sumbatyan, M.V. Skulachev, V.P. Skulachev, Penetrating cation/fatty acid anion pair as a mitochondria-targeted protonophore, *Proc. Natl. Acad. Sci. U. S. A.* 107 (2010) 663–668.
- [31] N.A. Stefanova, N.A. Muraleva, V.P. Skulachev, N.G. Kolosova, Alzheimer's disease-like pathology in senescence-accelerated OXYS rats can be partially retarded with mitochondria-targeted antioxidant SkQ1, *J. Alzheimers Dis.* 38 (2014) 681–694.
- [32] V.B. Saprunova, M.A. Lelekova, N.G. Kolosova, L.E. Bakeeva, SkQ1 slows development of age-dependent destructive processes in retina and vascular layer of eyes of wistar and OXYS rats, *Biochem. Mosc* 77 (2012) 648–658.
- [33] N.A. Muraleva, O.S. Kozhevnikova, A.A. Zhdankina, N.A. Stefanova, T.V. Karamysheva, A.Z. Fursova, N.G. Kolosova, The mitochondria-targeted antioxidant SkQ1 restores alphaB-crystallin expression and protects against AMD-like retinopathy in OXYS rats, *Cell Cycle* 13 (2014) 3499–3505.
- [34] A.V. Kalinovich, C.L. Mattsson, M.R. Youssef, N. Petrovic, M. Ost, V.P. Skulachev, I.G. Shabalina, Mitochondria-targeted dodecyl-triphenylphosphonium (C12TPP) combats high-fat diet-induced obesity in mice, *Int. J. Obes.* (2016).
- [35] L.S. Khaилоva, D.N. Silachev, T.I. Rokitskaya, A.V. Avetisyan, K.G. Lyamsaev, Severina II, T.M. Ilyasova, M.V. Gulyaev, V.I. Dedukhova, T.A. Trendeleva, E.Y. Plotnikov, R.A. Zvyagilskaya, B.V. Chernyak, D.B. Zorov, Y.N. Antonenko, V.P. Skulachev, A short-chain alkyl derivative of Rhodamine 19 acts as a mild uncoupler of mitochondria and a neuroprotector, *Biochim. Biophys. Acta* 1837 (2014) 1739–1747.
- [36] A.V. Kalinovich, I.G. Shabalina, Novel mitochondrial cationic uncoupler C4R1 is an effective treatment for combating obesity in mice, *Biochem. Mosc* 80 (2015) 620–628.
- [37] B.M. Kenwood, J.L. Weaver, A. Bajwa, I.K. Poon, F.L. Byrne, B.A. Murrow, J.A. Calderone, L. Huang, A.S. Divakaruni, J.L. Tomsig, K. Okabe, R.H. Lo, G. Cameron Coleman, L. Columbus, Z. Yan, J.J. Saucerman, J.S. Smith, J.W. Holmes, K.R. Lynch, K.S. Ravichandran, S. Uchiyama, W.L. Santos, G.W. Rogers, M.D. Okusa, D.A. Bayliss, K.L. Hoehn, Identification of a novel mitochondrial uncoupler that does not depolarize the plasma membrane, *Mol. Metab.* 3 (2014) 114–123.
- [38] B.Y. Qiu, N. Turner, Y.Y. Li, M. Gu, M.W. Huang, F. Wu, T. Pang, F.J. Nan, J.M. Ye, J.Y. Li, J. Li, High-throughput assay for modulators of mitochondrial membrane potential identifies a novel compound with beneficial effects on db/db mice, *Diabetes* 59 (2010) 256–265.
- [39] Y.Y. Fu, M. Zhang, N. Turner, L.N. Zhang, T.C. Dong, M. Gu, S.J. Leslie, J.Y. Li, F.J. Nan, J. Li, A novel chemical uncoupler ameliorates obesity and related phenotypes in mice with diet-induced obesity by modulating energy expenditure and food intake, *Diabetologia* 56 (2013) 2297–2307.
- [40] D.G. Hardie, K. Sakamoto, AMPK: a key sensor of fuel and energy status in skeletal muscle, *Physiology* 21 (2006) 48–60.
- [41] H. Tao, Y. Zhang, X. Zeng, G.I. Shulman, S. Jin, Niclosamide ethanolamine-induced mild mitochondrial uncoupling improves diabetic symptoms in mice, *Nat. Med.* 20 (2014) 1263–1269.
- [42] E.C. Weinbach, J. Garbus, Mechanism of action of reagents that uncouple oxidative phosphorylation, *Nature* 221 (1969) 1016–1018.
- [43] R.J. Perry, T. Kim, X.M. Zhang, H.Y. Lee, D. Pesta, V.B. Popov, D. Zhang, Y. Rahimi, M.J. Jurczak, G.W. Cline, D.A. Spiegel, G.I. Shulman, Reversal of hypertriglyceridemia, Fatty liver disease, and insulin resistance by a liver-targeted mitochondrial uncoupler, *Cell Metab.* 18 (2013) 740–748.
- [44] R.J. Perry, D. Zhang, X.M. Zhang, J.L. Boyer, G.I. Shulman, Controlled-release mitochondrial protonophore reverses diabetes and steatohepatitis in rats, *Science* 347 (2015) 1253–1256.
- [45] P.H. Lou, B.S. Hansen, P.H. Olsen, S. Tullin, M.P. Murphy, M.D. Brand, Mitochondrial uncouplers with an extraordinary dynamic range, *Biochem. J.* 407 (2007) 129–140.
- [46] J. Kopecky, G. Clarke, S. Enerback, B. Spiegelman, L.P. Kozak, Expression of the mitochondrial uncoupling protein gene from the aP2 gene promoter prevents genetic obesity, *J. Clin. Invest.* 96 (1995) 2914–2923.
- [47] J. Kopecky, Z. Hodny, M. Rossmeisl, I. Syrový, L.P. Kozak, Reduction of dietary obesity in aP2-Ucp transgenic mice: physiology and adipose tissue distribution, *Am. J. Physiol.* 270 (1996) E768–E775.
- [48] J. Kopecky, M. Rossmeisl, Z. Hodny, I. Syrový, M. Horakova, P. Kolarova, Reduction of dietary obesity in aP2-Ucp transgenic mice: mechanism and adipose tissue morphology, *Am. J. Physiol.* 270 (1996) E776–E786.
- [49] F. Baumruk, P. Flachs, M. Horakova, D. Floryk, J. Kopecky, Transgenic UCP1 in white adipocytes modulates mitochondrial membrane potential, *FEBS Lett.* 444 (1999) 206–210.
- [50] B. Stefl, A. Janovska, Z. Hodny, M. Rossmeisl, M. Horakova, I. Syrový, J. Bemova, B. Bendlova, J. Kopecky, Brown fat is essential for cold-induced thermogenesis but not for obesity resistance in aP2-Ucp mice, *Am. J. Physiol.* 274 (1998) E527–E533.
- [51] O. Matejkova, K.J. Mustard, J. Sponarova, P. Flachs, M. Rossmeisl, I. Miksik, M. Thomason-Hughes, D. Grahame Hardie, J. Kopecky, Possible involvement



- of AMP-activated protein kinase in obesity resistance induced by respiratory uncoupling in white fat, *FEBS Lett.* 569 (2004) 245–248.
- [52] M. Rossmesl, G. Barbatelli, P. Flachs, P. Brauner, M.C. Zingaretti, M. Marelli, P. Janovska, M. Horakova, I. Syrový, S. Cinti, J. Kopecky, Expression of the uncoupling protein 1 from the aP2 gene promoter stimulates mitochondrial biogenesis in unilocular adipocytes in vivo, *Eur. J. Biochem./FEBS* 269 (2002) 19–28.
- [53] P. Flachs, J. Novotný, F. Baumruk, K. Bardova, L. Bourova, I. Miksik, J. Sponarova, P. Svoboda, J. Kopecky, Impaired noradrenaline-induced lipolysis in white fat of aP2-Ucp1 transgenic mice is associated with changes in G-protein levels, *Biochem. J.* 364 (2002) 369–376.
- [54] M. Rossmesl, I. Syrový, F. Baumruk, P. Flachs, P. Janovska, J. Kopecky, Decreased fatty acid synthesis due to mitochondrial uncoupling in adipose tissue, *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* 14 (2000) 1793–1800.
- [55] T. Yamada, H. Katagiri, Y. Ishigaki, T. Ogihara, J. Imai, K. Uno, Y. Hasegawa, J. Gao, H. Ishihara, A. Nijijima, H. Mano, H. Aburatani, T. Asano, Y. Oka, Signals from intra-abdominal fat modulate insulin and leptin sensitivity through different mechanisms: neuronal involvement in food-intake regulation, *Cell Metab.* 3 (2006) 223–229.
- [56] S. Cinti, The adipose organ at a glance, *Dis. Models Mech.* 5 (2012) 588–594.
- [57] D. Loncar, B.A. Afzelius, B. Cannon, Epididymal white adipose tissue after cold stress in rats. II. Mitochondrial changes, *J. Ultrastruct. Mol. Struct. Res.* 101 (1988) 199–209.
- [58] D. Loncar, L. Bedrica, J. Mayer, B. Cannon, J. Nedergaard, B.A. Afzelius, A. Svajger, The effect of intermittent cold treatment on the adipose tissue of the rat. Apparent transformation from white to brown adipose tissue, *J. Ultrastruct. Mol. Struct. Res.* 97 (1986) 119–129.
- [59] M.L. Bonet, P. Oliver, A. Palou, Pharmacological and nutritional agents promoting browning of white adipose tissue, *Biochim. Biophys. Acta* 1831 (2013) 969–985.
- [60] T.J. Schulz, P. Huang, T.L. Huang, R. Xue, L.E. McDougall, K.L. Townsend, A.M. Cypess, Y. Mishina, E. Gussoni, Y.H. Tseng, Brown-fat paucity due to impaired BMP signalling induces compensatory browning of white fat, *Nature* 495 (2013) 379–383.
- [61] M. Ost, V. Coleman, A. Voigt, E.M. van Schothorst, S. Keipert, I. van der Stelt, S. Ringel, A. Graja, T. Ambrosi, A.P. Kipp, M. Jastroch, T.J. Schulz, J. Keijer, S. Klaus, Muscle mitochondrial stress adaptation operates independently of endogenous FGF21 action, *Mol. Metab.* 5 (2016) 79–90.
- [62] M. Harms, P. Seale, Brown and beige fat: development, function and therapeutic potential, *Nat. Med.* 19 (2013) 1252–1263.
- [63] I.G. Shabalina, N. Petrovich, J.M. de Jong, A.V. Kalinovich, B. Cannon, J. Nedergaard, UCP1 in brite/beige adipose tissue mitochondria is functionally thermogenic, *Cell Rep.* 5 (2013) 1196–1203.
- [64] J. Nedergaard, B. Cannon, The browning of white adipose tissue: some burning issues, *Cell Metab.* 20 (2014) 396–407.
- [65] P. Seale, B. Bjork, W. Yang, S. Kajimura, S. Chin, S. Kuang, A. Scime, S. Devarakonda, H.M. Conroe, H. Erdjument-Bromage, P. Tempst, M.A. Rudnicki, D.R. Beier, B.M. Spiegelman, PRDM16 controls a brown fat/skeletal muscle switch, *Nature* 454 (2008) 961–967.
- [66] F. Zurlo, K. Larson, C. Bogardus, E. Ravussin, Skeletal muscle metabolism is a major determinant of resting energy expenditure, *J. Clin. Invest.* 86 (1990) 1423–1427.
- [67] L. Simonsen, J. Bulow, J. Madsen, N.J. Christensen, Thermogenic response to epinephrine in the forearm and abdominal subcutaneous adipose tissue, *Am. J. Physiol.* 263 (1992) E850–E855.
- [68] I. Janssen, S.B. Heymsfield, Z.M. Wang, R. Ross, Skeletal muscle mass and distribution in 468 men and women aged 18–88 yr, *J. Appl. Physiol.* 89 (2000) 81–88.
- [69] R.A. DeFronzo, E. Jacot, E. Jequier, E. Maeder, J. Wahren, J.P. Felber, The effect of insulin on the disposal of intravenous glucose. Results from indirect calorimetry and hepatic and femoral venous catheterization, *Diabetes* 30 (1981) 1000–1007.
- [70] S.M. Furler, G.J. Cooney, B.D. Hegarty, M.Y. Lim-Fraser, E.W. Kraegen, N.D. Oakes, Local factors modulate tissue-specific NEFA utilization: assessment in rats using 3H-(R)-2-bromopalmitate, *Diabetes* 49 (2000) 1427–1433.
- [71] D.H. Bessesen, C.L. Rupp, R.H. Eckel, Trafficking of dietary fat in lean rats, *Obes. Res.* 3 (1995) 191–203.
- [72] V.B. Ritov, E.V. Menshikova, J. He, R.E. Ferrell, B.H. Goodpaster, D.E. Kelley, Deficiency of subsarcolemmal mitochondria in obesity and type 2 diabetes, *Diabetes* 54 (2005) 8–14.
- [73] C. Bonnard, A. Durand, S. Peyrol, E. Chanseaux, M.A. Chauvin, B. Morio, H. Vidal, J. Rieusset, Mitochondrial dysfunction results from oxidative stress in the skeletal muscle of diet-induced insulin-resistant mice, *J. Clin. Invest.* 118 (2008) 789–800.
- [74] E. Carafoli, Mitochondrial pathology: an overview, *Ann. N. Y. Acad. Sci.* 488 (1986) 1–18.
- [75] B. Li, L.A. Nolte, J.S. Ju, D.H. Han, T. Coleman, J.O. Holloszy, C.F. Semenkovich, Skeletal muscle respiratory uncoupling prevents diet-induced obesity and insulin resistance in mice, *Nat. Med.* 6 (2000) 1115–1120.
- [76] E. Couplan, C. Gelly, M. Goubern, C. Fleury, B. Quesson, M. Silberberg, E. Thiaudiere, P. Mateo, M. Lonchamp, N. Levens, C. De Montrion, S. Ortman, S. Klaus, M.D. Gonzalez-Barroso, A.M. Cassard-Douclier, D. Ricquier, A.X. Bigard, P. Diolez, F. Bouillaud, High level of uncoupling protein 1 expression in muscle of transgenic mice selectively affects muscles at rest and decreases their I1b fiber content, *J. Biol. Chem.* 277 (2002) 43079–43088.
- [77] S. Klaus, B. Rudolph, C. Dohrmann, R. Wehr, Expression of uncoupling protein 1 in skeletal muscle decreases muscle energy efficiency and affects thermoregulation and substrate oxidation, *Physiol. Genomics* 21 (2005) 193–200.
- [78] S. Keipert, S. Klaus, G. Heldmaier, M. Jastroch, UCP1 ectopically expressed in murine muscle displays native function and mitigates mitochondrial superoxide production, *Biochim. Biophys. Acta* 1797 (2010) 324–330.
- [79] Y. Katterle, S. Keipert, J. Hof, S. Klaus, Dissociation of obesity and insulin resistance in transgenic mice with skeletal muscle expression of uncoupling protein 1, *Physiol. Genomics* 32 (2008) 352–359.
- [80] M. Ost, F. Werner, J. Dokas, S. Klaus, A. Voigt, Activation of AMPK $\alpha$ 2 is not crucial for mitochondrial uncoupling-induced metabolic effects but required to maintain skeletal muscle integrity, *Plos One* 9 (2014) e94689.
- [81] S. Keipert, M. Ost, A. Chadt, A. Voigt, V. Ayala, M. Portero-Otin, R. Pamplona, H. Al-Hasani, S. Klaus, Skeletal muscle uncoupling-induced longevity in mice is linked to increased substrate metabolism and induction of the endogenous antioxidant defense system, *Am. J. Physiol. Endocrinol. Metab.* 304 (2013) E495–E506.
- [82] S. Keipert, A. Voigt, S. Klaus, Dietary effects on body composition, glucose metabolism, and longevity are modulated by skeletal muscle mitochondrial uncoupling in mice, *Aging Cell* 10 (2011) 122–136.
- [83] A. Voigt, Y. Katterle, M. Kahle, R. Kluge, A. Schurmann, H.G. Joost, S. Klaus, Skeletal muscle mitochondrial uncoupling prevents diabetes but not obesity in NZO mice, a model for polygenic diabetes, *Genes Nutr.* 10 (2015) 57.
- [84] S. Neschen, Y. Katterle, J. Richter, R. Augustin, S. Scherneck, F. Mirshahemi, A. Schurmann, H.G. Joost, S. Klaus, Uncoupling protein 1 expression in murine skeletal muscle increases AMPK activation, glucose turnover, and insulin sensitivity in vivo, *Physiol. Genomics* 33 (2008) 333–340.
- [85] S. Keipert, M. Ost, K. Johann, F. Imber, M. Jastroch, E.M. van Schothorst, J. Keijer, S. Klaus, Skeletal muscle mitochondrial uncoupling drives endocrine cross-talk through the induction of FGF21 as a myokine, *Am. J. Physiol. Endocrinol. Metab.* 306 (2014) E469–E482.
- [86] A.C. Gates, C. Bernal-Mizrachi, S.L. Chinault, C. Feng, J.G. Schneider, T. Coleman, J.P. Malone, R.R. Townsend, M.V. Chakravarthy, C.F. Semenkovich, Respiratory uncoupling in skeletal muscle delays death and diminishes age-related disease, *Cell Metab.* 6 (2007) 497–505.
- [87] A.D. van Dam, S. Kooijman, M. Schilperoort, P.C. Rensen, M.R. Boon, Regulation of brown fat by AMP-activated protein kinase, *Trends Mol. Med.* 21 (2015) 571–579.
- [88] D.R. Crooks, T.G. Natarajan, S.Y. Jeong, C. Chen, S.Y. Park, H. Huang, M.C. Ghosh, W.H. Tong, R.G. Haller, C. Wu, T.A. Rouault, Elevated FGF21 secretion, PGC-1 $\alpha$  and ketogenic enzyme expression are hallmarks of iron-sulfur cluster depletion in human skeletal muscle, *Hum. Mol. Genet.* 23 (2014) 24–39.
- [89] F.M. Fisher, S. Kleiner, N. Douris, E.C. Fox, R.J. Mepani, F. Verdeguez, J. Wu, A. Kharitonov, J.S. Flier, E. Maratos-Flier, B.M. Spiegelman, FGF21 regulates PGC-1 $\alpha$  and browning of white adipose tissues in adaptive thermogenesis, *Genes Dev.* 26 (2012) 271–281.
- [90] Y. Zhang, Y. Xie, E.D. Berglund, K.C. Coate, T.T. He, T. Katafuchi, G. Xiao, M.J. Potthoff, W. Wei, Y. Wan, R.T. Yu, R.M. Evans, S.A. Kliewer, D.J. Mangelsdorf, The starvation hormone, fibroblast growth factor-21, extends lifespan in mice, *elife* 1 (2012) e00065.
- [91] M.J. Hanssen, E. Broeders, R.J. Samms, M.J. Vosselman, A.A. van der Lans, C.C. Cheng, A.C. Adams, W.D. van Marken Lichtenbelt, P. Schrauwen, Serum FGF21 levels are associated with brown adipose tissue activity in humans, *Sci. Rep.* 5 (2015) 10275.
- [92] C. Bernal-Mizrachi, S. Weng, B. Li, L.A. Nolte, C. Feng, T. Coleman, J.O. Holloszy, C.F. Semenkovich, Respiratory uncoupling lowers blood pressure through a leptin-dependent mechanism in genetically obese mice, *Arterioscler. Thromb. Vasc. Biol.* 22 (2002) 961–968.
- [93] S. Keipert, M. Ost, A. Chadt, A. Voigt, V. Ayala, M. Portero-Otin, R. Pamplona, H. Al-Hasani, S. Klaus, Skeletal muscle uncoupling-induced longevity in mice is linked to increased substrate metabolism and induction of the endogenous antioxidant defense system, *Am. J. Physiol. Endoc. M.* 304 (2013) E495–E506.
- [94] R.S. Balaban, S. Nemoto, T. Finkel, Mitochondria, oxidants, and aging, *Cell* 120 (2005) 483–495.
- [95] J.A. Barbour, N. Turner, Mitochondrial stress signaling promotes cellular adaptations, *Int. J. Cell Biol.* 2014 (2014) 156020.
- [96] M. Ost, S. Keipert, E.M. van Schothorst, V. Donner, I. van der Stelt, A.P. Kipp, K.J. Petzke, M. Jove, R. Pamplona, M. Portero-Otin, J. Keijer, S. Klaus, Muscle mitochondrial uncoupling promotes cellular survival via serine/glycine pathway flux, *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* 29 (2015) 1314–1328.
- [97] C.N. Adjeitey, R.J. Mailloux, R.A. Dekemp, M.E. Harper, Mitochondrial uncoupling in skeletal muscle by UCP1 augments energy expenditure and glutathione content while mitigating ROS production, *Am. J. Physiol. Endocrinol. Metab.* 305 (2013) E405–E415.
- [98] S.C. Lu, Regulation of glutathione synthesis, *Curr. Top. Cell. Regul.* 36 (2000) 95–116.
- [99] R. Pearl, *The Rate of Living, Being an Account of Some Experimental Studies on the Biology of Life Duration.*, Alfred A. Knopf 1928.
- [100] P. Mitchell, Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism, *Nature* 191 (1961) 144–148.
- [101] M.D. Brand, Uncoupling to survive? The role of mitochondrial inefficiency in



- ageing, *Exp. Gerontol.* 35 (2000) 811–820.
- [102] J.R. Speakman, D.A. Talbot, C. Selman, S. Snart, J.S. McLaren, P. Redman, E. Krol, D.M. Jackson, M.S. Johnson, M.D. Brand, Uncoupled and surviving: individual mice with high metabolism have greater mitochondrial uncoupling and live longer, *Aging Cell* 3 (2004) 87–95.
- [103] I.G. Shabalina, J. Nedergaard, Mitochondrial ('mild') uncoupling and ROS production: physiologically relevant or not? *Biochem. Soc. Trans.* 39 (2011) 1305–1309.
- [104] M. Ristow, K. Zarse, How increased oxidative stress promotes longevity and metabolic health: the concept of mitochondrial hormesis (mitohormesis), *Exp. Gerontol.* 45 (2010) 410–418.
- [105] T.J. Schulz, K. Zarse, A. Voigt, N. Urban, M. Birringer, M. Ristow, Glucose restriction extends *Caenorhabditis elegans* life span by inducing mitochondrial respiration and increasing oxidative stress, *Cell Metab.* 6 (2007) 280–293.
- [106] S. Maglioni, A. Schiavi, A. Runci, A. Shaik, N. Ventura, Mitochondrial stress extends lifespan in *C. elegans* through neuronal hormesis, *Exp. Gerontol.* (2014).
- [107] E. Owusu-Ansah, W. Song, N. Perrimon, Muscle mitohormesis promotes longevity via systemic repression of insulin signaling, *Cell* 155 (2013) 699–712.
- [108] R. Luft, D. Ikkos, G. Palmieri, L. Ernster, B. Afzelius, A case of severe hypermetabolism of nonthyroid origin with a defect in the maintenance of mitochondrial respiratory control: a correlated clinical, biochemical, and morphological study, *J. Clin. Invest.* 41 (1962) 1776–1804.
- [109] M. Jastroch, S. Keipert, F. Perocchi, From explosives to physiological combustion: Next generation chemical uncouplers, *Mol. Metab.* 3 (2014) 86–87.

# Distinct functions for the paralogous RBM41 and U11/U12-65K proteins in the minor spliceosome

Antto J. Norppa<sup>1</sup>, Iftekhar Chowdhury<sup>2</sup>, Laura E. van Rooijen<sup>3</sup>, Janne J. Ravantti<sup>4</sup>, Berend Snel<sup>3</sup>, Markku Varjosalo<sup>2</sup> and Mikko J. Frilander<sup>1,\*</sup>

<sup>1</sup>Institute of Biotechnology, University of Helsinki, Helsinki, Finland

<sup>2</sup>Molecular Systems Biology Research Group and Proteomics Unit, Institute of Biotechnology, University of Helsinki, Helsinki, Finland

<sup>3</sup>Theoretical Biology and Bioinformatics, Department of Biology, Faculty of Science, Utrecht University, 3584 CH Utrecht, the Netherlands

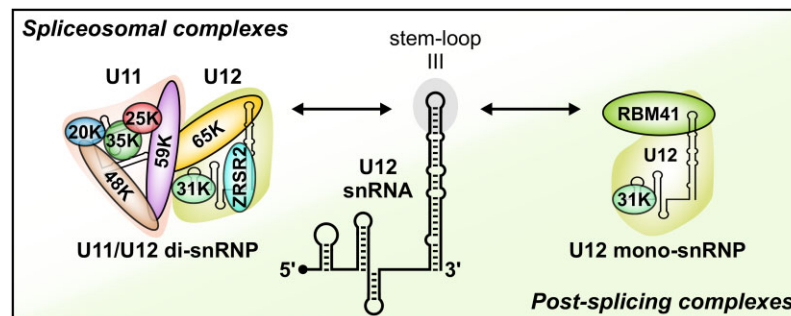
<sup>4</sup>Molecular and Integrative Biosciences Research Programme, University of Helsinki, Helsinki, Finland

\*To whom correspondence should be addressed. Tel: +358 2941 59509; Email: mikko.frilander@helsinki.fi

## Abstract

Here, we identify RBM41 as a novel unique protein component of the minor spliceosome. RBM41 has no previously recognized cellular function but has been identified as a paralog of U11/U12-65K, a known unique component of the U11/U12 di-snRNP. Both proteins use their highly similar C-terminal RRM to bind to 3'-terminal stem-loops in U12 and U6atac snRNAs with comparable affinity. Our BioID data indicate that the unique N-terminal domain of RBM41 is necessary for its association with complexes containing DHX8, an RNA helicase, which in the major spliceosome drives the release of mature mRNA from the spliceosome. Consistently, we show that RBM41 associates with excised U12-type intron lariats, is present in the U12 mono-snRNP, and is enriched in Cajal bodies, together suggesting that RBM41 functions in the post-splicing steps of the minor spliceosome assembly/disassembly cycle. This contrasts with U11/U12-65K, which uses its N-terminal region to interact with U11 snRNP during intron recognition. Finally, while RBM41 knockout cells are viable, they show alterations in U12-type 3' splice site usage. Together, our results highlight the role of the 3'-terminal stem-loop of U12 snRNA as a dynamic binding platform for the U11/U12-65K and RBM41 proteins, which function at distinct stages of the assembly/disassembly cycle.

## Graphical abstract



## Introduction

In the majority of metazoan species, the removal of spliceosomal introns from pre-mRNA is carried out by two parallel ribonucleoprotein machineries: the major spliceosome, which splices the U2-type (major) introns, and the minor spliceosome, responsible for splicing of the U12-type (minor) introns. The overall architecture of the major and minor spliceosomes is similar. Both are composed of five small nuclear ribonucleoprotein (snRNP) particles—U1, U2, U4, U5 and U6 in the major spliceosome or U11, U12, U4atac, U5 and U6atac in the minor spliceosome—and additional non-snRNP splicing factors (1–6). While the major spliceosome has been extensively characterized structurally (7), the structures of a minor spliceosome activated for the first step of splicing ( $B^{act}$ ) (8) and the U11 snRNP (9) are presently the only high-resolution

cryo-EM structures of minor spliceosomal complexes available. However, it is generally accepted that the overall assembly pathway of the minor spliceosome is similar to that of the major spliceosome as both spliceosomes utilize similar snRNP complexes and a two-step splicing mechanism with branching and exon ligation reactions (2,3,10). A key mechanistic difference is in the recognition of U12-type introns, which is carried out by a preformed U11/U12 di-snRNP that cooperatively binds to both the 5' splice site (5'ss) and branch point sequence (BPS) (11,12). In contrast to the earlier steps of the minor spliceosome assembly, very little is known about the post-catalytic events of the minor spliceosome cycle. In the major spliceosome, the DHX8 (hPrp22) helicase drives the release of the ligated exon product from the post-catalytic (P) complex and turns it to the intron lariat spliceosome (ILS)

Received: December 11, 2023. Revised: January 19, 2024. Editorial Decision: January 22, 2024. Accepted: March 11, 2024

© The Author(s) 2024. Published by Oxford University Press on behalf of Nucleic Acids Research.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

which is subsequently disassembled by the DHX15 (hPrp43) helicase (13–15). Recycling of snRNPs for subsequent rounds of splicing is thought to take place in the Cajal body, which is also the cellular site for other processes related to snRNP biogenesis (16). There, the Cajal body-localized recycling factor SART3 is thought to function in both spliceosomes (17,18).

Both the major and the minor spliceosome contain four unique small nuclear RNAs (snRNAs) and a number of unique protein components that are not found in the other spliceosome, while the majority of proteins are likely to be shared (8,19–21). The first set of minor spliceosome-specific proteins (20K, 25K, 31K, 35K, 48K, 59K, 65K) were identified over 20 years ago by affinity purification and mass spectrometry analysis of the U11/U12 di-snRNP and U11 mono-snRNP fractions (19,20). Of these, U11/U12-65K, U11-59K and U11-48K form a chain of interactions connecting the U11 and U12 mono-snRNPs into a di-snRNP (22,23) and are essential for the stability of the di-snRNP (23–25). The U11 snRNP-associated U11-48K protein recognizes the 5' splice site together with the U11 snRNA and interacts with the U11-59K protein (23), which is further engaged in an interaction with the N-terminal part of U11/U12-65K (22). The C-terminal RNA recognition motif (RRM) of U11/U12-65K binds the 3'-terminal stem-loop of U12 snRNA (26), but can also interact with the 3'-terminal stem-loop of the U6atac snRNA (27). ZRSR2, a component of the U11/U12 di-snRNP responsible for 3' splice site recognition (20), has also been shown to almost exclusively affect splicing of U12-type introns in the cell (28), even though it may also have a separate role in the major spliceosome (29).

The notion that specific protein components of the minor spliceosome are needed only during the intron recognition phase and not in the later assembly steps has been challenged only very recently. The first such factor, the plant ortholog of RBM48, was originally identified as a U12-type intron splicing factor from a transposon screen in maize (30), and its specificity for U12-type introns was later confirmed in human cells (31). A cryo-EM structure of the catalytically activated minor spliceosome ( $B^{act}$  complex) revealed that RBM48, together with additional three unique proteins ARMC7, SCNM1 and CRIPT, are all specific components of the  $B^{act}$  complex (8). The first unique protein component of the U4atac/U6atac di-snRNP and U4atac/U6atac.U5 tri-snRNP, CENATAC, was initially identified as a human disease gene and was shown to be specifically required for the splicing of U12-type introns with AT-AN termini (32). Similarly, mutations in DROL1, the plant ortholog of TXNL4B, the main interactor of CENATAC (32), were shown to lead to splicing defects with U12-type introns with AT-AC termini in *Arabidopsis thaliana* (33).

In this work, we provide evidence that the specific protein components in the minor spliceosome are not limited to intron recognition and catalytic steps. We show that RBM41, a closely related paralog of the U11/U12-65K protein, is a novel specific component of the minor spliceosome. The two paralogous proteins have a similar shared dual RNA binding specificity *in vitro*, interacting with the same 3'-terminal stem-loops of U12 and U6atac snRNAs with approximately equal affinity. However, while the U11/U12-65K functions during the intron recognition step, our results provide strong evidence that RBM41 functions in post-catalytic steps of the U12-type intron splicing. This further suggests that the 3'-terminal stem-loop of the U12 snRNA has a role as a protein-binding platform that is dynamically recognized first by the U11/U12-65K

during the early steps of the spliceosome assembly followed by an exchange to RBM41 during or after the catalytic steps of splicing.

## Materials and methods

### Antibodies

The following antibodies were used in this study at the dilutions indicated for western blot (WB) and immunofluorescence (IF): anti-HA (BioLegend, 16B12, WB: 1:2000, IF: 1:1000), anti-GAPDH (Proteintech, 60004-1, WB: 1:5000), anti-PDCD7 (Proteintech, 12485-1-AP, WB: 1:500), anti-RBM41 (Atlas Antibodies, HPA042881, WB: 1:1000, IF: 1:200), anti-RNPC3 (Santa Cruz, sc-514951, WB: 1:500, IF: 1:100), anti-Sm (Invitrogen, MA5-13449, WB: 1:500), anti-SNRNP48 (Proteintech, 24297-1-AP, WB: 1:500), anti-V5 (Invitrogen, R960-25, WB: 1:5000), anti-ZCRB1 (Bethyl Laboratories, A304-697A, WB: 1:2000), Goat anti-Rabbit Alexa Fluor 488 (Thermo Fischer, A11008, IF: 1:500), Goat anti-Mouse Alexa Fluor 568 (Abcam, Ab175473, IF: 1:500), Goat Anti-Mouse IgG HRP, light chain-specific (Jackson ImmunoResearch, 115-035-174, WB: 1:10 000), Goat Anti-Mouse IgG HRP (Millipore, AP308P, WB: 1:10000), Donkey Anti-Rabbit IgG HRP (Amersham, NA934V, WB: 1:10 000), Mouse IgG2a kappa Isotype Control (Invitrogen, 14-4724-82) served as a negative control antibody in IP experiments. eBioscience Avidin-HRP (Invitrogen, 18-4100, WB: 1:500) was used to detect biotinylated proteins. Protein G-HRP (Millipore, 18-161, WB: 1:5000) was used in western blotting after IP to avoid detection of heavy and light chains from the IP antibody.

### Plasmid construction

RBM41 cDNA sequence, corresponding to Ensembl transcript ENST00000372479.7, was amplified from HEK293 cDNA in two fragments. Gibson assembly was used to assemble the fragments into vector pCI-neo in-frame with an N-terminal V5 epitope tag, resulting in pCI-neo-V5-RBM41. For BioID cell line construction, full-length RBM41 (1–413), RBM41 N-terminal fragment (1–258), RBM41 C-terminal fragment (259–413) and full-length 65K were cloned into MAC-tag-N vector (Addgene #108078) using Gateway cloning as described (34). Human DHX8 cDNA (MGC clone 5529639) was obtained from the Genome Biology Unit at the University of Helsinki and cloned into pCI-neo in-frame with an N-terminal V5 tag. Point mutations were introduced into plasmids using site-directed mutagenesis with Phusion polymerase (Thermo).

### Expression and purification of recombinant proteins

C-terminal fragments of RBM41 (amino acids 267–413) and U11/U12-65K (amino acids 380–517) were expressed as GST-tagged proteins in *E. coli* Rosetta cells. Protein expression was induced with 0.5 mM IPTG for 3 h at 37°C and cells were harvested by centrifugation. Cell pellets were resuspended in 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 5 mM DTT, 1× cOmplete protease inhibitor cocktail (Roche) and lysed by sonication (30 s ON / 30 s OFF, 5 min total sonication time, Sonopuls HD 2070 with MS 73 microtip). After removal of cell debris (25 000 g, 15 min), GST-tagged proteins were captured from the lysate by incubation with glutathione agarose

(Pierce) for 1 h at 4°C. The resin was washed 3 times with 20 volumes of Wash/Cleavage buffer (50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 mM EDTA, 1 mM DTT) and GST tag cleaved by overnight digestion with PreScission Protease (Cytiva). Purified proteins were dialyzed against EMSA buffer (20 mM HEPES-KOH, 100 mM KCl, 1.5 mM MgCl<sub>2</sub>, 5% glycerol, pH 8.0) and protein concentrations measured by Pierce BCA protein assay (Thermo Scientific).

### Electrophoretic mobility shift assay

Recombinant proteins (0.25–25 μM) were incubated for 1 h on ice with 5 nM [ $\gamma$ -<sup>32</sup>P]-ATP-labeled RNA oligonucleotides in binding buffer (20 mM HEPES-KOH (pH 8.0), 100 mM KCl, 1.5 mM MgCl<sub>2</sub>, 5% glycerol, 0.1 μg/μl BSA, 1 μg/μl yeast RNA (Roche), 1 U/μl RiboLock RNase inhibitor (Thermo Scientific)) in 10 μl final volume. After addition of 2.5 μl of 5× loading buffer (20 mM HEPES-KOH, 100 mM KCl, 1.5 mM MgCl<sub>2</sub>, 50% glycerol, 0.1% bromophenol blue, 0.1% xylene cyanol), 4 μl of each binding reaction was loaded onto a native polyacrylamide gel (6%, 80:1 acrylamide:bisacrylamide ratio, 0.5× TBE, 5% glycerol). Electrophoresis was carried out at 120 V for 1.5 h at 4°C. After drying, the gel was exposed on an imaging plate and scanned using the FLA-5100 phosphorimager. Bound and free RNA bands were quantified using AIDA software (Raytest). Dissociation constants were determined by nonlinear regression using GraphPad Prism (one site – specific binding) from three independent replicates.

### Cell culture, transfection and nuclear extract preparation

HEK293 cells were grown in DMEM, 10% FBS, 1% penicillin-streptomycin and 2 mM L-glutamine at 37°C, 5% CO<sub>2</sub>. All plasmid transfections were carried out using Lipofectamine 2000 (Thermo Fisher) according to manufacturer's instructions. The growth conditions for HeLa S3 suspension cells and the subsequent nuclear extract preparation have been described by de Wolf *et al.* (32).

### CRISPR/Cas9-mediated knockout of RBM41

HEK293 cells were transfected with pSpCas9(BB)-2A-Puro vectors with sgRNA sequences targeting RBM41 exon 2 or 3. 24 h after transfection, puromycin was added at 3 μg/ml to enrich for transfected cells and puromycin treatment continued for 72 h. After puromycin selection, monoclonal cell lines were obtained by limiting dilution in 96-well plates. Genomic DNA was extracted from single-cell clones using the NucleoSpin Tissue kit (Macherey-Nagel), and the targeted areas amplified by PCR using primers in the introns flanking exons 2 and 3 (Supplementary Table S1). Clones were screened for editing using the Surveyor Mutation Detection Kit (IDT). Positive clones were verified by TOPO cloning of PCR products followed by sequencing, as well as by direct sequencing of the PCR products and deconvolution of sequencing traces using the DECODR tool (35).

### Northern and western blotting

Northern blotting was carried out exactly as described (36) using LNA or DNA oligonucleotide probes listed in Supplementary Table S1. For western blotting, protein samples were resolved on 4–12% NuPAGE gels and transferred

to Amersham Hybond P membrane using the Amersham TE 22 Mighty Small Transphor tank. After blocking for 1 h in 5% milk/TBS-T, membranes were incubated for 1 h with primary antibodies, followed by 3 × 5 min washes with TBS-T. After 1 h incubation with an HRP-conjugated secondary antibody, membranes were washed 5 × 5 min with TBS-T. Chemiluminescence was detected using the SuperSignal West Atto (Thermo Scientific) or Amersham ECL Prime detection reagent and the LAS-3000 imager (Fuji).

### RNA immunoprecipitation and protein co-immunoprecipitation

For RNA immunoprecipitation with V5-tagged proteins, pCIneo vectors for expressing V5-tagged proteins or empty pCIneo vector (4 μg) were transfected into HEK293 cells in 6-well plate format using Lipofectamine 2000. For RNA immunoprecipitation with endogenous proteins in cell lysate, typically ~20 × 10<sup>6</sup> cells were used for one IP. 24 h after transfection, cells were washed with 1× PBS, scraped into lysis buffer (20 mM HEPES, pH 7.9, 137 mM NaCl, 10% glycerol, 1% NP-40 + 1× cComplete protease inhibitor cocktail + 0.5 U/μl RiboLock) and sonicated (5 × 30 s, Bioruptor Twin, High setting). Cell debris was removed by centrifugation at 16000 g for 15 min. The supernatant was incubated o/n with rotation at 4°C with 2 μg IP antibody and for 1 h with Dynabeads Protein G (Invitrogen). Alternatively, the antibody was prebound to Dynabeads and incubated with lysate for 1 h at 4°C. After five washes with 200 μl lysis buffer, co-immunoprecipitated RNA was isolated by treatment of beads with Proteinase K, phenol:chloroform:isoamyl alcohol (pH 4.8) extraction and ethanol precipitation. For nuclear extract RIP experiments, 60 μl HeLa S3 nuclear extract was used for one IP. Nuclear extract was preincubated for 10 min at 37°C in 13 mM HEPES (pH 7.9), 2.4 mM MgCl<sub>2</sub>, 40 mM KCl, 2 mM DTT, 20 mM creatine phosphate, 0.5 mM ATP. After preincubation, the nuclear extract was diluted 1:5 and adjusted to 20 mM HEPES (pH 7.9), 137 mM KCl, 10% glycerol, 0.1% NP-40, 0.5 U/μl RiboLock, 1× cComplete protease inhibitor. The remaining steps were carried out as described above.

Protein co-immunoprecipitation experiments were carried out similarly to RIP experiments, except that the IP antibodies were crosslinked to Dynabeads Protein G with bis(sulfosuccinimidyl)suberate (BS<sub>3</sub>, Thermo) and complexes were eluted from the beads by heating in 1× NuPAGE LDS Sample Buffer for 5 min at 95°C.

### Glycerol gradient centrifugation

HeLa S3 nuclear extract (200 μl), prepared essentially as described in Tarn and Steitz (37), was adjusted to 13 mM HEPES (pH 7.9), 2.4 mM MgCl<sub>2</sub>, 40 mM KCl, 2 mM DTT, 20 mM creatine phosphate, 0.5 mM ATP, and pre-incubated for 10 min at 30°C in a final volume of 300 μl. After incubation, 210 μl of Gradient buffer (20 mM HEPES (pH 7.9), 40 mM KCl, 2 mM DTT, 2.4 mM MgCl<sub>2</sub>) was added, samples were centrifuged briefly (20 000 g, 1 min) and the supernatant was loaded on top of 10–30% glycerol gradient in Gradient buffer. Gradients were centrifuged in a Sorvall TH-641 rotor at 29 000 rpm, 4°C for 18 h and fractionated using BioComp Piston Gradient Fractionator. Gradient preparation, centrifugation and fractionation was carried out by the HiLIFE Bio-complex unit at the University of Helsinki. For RNA extraction, 20% of each fraction was treated with Proteinase K,



phenol:chloroform extracted and precipitated with ethanol. The remaining 80% of each fraction was precipitated with TCA for western blot analysis.

### BioID analysis

For each cell line, Flp-In™ T-REx™ 293 cells were grown in 5 × 15 cm plates to ~70% confluency. MAC-tagged protein expression and biotinylation was induced by addition of 2 µg/ml of tetracycline and 50 µM biotin. Cells were harvested 24 h after induction by pipetting up and down with PBS-EDTA, centrifugation at 1200 g for 5 min, and snap freezing the pellet in liquid nitrogen. Three independent replicates of 5 × 15 cm dishes were prepared for each cell line. BioID analysis was carried out essentially as described previously (34,38).

### Immunofluorescence

HEK293 or Flp-In™ T-REx™ 293 cells were grown on poly-L-lysine coated coverslips in a 12-well plate. Cells were fixed with 4% paraformaldehyde/1× PBS for 10 min at room temperature. After two washes with 1× PBS, cells were permeabilized with 0.2% Triton X-100/1× PBS for 10 min and blocked with 1% BSA/1× PBS for 30 min at room temperature. Incubation with primary antibodies, diluted in blocking buffer, was carried out overnight at 4°C. Following 3 × 5 min washes with 1× PBS, cells were incubated with Alexa Fluor 488 or 568 conjugated secondary antibodies and DAPI (0.5 µg/ml) for 1 h at room temperature. Antibodies used for immunofluorescence and their respective dilutions are listed in the Antibodies section above. After final washes with 1× PBS (3 × 5 min), coverslips were mounted onto slides with ProLong Diamond (Invitrogen). Imaging was carried out using a Leica DM5000B microscope.

### cDNA synthesis and RT-PCR

RNA was treated with RQ1 RNase-free DNase (Promega) to remove any genomic DNA contamination. cDNA synthesis was carried out using Maxima H Minus RT (Thermo) and random primers according to the manufacturer's instructions. For PCR, Phire polymerase (Thermo) and primers listed in [Supplementary Table S1](#) were used.

### High-throughput sequencing

Total RNA isolated using Trizol extraction followed by an additional acidic phenol (pH 5.0) extraction. RNAseq libraries were constructed using Illumina TruSeq Stranded Total RNA kit (Illumina) Human Ribo-Zero rRNA depletion kit (Illumina). Paired-end 150 + 150 bp sequencing was performed at the Institute for Molecular Medicine Finland FIMM Genomics unit with Illumina NovaSeq 6000 using partial S4 flow cell lane. The STAR aligner (39) was used for mapping the paired sequence reads to the genome (hg38/GRCh38). Transcript annotations were obtained from GENCODE (v29). The length of the genomic sequence flanking the annotated junctions (sjdbOverhang parameter) was set to 161. The Illumina adapter sequences AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC and AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTGATCTCGGTGGTCGCCGTATCATT were, respectively, clipped from the 3' of the first and the second pairs in the read libraries (using clip3pAdapterSeq parameter). Statistics of the RNAseq data are presented in the [Supplementary Table S5](#).

### Differential alternative splicing analysis and intron retention analysis

Differential alternative splicing (AS) analysis was done using Whippet (v0.11) (40). Both merged aligned reads (bam files) and AS event annotations from GENCODE (v29) were used to build the index reference for AS events. To detect the significantly differential events, probability cutoff of  $Pr > 0.9$  and Percentage Spliced In deviation cutoff of  $|\Delta\Psi| > 0.05$  were used. Differential intron retention was analyzed with IRFinder-S using SUPPA2 wrapper (41,42). A custom list of human U12-type intron coordinates ([Supplementary Table S8](#)) combining high-confidence U12-type introns from IntERESt (43), IAOD (44) and MIDB (45) databases was used in the annotation of U12-type introns and their host genes.

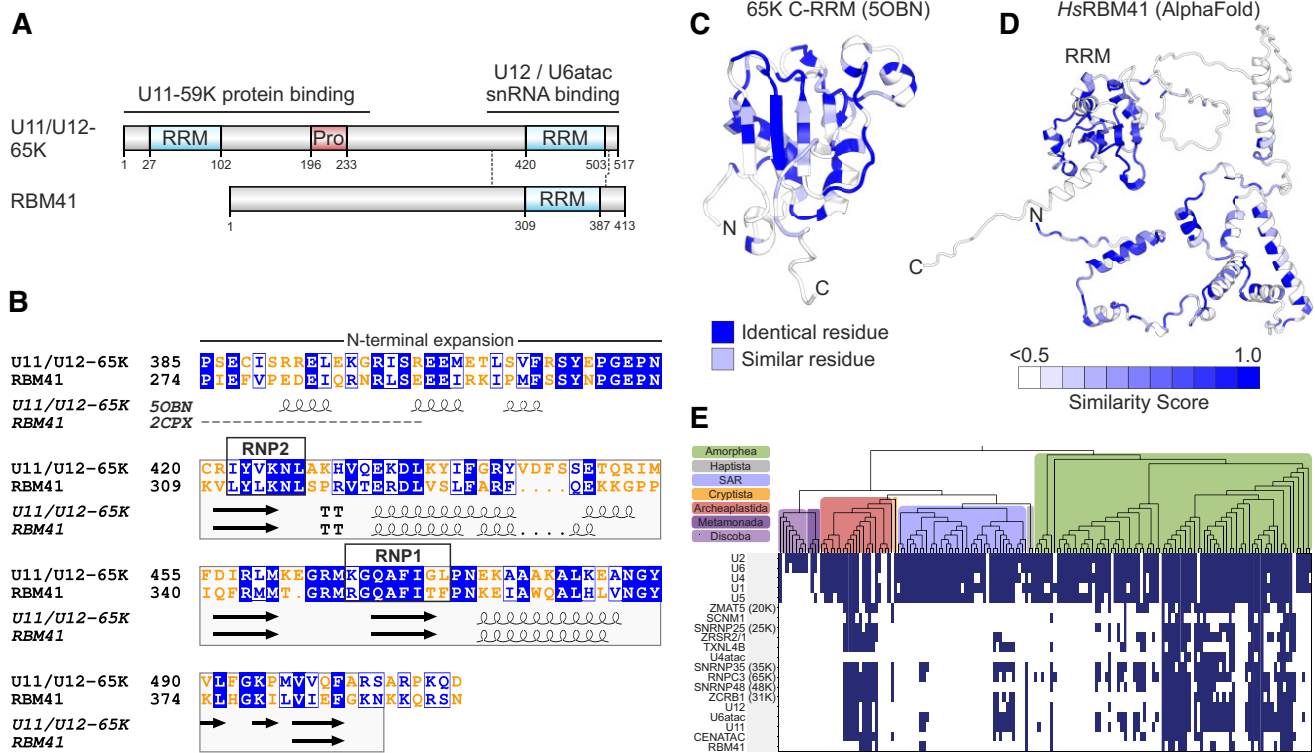
### Phylogenetic profiling-based co-evolution analysis

To conduct the phylogenetic profiling analysis of RBM41, we utilized a diverse set of eukaryotic proteomes. This dataset, consisting of 167 eukaryotic species, was previously compiled to represent the eukaryotic tree of life, and the species were selected based on their representation in the tree. We used automatic orthologous groups (OG) based on previous work (46,47), generated using methods such as Orthofinder (48), eggNOG (49) hmm profile database, and OGs from Vosseberg *et al.* (47). However, as RBM41 was not accurately represented in these automatically generated OGs, we manually created the OG for RBM41. This was achieved by performing a blast search with RBM41 against our in-house eukaryotic dataset and creating a Hidden Markov model (HMM) of RBM41, which was then used to perform a HMM search (50) against the same dataset. To determine the OG, a phylogenetic analysis was carried out with the top 100 entries of the HMM search using mafft E-INS-I (51) and IQtree (52). In this phylogeny, a cluster with representatives from diverse eukaryotic groups indicated the RBM41 Orthologous group. Besides the support-value, the species overlap of this cluster with the other putative orthologous groups in the tree solidifies it as resulting from an ancient duplication and having been present in LECA. To compare the phylogenetic similarity, we computed cosine distances between the phylogenetic profile of each automatically generated OG and RBM41. As we suspected that RBM41 was part of the minor spliceosome, we also compared the phylogenetic profile of RBM41 to the profiles of other known spliceosome proteins obtained from Vosseberg *et al.* (53).

## Results

### RBM41 is a paralog of the U11/U12-65K protein and a putative component of the minor spliceosome

Human RBM41 is a ~47 kDa protein with a single annotated domain, an RNA recognition motif (RRM) near the C-terminus of the protein (at positions 309–387) (Figure 1A). While no specific cellular function has been assigned to RBM41, it has been listed as a paralog of the U11/U12-65K protein (RNPC3), a structural component of the U11/U12 di-snRNP complex in the minor spliceosome (22,53). The paralog assignment is based on the local sequence similarity between the C-terminal regions of the two proteins, that encompass the core RRM and its N-terminal expansion



**Figure 1.** RBM41 is a paralog of the U11/U12-65K protein. **(A)** Domain structures of human U11/U12-65K and RBM41 proteins. **(B)** Pairwise sequence alignment of RBM41 and U11/U12-65K. Local sequence alignment was carried out using Matcher and visualized using ESPrnt 3.0 (80). Identical residues are shown in white text with blue background and similar residues in blue text with white background. Protein secondary structure elements extracted from NMR structures (U11/U12-65K: 50BN, RBM41: 2CPX) are shown below the alignment. **(C)** Structure of the U11/U12-65K C-terminal RRM (50BN) showing identical and similar residues between RBM41 and U11/U12-65K. **(D)** AlphaFold-predicted structure of human RBM41 colored for sequence conservation. Conservation is based on a multiple sequence alignment of RBM41 orthologues from 15 animal species (Supplementary Figure S3). Conservation was mapped to the structure with ESPrnt 3.0 and structure rendered using PyMOL. **(E)** Phylogenetic profile of RBM41 compared to the known minor spliceosome-specific proteins and minor and major spliceosomal snRNAs.

(Figure 1B, C) which in the U11/U12-65K protein is essential for the stability of the C-terminal RRM and its interaction with the U12 snRNA (26). The RRM sequences of 65K and RBM41 conform with the general RRM consensus (54) except for an aromatic to nonaromatic amino acid substitution F/Y352Q at position 3 of the RNP1 motif (Figure 1B). The same substitution is present of in the homologous N-terminal RRMs of the U1A/U2B''/SNF family of spliceosome components, which are characterized by a YQF triad of RNP2 tyrosine and RNP1 glutamine and phenylalanine (55,56; Supplementary Figure S1). These three residues are displayed on the  $\beta$ -sheet surface of the RRM and engage in stacking interactions with RNA nucleobases (55,56). Conservation of the corresponding residues in the RBM41 RRM (Y312/Q352/F354; Supplementary Figure S1) suggests that RBM41 may employ a similar mode of RNA binding as utilized by the U1A/U2B''/SNF proteins.

The homology between RBM41 and U11/U12-65K raises the question of whether RBM41 also functions in the minor spliceosome. We carried out a phylogenetic profiling-based co-evolution analysis in 167 eukaryotic species to identify proteins that show a similar phylogenetic presence/absence profile as RBM41 and may therefore function in the same molecular process (Figure 1E). Notably, three minor spliceosome-specific proteins, CENATAC, U11-48K (SNRP48) and U11/U12-31K (ZCRB1) were found

among the proteins showing the strongest co-occurrence with RBM41 (Supplementary Table S3, Supplementary Figure S2), suggesting a role in the minor spliceosome. This was further supported by earlier reciprocal coevolution analysis for CENATAC (32), which similarly identified RBM41 as one of the top co-occurring proteins among the known minor spliceosome components. Upon comparing RBM41's phylogenetic profile with that of other known minor spliceosome proteins (Figure 1E, Supplementary Figure S2), it becomes evident that RBM41 adheres to the typical presence/absence profile of these proteins. However, the presence of RBM41 appears to be even less common than that of other minor spliceosome proteins.

In contrast to the homologous C-terminal RRMs, the N-termini of the two proteins do not share sequence similarity. This suggests a functional difference between RBM41 and U11/U12-65K, that uses the N-terminus to promote the formation of the U11/U12 di-snRNP via an interaction with the U11-59K protein (22; Figure 1A). RBM41 N-terminus is highly conserved among animal orthologs (Supplementary Figure S3, Supplementary Table S2, Figure 1D) but lacks annotated domains and adopts a predominantly helical conformation in an AlphaFold prediction (57) (Figure 1D). Together the sequence analysis suggests similar RNA binding properties, but otherwise divergent functions for the two paralogous proteins.

## RBM41 RRM interacts with the U12 and U6atac snRNAs *in vitro*

Previously, two independent systematic high-throughput studies of RNA-binding protein specificities (58,59) reported nearly identical consensus RNA motifs for RBM41 (Figure 2A) which are preferentially located in a loop sequence within an RNA stem-loop context. These RNA motifs bear a striking similarity to the loop sequences of the U12 snRNA 3'-terminal stem-loop and U6atac snRNA 3'-terminal stem-loop, both of which are bound by the U11/U12-65K protein (Figure 2A,B; 22,27). Given the apparent RNA-binding similarities between the RBM41 and U11/U12-65K proteins, we compared their RNA-binding characteristics *in vitro*, using recombinant C-terminal RRMs containing additional N- and C-terminal regions known to be essential for RNA binding (25,26). RNA binding properties were analyzed using electrophoretic mobility shift assay (EMSA) with untagged RRMs and RNA oligonucleotides corresponding to the apical hairpins of the U12 snRNA and U6atac 3'-terminal stem-loops (Figure 2C,D; 22,25,26).

Our EMSA analyses revealed that U11/U12-65K C-RRM and RBM41 RRM have similar overall binding characteristics, binding to both U12 and U6atac hairpins (Figure 2C, compare lanes 1–12 and 13–23). In contrast, no complex formation was observed with either RRM when a control hairpin (complementary to the U12 hairpin) was used as a ligand (Figure 2C, lanes 24 and 25). A further determination of the dissociation constants ( $K_d$ ) revealed that the RBM41 RRM has a low micromolar affinity to both U12 ( $K_d = 2.14 \mu\text{M}$ ) and U6atac hairpins ( $K_d = 3.78 \mu\text{M}$ ; Figure 2D). The 65K C-RRM showed approximately 2-fold higher affinity to both hairpins (U12:  $K_d = 1.04 \mu\text{M}$ ; U6atac:  $K_d = 2.07 \mu\text{M}$ ). Furthermore, both RRMs bound the U12 hairpin with ~2-fold higher affinity compared to the U6atac hairpin (Figure 2D). Thus, both RBM41 and U11/U12-65K C-terminal RRMs show dual snRNA binding specificity *in vitro* with only slight differences in RNA affinity.

## RBM41 specifically associates with minor spliceosomal snRNPs

We next asked if RBM41 also associates with spliceosomal snRNAs *in vivo*. We transfected HEK293 cells with V5-tagged 65K or RBM41 followed by RNA immunoprecipitations (RIP) with anti-V5 antibody and northern blot analysis of minor and major spliceosomal snRNAs (Figure 3A). Consistent with its function in the U11/U12 di-snRNP, V5-65K strongly co-immunoprecipitated the U11 and U12 snRNAs and to a lesser extent U6atac and U4atac snRNAs (Figure 3A, lane 8). In contrast, V5-RBM41 co-immunoprecipitated the U12 snRNA, U4atac and U6atac snRNAs, but not the U11 snRNA (Figure 3A, lane 7). Similar results were obtained with an antibody against the endogenous RBM41 in HeLa nuclear extract (Figure 3B, lanes 1–4) or HEK293 total cell lysate (Figure 3B, lanes 5–8), showing that these interactions are not artifacts resulting from RBM41 overexpression. U1, U2, U4, U5 and U6 snRNAs were variably detected slightly above control IP levels with transiently overexpressed V5-RBM41 (Figure 3A,D) and endogenous RBM41 (Figure 3B), likely due to non-specific association of RBM41 with these highly abundant snRNPs (see below).

Next, we tested the role of the RBM41 RRM and the large N-terminal region lacking any identifiable domains for

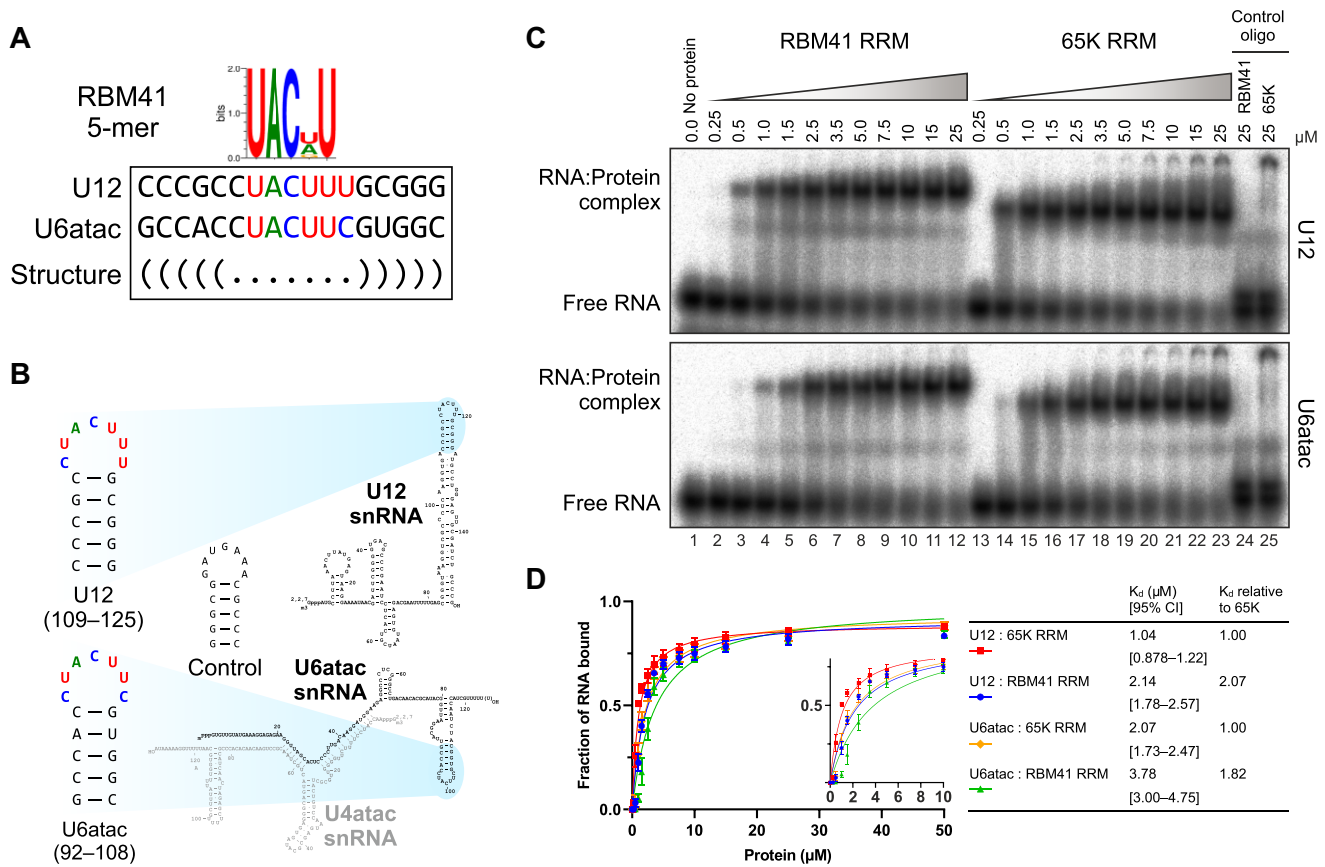
their association with snRNPs. We carried out RIP experiments using wild-type V5-RBM41, V5-RBM41 lacking the RRM domain (V5-RBM41(1–258)) or the entire N-terminal region (V5-RBM41(259–413)), as well as V5-RBM41 constructs with alanine substitutions of Gln352 or Phe354 of the conserved YQF triad in the RRM (V5-RBM41-Q352A and V5-RBM41-F354A) (Figure 3C). The Q352A and F354A mutations and deletion of the RRM led to dramatic loss of the U12, U4atac and U6atac interactions (Figure 3D, lanes 8–11), while the V5-RBM41(259–413) construct still showed robust co-immunoprecipitation of all three snRNAs (Figure 3D, lane 12). Notably, while major spliceosome-specific snRNAs (U1, U2, U4, U6) and the shared U5 snRNA were also detected above control IP levels in the V5-RBM41 anti-V5 IP (Figure 3D, lane 8), these were unaffected by the RRM mutations or the truncations, indicating that these IP signals represent non-specific background rather than specific interactions. Taken together, our RIP experiments show that RBM41 specifically associates with minor spliceosomal snRNPs in the cell and suggest that the snRNP association of RBM41 is primarily mediated through the RRM binding with its target snRNAs.

## RBM41 and U11/U12-65K partition into distinct snRNP complexes

While our *in vitro* binding experiments demonstrated similar RNA-binding properties for RBM41 and U11/U12-65K, the distinct snRNA IP profiles and the lack of sequence similarity outside the RRM suggested functional divergence of the two proteins within the minor spliceosome. As a complementary method to study snRNP complex association of RBM41 and U11/U12-65K, we carried out glycerol gradient fractionation of HeLa nuclear extract and analyzed the sedimentation behavior of the two proteins by western blot and spliceosomal snRNAs by northern blot (Figure 4A). Consistent with our RIP experiments, RBM41 and U11/U12-65K were largely found in different fractions, indicating association with different molecular complexes. RBM41 peak co-migrated with the U12 mono-snRNP (Figure 4A, fractions 8–10), whereas the U11/U12-65K peak co-migrated with the U11/U12 di-snRNP (fractions 12–14). Though the U4atac or U6atac snRNAs did not show clear co-migration with RBM41, a minor fraction of these snRNAs was present in RBM41 peak gradient fractions. Taken together with our RIP data, while the two proteins share similar RNA binding specificity *in vitro*, in the cell they partition into distinct snRNP complexes.

To define the composition and interactions of the RBM41 in the U12 mono-snRNP related to the 65K in U11/U12 di-snRNP, we used proximity-based labeling (BioID) to map the proximity interactors of the two proteins. A panel of RBM41 constructs and a U11/U12-65K construct (Figure 4B), each carrying an N-terminal MAC-tag (consisting of BirA\* biotin ligase, hemagglutinin (HA) and StrepIII tags), were integrated into the Flp-In™ T-REx™ 293 cell line, enabling both inducible transgene control (tetracycline) and inducible biotinylation of proteins coming into proximity with the bait protein. Wild-type U11/U12-65K, wild-type RBM41, four RNA binding deficient RBM41 mutants (Y312A, Q352A, F354A, Y312A + F354A) and two RBM41 truncation constructs (1–258 and 259–413) were analyzed (Figure 4B). Western blot analysis with anti-HA antibody confirmed correct expression of MAC-tagged bait proteins (Supplementary Figure S4A), while biotinylation was visualized with avidin-HRP





**Figure 2.** RBM41 interacts with the U12 and U6atac snRNAs *in vitro*. **(A)** Consensus RNA motifs bound by RBM41 *in vitro* and matching sequences in the U12 and U6atac snRNAs. The consensus motif (obtained from ENCODE database (81), entry ENCSR637HFY) determined by Ray *et al.* (58) using the RNAcompete method is shown. **(B)** RNA hairpins used in EMSA experiments and their location in the U12 and U6atac snRNAs. **(C)** EMSA analysis of RBM41 and U11/U12-65K RRM binding to U12 (top panel) and U6atac snRNA (bottom panel) hairpins. EMSA was carried out using recombinant RBM41 RRM (residues 267–413) or 65K C-terminal RRM (residues 380–517) and  $^{32}\text{P}$ -labeled U12, U6atac or negative control RNA hairpins shown in panel B. **(D)** Binding curves and dissociation constants for the interaction of RBM41 and 65K RRMs with U12 and U6atac hairpins. The inset shows a low protein concentration range (0–10  $\mu\text{M}$ ) of the same binding curves.

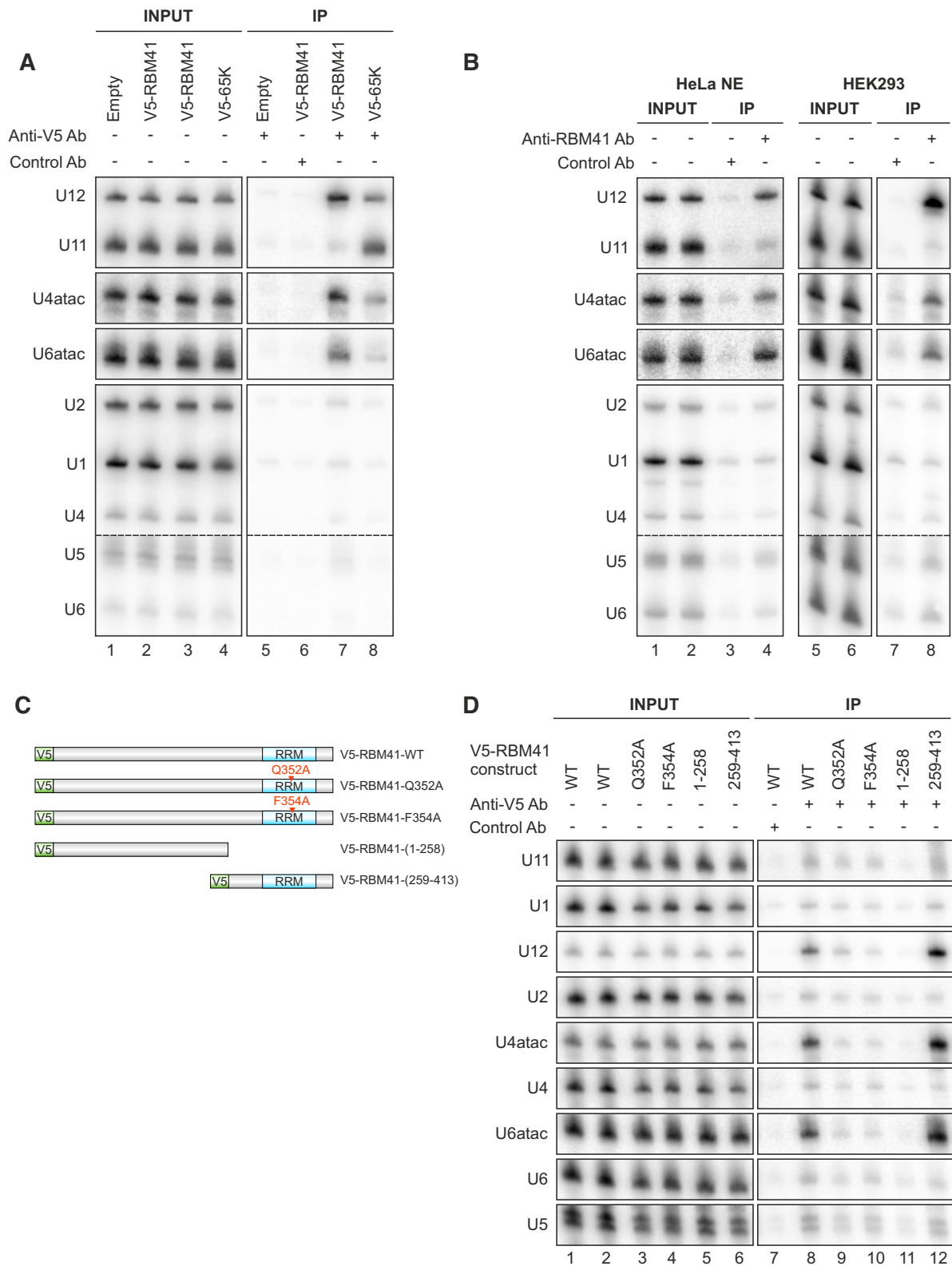
(Supplementary Figure S4B). Immunofluorescence with anti-HA was used to detect localization of the MAC-tagged proteins (Supplementary Figure S5A). MAC-RBM41-WT and MAC-U11/U12-65K proteins showed a predominantly nuclear localization, with MAC-RBM41-WT showing a more prominent cytoplasmic subpopulation. In contrast, MAC-RBM41-(1–258) was almost uniformly distributed between cytoplasm and nucleus, suggesting that deletion of the C-terminal RRM impairs nuclear import of RBM41.

MAC-tagged protein expression was induced with tetracycline in the presence of biotin for 24 h before harvesting of cells. Three independent replicates were analyzed for each cell line using an established BioID pipeline that has been described in detail (34,38). Consistent with its role in U11/U12 di-snRNP, MAC-U11/U12-65K interacted with all 8 minor spliceosome-specific proteins of the U11/U12 di-snRNP complex (Figure 4C, Supplementary Table S4), with its top interactors being U11-59K (PDCD7), U11/U12-20K (ZMAT5), U11-48K (SNRNP48) and U11-35K (SNRNP35). In contrast, MAC-RBM41-WT only interacted with one of the U11/U12 di-snRNP proteins, U11/U12-31K (ZCRB1) (Figure 4C, Supplementary Table S4). This interaction was dependent on a functional RBM41–U12 snRNA interaction as the construct encoding the RRM-only version of MAC-RBM41-(259–413) was still able to support the in-

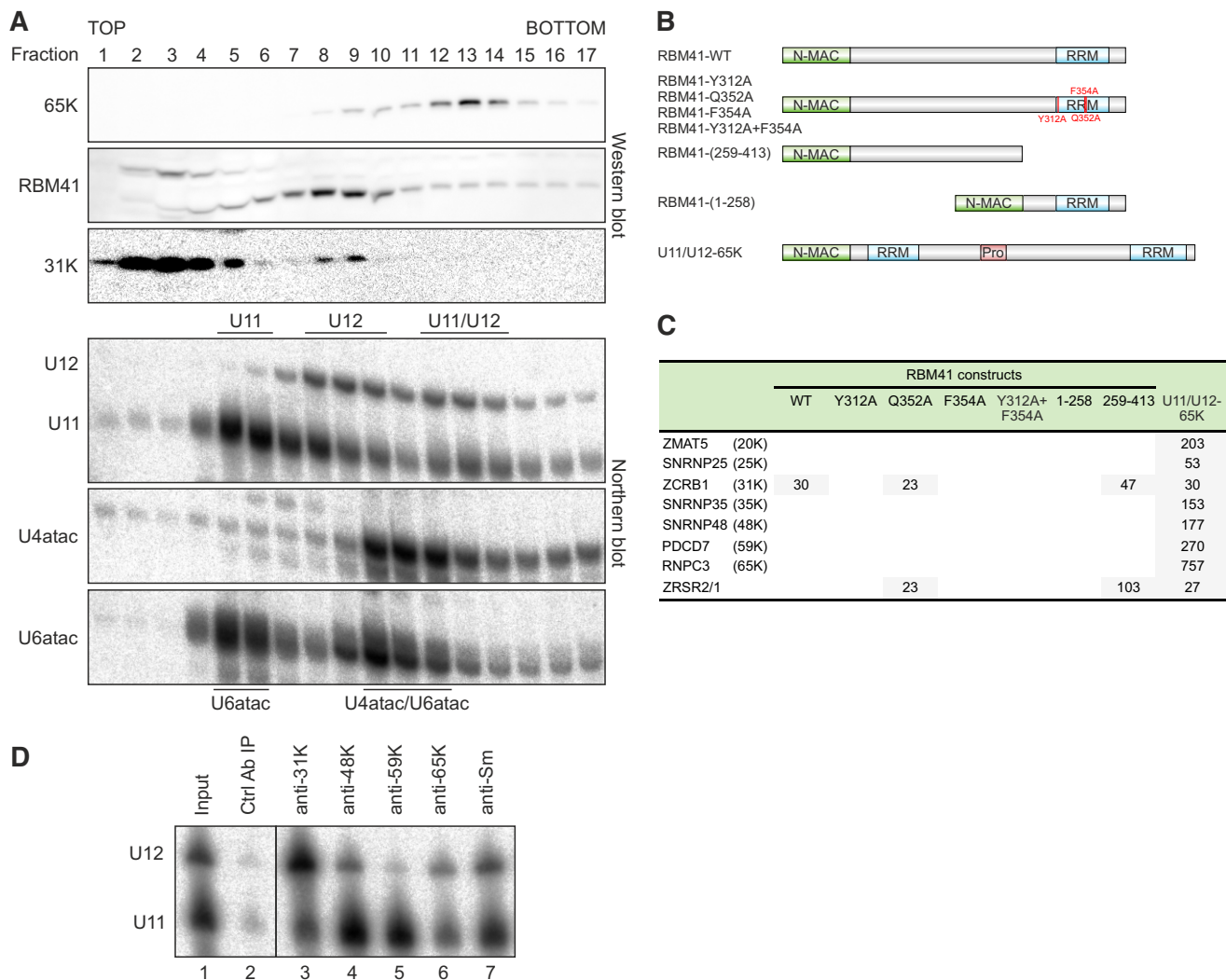
teraction with U11/U12-31K while the Y312A, F354A and Y312A + F354A RRM mutants and the RRM deletion construct (MAC-RBM41-(1–258)) did not. The Q352A mutant likely retains some residual binding for U12 snRNP, as the BioID signal with U11/U12-31K was reduced but not completely eliminated (see also Figure 5E). Additionally, the loss of the RBM41 N-terminal domain resulted in a strong interaction with ZRSR2/1, which functions in 3' ss recognition of U12-type introns.

A further support for the direct or indirect interaction between RBM41 and U11/U12-31K within the U12 mono-snRNP is obtained from glycerol gradient centrifugation which shows that RBM41 and U11/U12-31K proteins cosediment in the same U12 mono-snRNP fractions (Figure 4A, lanes 8–9). Similarly, co-IP experiments show that the endogenous U11/U12-31K is preferentially associated with U12 snRNA, unlike the U11/U12-65K which shows an even co-IP efficiency for both U11 and U12 snRNAs (Figure 4D, cf. lanes 3 and 6). Association of the U11/U12-31K with the mono-snRNP is consistent with direct recognition of 2'-O-methylated A8 residue of the U12 snRNA (60). Conversely, the inverted co-IP pattern with U11-48K and U11-59K is consistent with their role as components of both U11 mono-snRNP and the U11/U12 di-snRNP (Figure 4D, lanes 4 and 5). Together, our RIP, glycerol gradient and BioID data show that RBM41





**Figure 3.** RBM41 specifically associates with minor spliceosomal snRNPs. **(A)** RNA immunoprecipitation with V5-tagged RBM41 and 65K. V5-RBM41 or V5-65K expression vector or empty vector were transfected into HEK293 cells. 24 h later, RNA immunoprecipitation with anti-V5 antibody or control antibody was carried out in native conditions and co-immunoprecipitated RNA analyzed by northern blot using the indicated probes. **(B)** RNA immunoprecipitation with endogenous RBM41. RIP was carried out in native conditions in either HeLa nuclear extract (left) or HEK293 total lysate (right) using an antibody against endogenous RBM41 or control antibody. **(C)** V5-RBM41 constructs used for RNA immunoprecipitation in panel D. **(D)** Effect of truncations and RRM mutations on the snRNP association of RBM41. V5-tagged RBM41 constructs shown in C were transfected into HEK293 cells and RNA immunoprecipitation carried out using anti-V5 or control antibody.



**Figure 4.** RBM41 and U11/U12-65K partition into distinct snRNP complexes. **(A)** Glycerol gradient analysis of RBM41 and U11/U12-65K in HeLa nuclear extract. Nuclear extract was loaded on top of a 10–30% glycerol gradient. After ultracentrifugation, the gradient was fractionated, protein and RNA isolated and analyzed by western and northern blot using the antibodies and probes indicated on the left. Location of the U11, U12 and U6atac mono-snRNPs, U11/U12 di-snRNP and U4atac/U6atac di-snRNP are inferred based on the snRNA profiles. **(B)** Domain structures of MAC-tagged RBM41 and 65K constructs used for BioID. N-terminal MAC tag is not drawn to scale. **(C)** Spectral count fold changes for U11/U12 di-snRNP proteins in BioID datasets. **(D)** Immunoprecipitation of U11 and U12 snRNAs by anti-31K, anti-48K, anti-59K, anti-65K and anti-Sm antibodies in HEK293 total lysate followed by Northern blot detection of the U11 and U12 snRNAs.

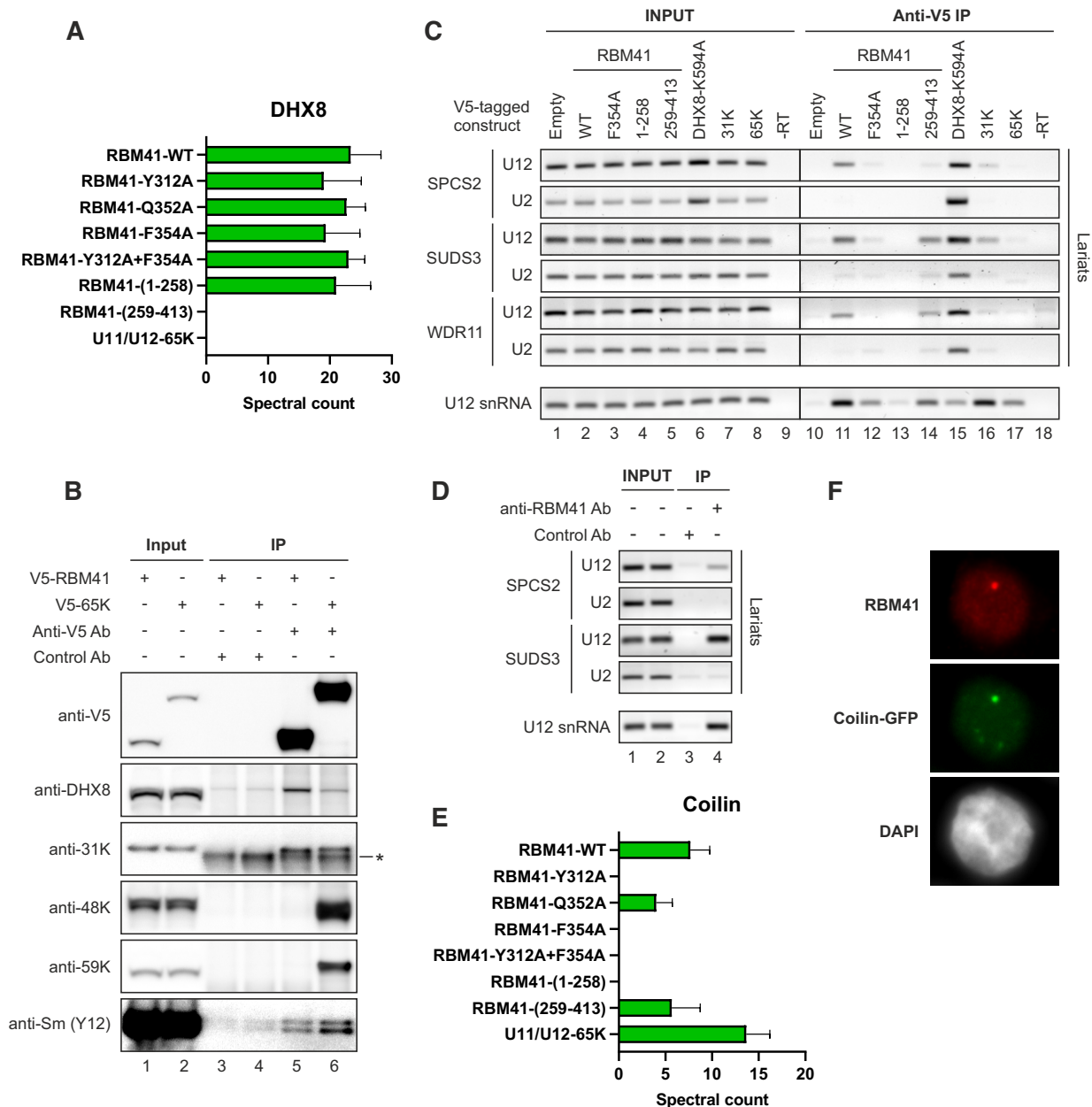
and U11/U12-65K partition into distinct snRNP complexes and are likely to play distinct roles within the minor spliceosome.

### RBM41 interacts with DHX8 and localizes to Cajal bodies

One of the top BioID proximity-labeling hits of MAC-RBM41-WT was the DEAH-box RNA helicase DHX8 (Figure 5A, Supplementary Table S4). The interaction with DHX8 was dependent on the highly conserved N-terminal region of RBM41, while the RRM mutations and deletion of the RRM had no effect. In contrast, no interaction of U11/U12-65K with DHX8 was detected in BioID analysis. The DHX8:RBM41 interaction was also reported in a recently published reciprocal DHX8 BioID dataset (61). To validate these findings, we carried out co-immunoprecipitation experiments in Flp-In™ T-REx™ 293 cell lines expressing V5-RBM41 or V5-65K (Figure 5B). Consistent with BioID, V5-

RBM41 co-immunoprecipitated DHX8, while close to background levels of DHX8 were detected in V5-65K immunoprecipitates. Furthermore, both V5-RBM41 and V5-65K co-immunoprecipitated the U11/U12-31K protein at similar levels, consistent with U11/U12-31K associating with both the U12 mono-snRNP and the U11/U12 di-snRNP. Although not detected in BioID, Sm proteins were detected in both V5-RBM41 and V5-65K immunoprecipitates. In contrast, the U11 snRNP and U11/U12 di-snRNP-associated U11-48K and U11-59K proteins were only detected in V5-65K IPs.

DHX8 has not been shown to function in the minor spliceosome. In the major spliceosome, it is recruited before exon ligation, during the transition from C to C\* complex (62), and drives the P-to-ILS1 transition presumably by pulling on the ligated exon, leading to its release and dissociation of at least nine proteins (63). We thus hypothesized that RBM41 could have a function in the late stages of minor splicing, possibly in the post-catalytic complexes. To test the association of RBM41 and DHX8 with late-stage minor spliceosomes, we



**Figure 5.** RBM41 interacts with DHX8 and localizes to Cajal bodies. **(A)** Spectral counts for DHX8 in RBM41 and U11/U12-65K BioID datasets. **(B)** Immunoprecipitation with anti-V5 or control antibody followed by western blot in Flp-In™ T-REx™ 293 cell lines expressing V5-RBM41 or V5-65K. The asterisk indicates a non-specific band detected in both control and anti-31K IPs and likely represents cross-reaction of the anti-rabbit secondary antibody with light chain from the IP antibody. **(C)** RNA immunoprecipitation with exogenously expressed V5-tagged proteins followed by RT-PCR. The indicated pCl-neo constructs for expressing V5-tagged proteins or empty pCl-neo vectors were transfected into HEK293 cells. 24 h later, RIP was carried out using anti-V5 antibody and RNA extracted from the beads analyzed by RT-PCR. Amplification across the branch junction was used to detect U2- and U12-type intron lariats and lariat intermediates from the following introns: SPCS2 introns 3–4 (U12) and 2–3 (U2), SUDS3 introns 7–8 (U12) and 9–10 (U2), WDR11 introns 28–29 (U12) and 27–28 (U2). **(D)** RNA immunoprecipitation with endogenous RBM41 in HEK293 cells followed by RT-PCR. **(E)** Spectral counts for coilin in RBM41 and U11/U12-65K BioID datasets. **(F)** Anti-RBM41 immunofluorescence in HEK293 cells transfected with a vector for expressing coilin-GFP.

carried out a RIP experiment in HEK293 cells transfected with V5-RBM41, V5-RBM41 mutant and truncation constructs (F354A, 1–258 and 259–413), or V5-tagged DHX8 carrying a helicase mutation (K594A) that stalls splicing in the P complex stage (14,64; Figure 5C). V5-65K and V5-31K were analyzed as controls for U11/U12 proteins. As DHX8 associates with post-branching spliceosomes, we used RT-PCR across branch sites (65,66) to detect U2-type and U12-type lariat intermediates and excised intron lariats in the immunoprecipitates.

V5-DHX8-K594A co-immunoprecipitated both U2-type and U12-type lariats from the same genes (Figure 5C, lane 15), suggesting that DHX8 functions in both spliceosomes. In contrast, RBM41 preferentially co-immunoprecipitated U12-type but not U2-type intron lariats (Figure 5C, lane 11). Similar results were obtained when RIP was carried out using an antibody against the endogenous RBM41 protein (Figure 5D). While U11/U12-65K did not co-IP any lariats (Figure 5C, lane 17), the U11/U12-31K showed variable and low levels of

lariat co-IP (Figure 5C, lane 16), suggesting that RBM41 BioID hits for DHX8 and U11/U12-31K may be originating from two separate complexes.

Another major interactor of MAC-RBM41-WT in our BioID data was coilin (Figure 5E, Supplementary Table S4), a key scaffolding protein and widely used marker for Cajal bodies. Consistently, we found that endogenous RBM41 localizes to Cajal bodies, as shown by colocalization with coilin-GFP in HEK293 cells (Figure 5F). While U11/U12-65K also interacted with coilin, the nuclear bodies labeled by the anti-RBM41 antibody did not colocalize with endogenous U11/U12-65K (Supplementary Figure S5B). This suggests that U11/U12-65K enters Cajal bodies only transiently. The coilin:RBM41 interaction was dependent on U12 snRNA binding, as mutating or deleting the RBM41 RRM reduced or completely eliminated the interaction, while MAC-RBM41-(259–413) was still able to interact with coilin (Figure 5E). Similarly, anti-HA immunofluorescence staining in BioID cell lines detected nuclear bodies in cells expressing MAC-RBM41-WT, but not in cells expressing any of the RBM41 RRM mutants or MAC-RBM41-(1–258) (Supplementary Figure S5A). This suggests that RBM41 localizes to Cajal bodies in an RNA binding-dependent manner.

### RBM41 is not essential for cell viability but affects the splicing of a subset of U12-type introns

To assess the effect of RBM41 loss on the splicing of the U12-type introns, we generated several independent RBM41 full knockout lines with HEK293 cells using CRISPR-Cas9 editing. The loss of functional RBM41 loci in each of the three X chromosomes in HEK293 cells was confirmed by Sanger sequencing and western blot analyses (Figure 6A; Supplementary Figure S6). The knockout cells were not only viable, but the loss of RBM41 did not lead to any noticeable growth phenotypes. This is consistent with the earlier investigations on human essentialomes that consistently indicated that RBM41 locus is not essential for the cell viability (67,68).

To analyze the effects of RBM41 knockout on splicing we carried out RNAseq analysis of three independent knockout cell lines and matching unedited lines. Subsequent bioinformatics analysis concentrated on U12-type intron retention and alternative/cryptic splice site activation with the U12-type intron containing genes. Intron retention analysis revealed weak splicing defects in the knockout cell lines for a small subset of genes (14 genes), such as the *NOL11* (Figure 6C; Supplementary Table S6), but also identified 15 genes that instead showed the opposite, that is, a reduction in the read levels mapping to the U12-type introns (Supplementary Table S6). Additionally, we investigated alternative splice donor (AD), splice acceptor (AA) and core exon (CE) usage. We further focused on the events within the U12-type introns and the surrounding exons and introns as these are the potential direct targets of RBM41 knockdown. We detected a total of 37 statistically significant alternative splicing events in 26 genes, as several genes showed multiple AS events being activated as a result of RBM41 knockout (Supplementary Table S7). Of these, the most notable were the ~3-fold enrichment in alternative 3'ss (AA) usage (Figures 6B–D;  $P = 3.4 \times 10^{-6}$ , hypergeometric test) and the 1.7-fold reduction in core exon (CE) events (Figures 6B–D;  $P = 1.2 \times 10^{-4}$ , hypergeometric test) when compared to the alternative splicing events detected in genes containing only major introns. Notably, most of the

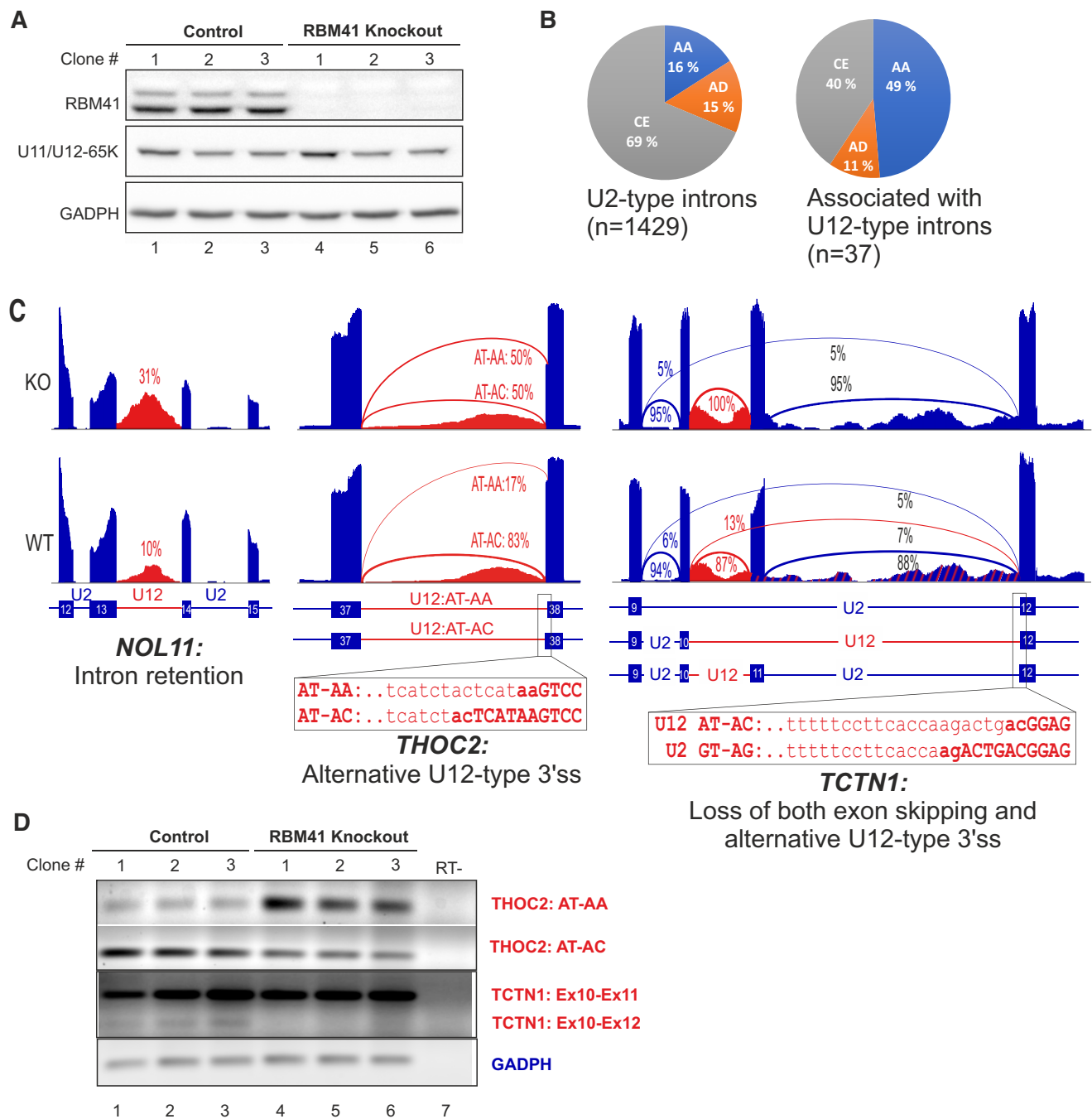
identified alternative 3'ss events (13/18) affected the U12-type intron 3'ss choice (Supplementary Table S6), suggesting that the loss of RBM41 has a weak, but nevertheless statistically significant effect on the splicing of a subset of U12-type introns.

## Discussion

In this work, we have expanded the repertoire of unique protein components specific to minor spliceosome by providing evidence that RBM41 functions in post-splicing steps of the minor spliceosome assembly/disassembly cycle. RBM41 shows a similar phylogenetic co-evolution pattern as several other minor spliceosome components (Figure 1E) and it has earlier been annotated as a paralog of the U11/U12-65K protein, due to the highly similar C-terminal RRMs found in the two proteins. Here, we show that the C-terminal RRM of RBM41 binds to the 3'-terminal stem-loops of U12 and U6atac snRNAs both *in vitro* and *in vivo*. Compared to the U11/U12-65K C-terminal RRM, RBM41 has approximately 2× lower affinity to its RNA ligands. We further show that unlike U11/U12-65K, which is a component of the U11/U12 di-snRNP, RBM41 associates with a distinct U12 mono-snRNP. Both U12 mono-snRNP and U11/U12 di-snRNP complexes have been described previously (11,20,69), but the function or composition of the U12 mono-snRNP has not been studied further. Here, our ultracentrifugation and BioID analysis provides evidence that the U12 mono-snRNP is a distinct functional complex in the minor spliceosome and contains, in addition to RBM41, the U11/U12-31K (ZCRB1) protein as the specific protein components. Additionally, we show that RBM41 associates specifically with excised U12-type intron lariats and uses its unique N-terminal domain to interact with the DHX8 helicase, and likely cycles through the Cajal bodies. Together, our data suggests that the two paralogous proteins have distinct functions in U12-type intron splicing with U11/U12-65K functioning in the early steps of U12-type intron recognition and RBM41 in the post-splicing steps and during minor spliceosome disassembly (Figure 7).

Our results highlight the role of the 3' stem-loop of U12 snRNA in the minor spliceosome assembly-disassembly cycle. The significance of the 5' end of the U12 snRNA has long been recognized due to its function in the BPS recognition and the interactions with the U6atac snRNA in the catalytic core of the minor spliceosome (70). In contrast, the 3'-terminal stem-loop of the U12 snRNA has appeared as a static binding site for the U11/U12-65K protein, necessary for the formation of the U11/U12 di-snRNP. Our identification of RBM41 binding to the 3'-terminal stem-loop during minor spliceosome disassembly suggests more dynamic recognition events where the 3'-terminal stem-loop serves as a platform for the two paralogous proteins which guide the U12 snRNA through the minor spliceosome assembly and disassembly cycle. The previously characterized steps include the recognition of the 3'-terminal stem-loop of the U12 snRNA by the U11/U12-65K protein, which uses its N-terminus to interact with the U11-59K protein (22) to form the U11/U12 di-snRNP, which in turn is necessary for the U12-type intron recognition (12). Furthermore, during the formation of the catalytically active spliceosome ( $B^{act}$  complex) the U11/U12-65K protein remain attached to the  $B^{act}$  complex (presumably to the 3'-terminal stem-loop), while U11 snRNP and all the other specific protein components of the di-snRNP are released from the

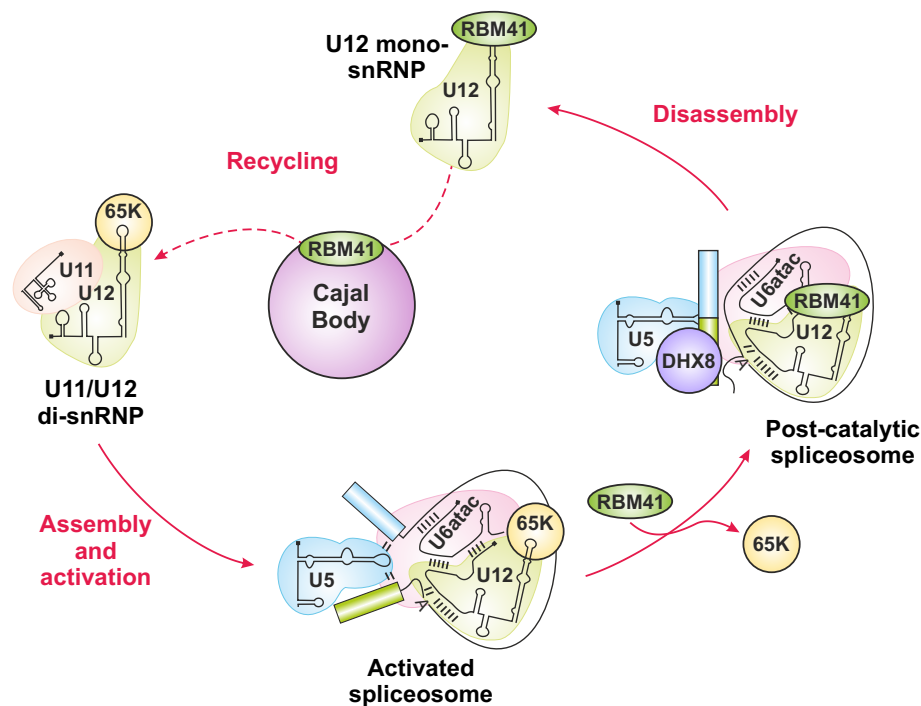




**Figure 6.** RBM41 knockout influences the splicing of U12-type introns. **(A)** Western blot analysis of RBM41 knockout and matching control cell lines used in the RNAseq analysis. **(B)** Comparison of the statistically significant (Whippet Probability > 0.9) alternative splicing events in the genes containing only U2-type introns and events either within or near proximity (immediate up- or downstream exons and introns) of the U12-type introns. AA - alternative acceptor, AD - alternative donor, CE - core exon. **(C)** Representative sashimi plots showing Intron retention (*NOL11*), Alternative U12-type 3'ss choice (*THOC2*) and loss of both exon skipping and alternative U12-type 3'ss usage in RBM41 knockout cells (*TCTN1*). The percentages refer to the intron retention levels (*NOL11*), the alternative 3' splice usage levels (*THOC2*) or exon skipping levels (*TCTN1*) as indicated by the arches in the Sashimi plot. **(D)** Validation of the *THOC2* and *TCTN1* alternative splicing changes using a set of three independent RBM41 knockout cell lines and their matching controls.

activated spliceosome (8). Our data indicate that later in the splicing process there is an exchange in the 3'-terminal stem-loop binding partner from U11/U12-65K to RBM41 which can be detected in post-splicing complexes containing excised minor intron lariats, and which is also in close proximity with the DHX8/hPrp22 helicase. Together these results suggest that RBM41 is present in the minor spliceosome post catalytic (P) and intron lariat spliceosome (ILS) complexes. Fur-

thermore, proximity labeling of Cajal body marker coilin (Figure 5E; (61) and co-localization of RBM41 and coilin in Cajal bodies (Figure 5F, Supplementary Figure S5) suggests that RBM41 remains bound to the U12 snRNA during the entire spliceosome disassembly process and follows the snRNA to the Cajal body. This is the likely location for a new round of assembly of the U11/U12 di-snRNP complex. The association of U11/U12-31K with both the U12 mono-snRNP and



**Figure 7.** A model of the dynamic exchanges between RBM41 and 65K binding to U12 snRNA during the splicing cycle. In the U11/U12 di-snRNP, the 3'-terminal stem-loop is bound by U11/U12-65K, which mediates the connection between the U11 and U12 snRNPs. U11/U12-65K likely remains bound to the stem-loop throughout minor spliceosome assembly and activation but is exchanged for RBM41 during or after the catalytic steps of splicing. After spliceosome disassembly, RBM41 remains bound to the post-spliceosomal U12 mono-snRNP and accompanies it to Cajal bodies. During U11/U12 di-snRNP recycling, which likely takes place in Cajal bodies, RBM41 is again replaced by U11/U12-65K at the 3'-terminal stem-loop.

U11/U12 di-snRNP (Figures 4A, C, D) suggests that it remains bound with U12 snRNA (60) in the new round of U11/U12 di-snRNP assembly initiated by the exchange from RBM41 to U11/U12-65K.

RBM41 and U11/U12-65K proteins interact with both U12 and U6atac 3' terminal stem-loops *in vitro* and *in vivo*. However, the U12 interactions appear more significant, given the sensitivity of U12-type intron splicing to mutations in the U12 single-stranded loop and the insensitivity to U6atac loop mutations (27,71). Furthermore, an 84C > U mutation that compromises the U12 3'-terminal stem-loop integrity leads to early onset cerebellar ataxia due to overtrimming of the 3'-terminal stem-loop which removes the binding site of the U11/U12-65K and RBM41 proteins (36,72). Similarly, the U11/U12-65K P474T mutation associated with isolated growth hormone deficiency has been shown to reduce the U11/U12 di-snRNP levels due to a folding defect of the U11/U12-65K C-RRM, which reduces its affinity to the 3'-terminal stem-loop (24,25). However, in that case the potential additional effects on U6atac binding *in vivo* or on the recycling of the U12 snRNA have not been ruled out.

Our data portrays a somewhat conflicting view on the significance of the RBM41 and the need for a specific protein factor(s) for the minor spliceosome disassembly process. The strong sequence conservation observed with the domains of RBM41 that interact either with the U12 snRNA or DHX8 (Supplementary Figure S2) suggests a strong selection pressure at the organismal level to maintain these interactions. This is reflected in phylogenetic co-occurrence in multiple evolutionary lineages (Figure 1E), though we note that the locus encoding RBM41 is more frequently absent in multiple evolutionary lineages than several other components of the minor spliceo-

some. On the other hand, both our knockout data (Figure 6A) and data from essentialomes (67,68) indicate that RBM41 is dispensable at least at the cellular level. Based on the weak, yet statistically significant effects of RBM41 knockout in human HEK293 cells specifically on 3'ss selection of U12-type introns (Figure 6B–D), we hypothesize that while RBM41 is dispensable at the cellular level, it may nevertheless be able to exert a weak kinetic effect on splicing in addition to later participating in the disassembly process. The effect on 3'ss choice is similar to that observed after major spliceosome catalytic step II factor knockdowns, which similarly influence the 3'ss choice, particularly with NAGNAG introns (73), further suggesting that the exchange from the U11/U12-65K to RBM41 may take place prior to step II. However, given that our BioID analysis did not provide supporting evidence for this possibility, it is also possible that the exchange from U11/U12-65K to RBM41 takes place at a later step and the effects on minor intron splicing are secondary effects of downstream processes being disturbed. Finally, while RBM41 protein is not absolutely needed for cell viability or splicing in the highly proliferative cell types used in essentialome and our knockout studies, it may provide selective advantage in specific cell types or in the context of whole organisms to account for the observed evolutionary conservation.

RBM41 may also play a role in substituting structures or interactions that are present in the major but not in the minor spliceosome. Specifically, the human minor B<sup>act</sup> complex lacks several key protein components that are present in the major B<sup>act</sup> complex. These include NTC complex proteins (PRPF19, SPF27 and SYF1), NTR complex proteins (BUD31 and RBM22), SF3a complex, and phosphoprotein isomerases (PP1L1 and CypE). Conversely, the minor B<sup>act</sup>

complex contains four unique proteins RBM48, ARMC7, SCNM1 and CRIPT that are not present in the major B<sup>act</sup> complex (8,30,31). At least a subset of these differences in protein composition can be explained by the presence of distinct snRNA components in the two spliceosomes which influences the availability of the specific RNA elements for the binding of protein factors. For example, U6atac snRNA lacks the 5' stem-loop structure which is present in U6 snRNA and recognized by the BUD31 and RBM22 proteins in major B<sup>act</sup> complex. Instead, in minor B<sup>act</sup> these proteins are replaced by RBM48 and ARMC7 proteins binding to the  $\gamma$ -monomethyl cap of U6atac snRNA (8). Similarly, as the U11/U12 di-snRNP lacks the SF3a complex (19) that is present both in the 17S U2 snRNP and the major B<sup>act</sup> complex (74,75), the functions of SF3a in minor B<sup>act</sup> has been taken over by the SCNM1, a unique protein component of the minor spliceosome (8). In this respect the absence of CypE and SYF1 from the minor B<sup>act</sup> (8) is intriguing as those are loaded to major B<sup>act</sup> and are later involved in post-spliceosomal complex transitions (63,64). From the available limited structural information, we hypothesize that there may be similar differences in the post-spliceosomal complex composition between the two spliceosomes such that RBM41 may substitute some of the major spliceosome structures and /or interactions that are missing from the minor spliceosome. Testing of this hypothesis would require similar high-resolution structures of the post-splicing minor spliceosome complexes that are available for the major spliceosome.

Unique protein components in the minor spliceosome can also offer potential opportunities for differential regulation of the two spliceosomes. Given the function of RBM41 in the post-splicing complexes, the regulation of minor intron splicing via RBM41 abundance or activity is unlikely. However, earlier work has nevertheless provided evidence of translational regulation of RBM41 via differential 3' UTR isoform usage (76), particularly during neuronal development (77). As the abundance of U11/U12-65K is also regulated by a feedback/cross-regulation pathway, particularly during neuronal differentiation (78,79), it is possible that the cellular levels of RBM41 protein are similarly linked to the abundance of minor spliceosome components or to global regulation of mRNA processing pathways.

## Data availability

The RNAseq data generated in this study is available at NCBI Gene Expression Omnibus (GEO) with accession number GSE246283 and the mass spectrometry data at MassIVE with dataset ID MSV000092983.

## Supplementary data

Supplementary Data are available at NAR Online.

## Acknowledgements

We thank the members of the Frilander lab for suggestions during this work and Marja-Leena Peltonen for technical assistance. We acknowledge the assistance of University of Helsinki Life Science Research Infrastructures, specifically, Viikki DNA Sequencing unit, Viikki Proteomics unit, Institute of Biotechnology Light Microscopy unit and Instruct-HiLIFE

Biocomplex unit (a member of Instruct-ERIC Centre Finland, FINStruct and Biocenter Finland), during this work.

## Funding

Sigrid Jusélius Foundation [to M.J.F and M.V.]; Academy of Finland [308657 and 1341477 to M.J.F.]; Jane and Aatos Erkko Foundation [to M.J.F.]; Biocentrum Finland and Helsinki Institute of Life Sciences (HiLIFE) infrastructure funding [to M.V.]; Netherlands Organisation for Scientific Research (NWO) [VICI program, grant 016.160.638 to B.S.]; A.J.N. was supported by the Integrative Life Science doctoral program at the University of Helsinki. Funding for open access charge: Helsinki University Library.

## Conflict of interest statement

None declared.

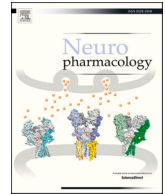
## References

- Hall,S.L. and Padgett,R.A. (1996) Requirement of U12 snRNA for in vivo splicing of a minor class of eukaryotic nuclear pre-mRNA introns. *Science*, **271**, 1716–1718.
- Tarn,W.-Y. and Steitz,J.A. (1996) Highly diverged U4 and U6 small nuclear RNAs required for splicing rare AT-AC introns. *Science*, **273**, 1824–1832.
- Tarn,W.-Y. and Steitz,J.A. (1996) A novel spliceosome containing U11, U12 and U5 snRNPs excises a minor class (AT-AC) intron *in vitro*. *Cell*, **84**, 801–811.
- Kolossova,I. and Padgett,R.A. (1997) U11 snRNA interacts in vivo with the 5' splice site of U12-dependent (AU-AC) pre-mRNA introns. *RNA*, **3**, 227–233.
- Yu,Y.-T. and Steitz,J.A. (1997) Site-specific crosslinking of mammalian U11 and U6atac to the 5' splice site of an AT-AC intron. *Proc. Natl. Acad. Sci. U.S.A.*, **94**, 6030–6035.
- Hastings,M.L. and Krainer,A.R. (2001) Functions of SR proteins in the U12-dependent AT-AC pre-mRNA splicing pathway. *RNA*, **7**, 471–482.
- Wilkinson,M.E., Charenton,C. and Nagai,K. (2020) RNA Splicing by the Spliceosome. *Annu. Rev. Biochem.*, **89**, 359–388.
- Bai,R., Wan,R., Wang,L., Xu,K., Zhang,Q., Lei,J. and Shi,Y. (2021) Structure of the activated human minor spliceosome. *Science*, **371**, eabg0879.
- Zhao,J., Peter,D., Brandina,I., Liu,X. and Galej,W.P. (2023) Structure of the minor spliceosomal U11 snRNP. bioRxiv doi: <https://doi.org/10.1101/2023.12.22.573053>, 22 December 2023, preprint: not peer reviewed.
- Frilander,M.J. and Steitz,J.A. (2001) Dynamic exchanges of RNA interactions leading to catalytic core formation in the U12-dependent spliceosome. *Mol. Cell*, **7**, 217–226.
- Wassarman,K.M. and Steitz,J.A. (1992) The low-abundance U11 and U12 small nuclear ribonucleoproteins (snRNPs) interact to form a two-snRNP complex. *Mol. Cell Biol.*, **12**, 1276–1285.
- Frilander,M.J. and Steitz,J.A. (1999) Initial recognition of U12-dependent introns requires both U11/5' splice-site and U12/branchpoint interactions. *Genes Dev.*, **13**, 851–863.
- Company,M., Arenas,J. and Abelson,J. (1991) Requirement of the RNA helicase-like protein PRP22 for release of messenger RNA from spliceosomes. *Nature*, **349**, 487–493.
- Ohno,M. and Shimura,Y. (1996) A human RNA helicase-like protein, HRH1, facilitates nuclear export of spliced mRNA by releasing the RNA from the spliceosome. *Genes Dev.*, **10**, 997–1007.
- Arenas,J.E. and Abelson,J.N. (1997) Prp43: an RNA helicase-like factor involved in spliceosome disassembly. *Proc. Natl. Acad. Sci. U.S.A.*, **94**, 11798–11802.

16. Staněk, D. (2017) Cajal bodies and snRNPs - friends with benefits. *RNA Biol*, **14**, 671–679.
17. Damianov, A., Schreiner, S. and Bindereif, A. (2004) Recycling of the U12-type spliceosome requires p110, a component of the U6atac snRNP. *Mol. Cell. Biol*, **24**, 1700–1708.
18. Bell, M., Schreiner, S., Damianov, A., Reddy, R. and Bindereif, A. (2002) p110, a novel human U6 snRNP protein and U4/U6 snRNP recycling factor. *EMBO J*, **21**, 2724–2735.
19. Will, C.L., Schneider, C., Reed, R. and Lührmann, R. (1999) Identification of both shared and distinct proteins in the major and minor spliceosomes. *Science*, **284**, 2003–2005.
20. Will, C.L., Schneider, C., Hossbach, M., Urlaub, H., Rauhut, R., Elbashir, S., Tuschl, T. and Lührmann, R. (2004) The human 18S U11/U12 snRNP contains a set of novel proteins not found in the U2-dependent spliceosome. *RNA*, **10**, 929–941.
21. Schneider, C., Will, C.L., Makarova, O.V., Makarov, E.M. and Lührmann, R. (2002) Human U4/U6.U5 and U4atac/U6atac.U5 tri-snRNPs exhibit similar protein compositions. *Mol. Cell. Biol*, **22**, 3219–3229.
22. Benecke, H., Lührmann, R. and Will, C.L. (2005) The U11/U12 snRNP 65K protein acts as a molecular bridge, binding the U12 snRNA and U11-59K protein. *EMBO J*, **24**, 3057–3069.
23. Turunen, J.J., Will, C.L., Grote, M., Lührmann, R. and Frilander, M.J. (2008) The U11-48K protein contacts the 5' splice site of U12-type introns and the U11-59K protein. *Mol. Cell. Biol*, **28**, 3548–3560.
24. Argente, J., Flores, R., Gutierrez-Arumi, A., Verma, B., Martos-Moreno, G.A., Cusco, I., Oghabian, A., Chowen, J.A., Frilander, M.J. and Perez-Jurado, L.A. (2014) Defective minor spliceosome mRNA processing results in isolated familial growth hormone deficiency. *EMBO Mol. Med.*, **6**, 299–306.
25. Norppa, A.J., Kauppala, T., Heikkinen, H.A., Verma, B., Iwai, H. and Frilander, M.J. (2018) Mutations in the U11/U12-65K protein associated with isolated growth hormone deficiency lead to structural destabilization and impaired binding of U12 snRNA. *RNA*, **24**, 396–409.
26. Netter, C., Weber, G., Benecke, H. and Wahl, M.C. (2009) Functional stabilization of an RNA recognition motif by a noncanonical N-terminal expansion. *RNA*, **15**, 1305–1313.
27. Singh, J., Sikand, K., Conrad, H., Will, C.L., Komar, A.A. and Shukla, G.C. (2016) U6atac snRNA stem-loop interacts with U12 p65 RNA binding protein and is functionally interchangeable with the U12 apical stem-loop III. *Sci. Rep.*, **6**, 31393.
28. Madan, V., Kanojia, D., Li, J., Okamoto, R., Sato-Otsubo, A., Kohlmann, A., Sanada, M., Grossmann, V., Sundaresan, J., Shiraiishi, Y., et al. (2015) Aberrant splicing of U12-type introns is the hallmark of ZRSR2 mutant myelodysplastic syndrome. *Nat. Commun.*, **6**, 6042.
29. Shen, H., Zheng, X., Luecke, S. and Green, M.R. (2010) The U2AF35-related protein Urp contacts the 3' splice site to promote U12-type intron splicing and the second step of U2-type intron splicing. *Genes Dev.*, **24**, 2389–2394.
30. Bai, F., Corll, J., Shodja, D.N., Davenport, R., Feng, G., Mudunkothge, J., Brigolin, C.J., Martin, F., Spielbauer, G., Tseung, C.-W., et al. (2019) RNA binding motif protein 48 is required for U12 splicing and maize endosperm differentiation. *Plant Cell*, **31**, 715.
31. Siebert, A.E., Corll, J., Paige Gronevelt, J., Levine, L., Hobbs, L.M., Kenney, C., Powell, C.L.E., Battistuzzi, F.U., Davenport, R., Mark Settles, A., et al. (2022) Genetic analysis of human RNA binding motif protein 48 (RBM48) reveals an essential role in U12-type intron splicing. *Genetics*, **222**, iyac129.
32. de Wolf, B., Oghabian, A., Akinyi, M.V., Hanks, S., Tromer, E.C., van Hooff, J.J.E., van Voorthuisen, L., van Rooijen, L.E., Verbeeren, J., Uijttewaai, E.C.H., et al. (2021) Chromosomal instability by mutations in the novel minor spliceosome component CENATAC. *EMBO J*, **40**, e106536.
33. Suzuki, T., Shinagawa, T., Niwa, T., Akeda, H., Hashimoto, S., Tanaka, H., Hiroaki, Y., Yamasaki, F., Mishima, H., Kawai, T., et al. (2022) The DROL1 subunit of U5 snRNP in the spliceosome is specifically required to splice AT-AC-type introns in Arabidopsis. *Plant J*, **109**, 633–648.
34. Liu, X., Salokas, K., Weldatsadik, R.G., Gawryski, L. and Varjosalo, M. (2020) Combined proximity labeling and affinity purification-mass spectrometry workflow for mapping and visualizing protein interaction networks. *Nat. Protoc*, **15**, 3182–3211.
35. Bloh, K., Kanchana, R., Bialk, P., Banas, K., Zhang, Z., Yoo, B.C. and Kmiec, E.B. (2021) Deconvolution of complex DNA repair (DECODR): establishing a novel deconvolution algorithm for comprehensive analysis of CRISPR-edited Sanger sequencing data. *CRISPR J*, **4**, 120–131.
36. Norppa, A.J. and Frilander, M.J. (2021) The integrity of the U12 snRNA 3' stem-loop is necessary for its overall stability. *Nucleic Acids Res.*, **49**, 2835–2847.
37. Tarn, W.-Y. and Steitz, J.A. (1994) SR proteins can compensate for the loss of U1 snRNP functions *in vitro*. *Genes Dev.*, **8**, 2704–2717.
38. Liu, X., Salokas, K., Tamene, F., Jiu, Y., Weldatsadik, R.G., Öhman, T. and Varjosalo, M. (2018) An AP-MS- and BioID-compatible MAC-tag enables comprehensive mapping of protein interactions and subcellular localizations. *Nat. Commun.*, **9**, 1188.
39. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T.R. (2012) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.
40. Sterne-Weiler, T., Weatheritt, R.J., Best, A.J., Ha, K.C.H. and Blencowe, B.J. (2018) Efficient and accurate quantitative profiling of alternative splicing patterns of any complexity on a laptop. *Mol. Cell*, **72**, 187–200.
41. Middleton, R., Gao, D., Thomas, A., Singh, B., Au, A., Wong, J.J.-L., Bomane, A., Cosson, B., Eyraes, E., Rasko, J.E.J., et al. (2017) IRFinder: assessing the impact of intron retention on mammalian gene expression. *Genome Biol.*, **18**, 51.
42. Lorenzi, C., Barriere, S., Arnold, K., Luco, R.F., Oldfield, A.J. and Ritchie, W. (2021) IRFinder-S: a comprehensive suite to discover and explore intron retention. *Genome Biol.*, **22**, 307.
43. Oghabian, A., Greco, D. and Frilander, M.J. (2018) IntERESt: intron-exon retention estimator. *BMC Bioinform.*, **19**, 130.
44. Moyer, D.C., Larue, G.E., Hershberger, C.E., Roy, S.W. and Padgett, R.A. (2020) Comprehensive database and evolutionary dynamics of U12-type introns. *Nucleic Acids Res.*, **48**, 7066–7078.
45. Olthof, A.M., Hyatt, K.C. and Kanadia, R.N. (2019) Minor intron splicing revisited: identification of new minor intron-containing genes and tissue-dependent retention and alternative splicing of minor introns. *BMC Genom.*, **20**, 686.
46. Deutekom, E.S., Snel, B. and van Dam, T.J.P. (2021) Benchmarking orthology methods using phylogenetic patterns defined at the base of Eukaryotes. *Brief. Bioinform.*, **22**, bbaa206.
47. Vosseberg, J., van Hooff, J.J.E., Marcet-Houben, M., van Vlimmeren, A., van Wijk, L.M., Gabaldón, T. and Snel, B. (2021) Timing the origin of eukaryotic cellular complexity with ancient duplications. *Nat. Ecol. Evol.*, **5**, 92–100.
48. Emms, D.M. and Kelly, S. (2019) OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.*, **20**, 238.
49. Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S.K., Cook, H., Mende, D.R., Letunic, I., Rattei, T., Jensen, L.J., et al. (2019) eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.*, **47**, D309–D314.
50. Mistry, J., Finn, R.D., Eddy, S.R., Bateman, A. and Punta, M. (2013) Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Res.*, **41**, e121.
51. Katoh, K., Kuma, K.-i., Toh, H. and Miyata, T. (2005) MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.*, **33**, 511–518.
52. Nguyen, L.-T., Schmidt, H.A., von Haeseler, A. and Minh, B.Q. (2015) IQ-TREE: a fast and effective stochastic algorithm for



- estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.*, **32**, 268–274.
53. Vosseberg, J., Stolker, D., von der Dunk, S.H.A. and Snel, B. (2023) Integrating phylogenetics with intron positions illuminates the origin of the complex spliceosome. *Mol. Biol. Evol.*, **40**, msd011.
  54. Muto, Y. and Yokoyama, S. (2012) Structural insight into RNA recognition motifs: versatile molecular Lego building blocks for biological systems. *WIREs RNA*, **3**, 229–246.
  55. DeKoster, G.T., Delaney, K.J. and Hall, K.B. (2014) A compare-and-contrast NMR dynamics study of two related RRM: U1A and SNF. *Biophys. J.*, **107**, 208–219.
  56. Weber, G., DeKoster, G.T., Holton, N., Hall, K.B. and Wahl, M.C. (2018) Molecular principles underlying dual RNA specificity in the *Drosophila* SNF protein. *Nat. Commun.*, **9**, 2220.
  57. Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., *et al.* (2022) AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.*, **50**, D439–D444.
  58. Ray, D., Kazan, H., Cook, K.B., Weirauch, M.T., Najafabadi, H.S., Li, X., Gueroussov, S., Albu, M., Zheng, H., Yang, A., *et al.* (2013) A compendium of RNA-binding motifs for decoding gene regulation. *Nature*, **499**, 172–177.
  59. Dominguez, D., Freese, P., Alexis, M.S., Su, A., Hochman, M., Palden, T., Bazile, C., Lambert, N.J., Van Nostrand, E.L., Pratt, G.A., *et al.* (2018) Sequence, structure, and context preferences of human RNA binding proteins. *Mol. Cell*, **70**, 854–867.
  60. Li, B., Liu, S., Zheng, W., Liu, A., Yu, P., Wu, D., Zhou, J., Zhang, P., Liu, C., Lin, Q., *et al.* (2024) RIP-PEN-seq identifies a class of kink-turn RNAs as splicing regulators. *Nat. Biotechnol.*, **42**, 119–131.
  61. Go, C.D., Knight, J.D.R., Rajasekharan, A., Rathod, B., Hesketh, G.G., Abe, K.T., Youn, J.Y., Samavarchi-Tehrani, P., Zhang, H., Zhu, L.Y., *et al.* (2021) A proximity-dependent biotinylation map of a human cell. *Nature*, **595**, 120–124.
  62. Zhan, X., Yan, C., Zhang, X., Lei, J. and Shi, Y. (2018) Structure of a human catalytic step I spliceosome. *Science*, **359**, 537–545.
  63. Zhang, X., Zhan, X., Yan, C., Zhang, W., Liu, D., Lei, J. and Shi, Y. (2019) Structures of the human spliceosomes before and after release of the ligated exon. *Cell Res.*, **29**, 274–285.
  64. Fica, S.M., Oubridge, C., Wilkinson, M.E., Newman, A.J. and Nagai, K. (2019) A human postcatalytic spliceosome structure reveals essential roles of metazoan factors for exon ligation. *Science*, **363**, 710–714.
  65. Lorsch, J.R., Bartel, D.P. and Szostak, J.W. (1995) Reverse transcriptase reads through a 2′–5′ linkage and a 2′-thiophosphate in a template. *Nucleic Acids Res.*, **23**, 2811–2814.
  66. Suzuki, H., Zuo, Y., Wang, J., Zhang, M.Q., Malhotra, A. and Mayeda, A. (2006) Characterization of RNase R-digested cellular RNA source that consists of lariat and circular RNAs from pre-mRNA splicing. *Nucleic Acids Res.*, **34**, e63.
  67. Blomen, V.A., Májek, P., Jae, L.T., Bigenzahn, J.W., Nieuwenhuis, J., Staring, J., Sacco, R., van Diemen, F.R., Olk, N., Stukalov, A., *et al.* (2015) Gene essentiality and synthetic lethality in haploid human cells. *Science*, **350**, 1092–1096.
  68. Yilmaz, A., Peretz, M., Aharoni, A., Sagi, I. and Benvenisty, N. (2018) Defining essential genes for human pluripotent stem cells by CRISPR–Cas9 screening in haploid cells. *Nat. Cell Biol.*, **20**, 610–619.
  69. Montzka, K.A. and Steitz, J.A. (1988) Additional low-abundance human small nuclear ribonucleoproteins: U11, U12 *etc.* *Proc. Natl. Acad. Sci. U.S.A.*, **85**, 8885–8889.
  70. Turunen, J.J., Niemelä, E.H., Verma, B. and Frilander, M.J. (2013) The significant other: splicing by the minor spliceosome. *WIREs RNA*, **4**, 61–76.
  71. Sikand, K. and Shukla, G.C. (2011) Functionally important structural elements of U12 snRNA. *Nucleic Acids Res.*, **39**, 8531–8543.
  72. Elsaid, M.F., Chalhoub, N., Ben-Omran, T., Kumar, P., Kamel, H., Ibrahim, K., Mohamoud, Y., Al-Dous, E., Al-Azwani, I., Malek, J.A., *et al.* (2017) Mutation in noncoding RNA RNU12 causes early onset cerebellar ataxia. *Ann. Neurol.*, **81**, 68–78.
  73. Dybkov, O., Preußner, M., El Ayoubi, L., Feng, V.-Y., Harnisch, C., Merz, K., Leupold, P., Yudichev, P., Agafonov, D.E., Will, C.L., *et al.* (2023) Regulation of 3′ splice site selection after step 1 of splicing by spliceosomal C\* proteins. *Sci. Adv.*, **9**, eadf1785.
  74. Bessonov, S., Anokhina, M., Krasauskas, A., Golas, M.M., Sander, B., Will, C.L., Urlaub, H., Stark, H. and Lührmann, R. (2010) Characterization of purified human Bact spliceosomal complexes reveals compositional and morphological changes during spliceosome activation and first step catalysis. *RNA*, **16**, 2384–2403.
  75. Schmidt, C., Grønborg, M., Deckert, J., Bessonov, S., Conrad, T., Lührmann, R. and Urlaub, H. (2014) Mass spectrometry–based relative quantification of proteins in pre-catalytic and catalytically active spliceosomes by metabolic labeling (SILAC), chemical labeling (iTRAQ), and label-free spectral count. *RNA*, **20**, 406–420.
  76. Sterne-Weiler, T., Martinez-Nunez, R.T., Howard, J.M., Cvitovik, J., Katzman, S., Tariq, M.A., Pourmand, N. and Sanford, J.R. (2013) Frac-seq reveals isoform-specific recruitment to polyribosomes. *Genome Res.*, **23**, 1615–1623.
  77. Yap, K., Xiao, Y., Friedman, B.A., Je, H.S. and Makeyev, E.V. (2016) Polarizing the neuron through sustained co-expression of alternatively spliced isoforms. *Cell Rep.*, **15**, 1316–1328.
  78. Verbeeren, J., Niemelä, E.H., Turunen, J.J., Will, C.L., Ravanti, J.J., Lührmann, R. and Frilander, M.J. (2010) An ancient mechanism for splicing control: U11 snRNP as an activator of alternative splicing. *Mol. Cell*, **37**, 821–833.
  79. Verbeeren, J., Verma, B., Niemelä, E.H., Yap, K., Makeyev, E.V. and Frilander, M.J. (2017) Alternative exon definition events control the choice between nuclear retention and cytoplasmic export of U11/U12-65K mRNA. *PLoS Genet.*, **13**, e1006824.
  80. Robert, X. and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.*, **42**, W320–W324.
  81. Luo, Y., Hitz, B.C., Gabdank, I., Hilton, J.A., Kagda, M.S., Lam, B., Myers, Z., Sud, P., Jou, J., Lin, K., *et al.* (2020) New developments on the Encyclopedia of DNA Elements (ENCODE) data portal. *Nucleic Acids Res.*, **48**, D882–D889.



# Quantitative T<sub>2</sub> mapping-based longitudinal assessment of brain injury and therapeutic rescue in the rat following acute organophosphate intoxication

Alita Jesal D. Almeida<sup>a,b,1</sup>, Brad A. Hobson<sup>c,1</sup>, Naomi Saito<sup>d</sup>, Donald A. Bruun<sup>e</sup>,  
Valerie A. Porter<sup>a,b</sup>, Danielle J. Harvey<sup>d</sup>, Joel R. Garbow<sup>f</sup>, Abhijit J. Chaudhari<sup>b,c,\*\*</sup>, Pamela  
J. Lein<sup>e,\*</sup>

<sup>a</sup> Department of Biomedical Engineering, University of California-Davis College of Engineering, Davis, CA, 95616, USA

<sup>b</sup> Department of Radiology, University of California-Davis School of Medicine, Sacramento, CA, 95817, USA

<sup>c</sup> Center for Molecular and Genomic Imaging, Department of Biomedical Engineering, University of California-Davis College of Engineering, Davis, CA, 95616, USA

<sup>d</sup> Department of Public Health Sciences, University of California-Davis School of Medicine, Davis, CA, 95616, USA

<sup>e</sup> Department of Molecular Biosciences, University of California-Davis School of Veterinary Medicine, Davis, CA, 95616, USA

<sup>f</sup> Department of Radiology, Washington University School of Medicine, St Louis, MO, 63110, USA

## ARTICLE INFO

Handling Editor: Prof D Boison

### Keywords:

Allopregnanolone  
Diisopropylfluorophosphate  
Magnetic resonance imaging  
Midazolam  
Neurosteroid  
T<sub>2</sub> mapping

## ABSTRACT

Acute intoxication with organophosphate (OP) cholinesterase inhibitors poses a significant public health risk. While currently approved medical countermeasures can improve survival rates, they often fail to prevent chronic neurological damage. Therefore, there is need to develop effective therapies and quantitative metrics for assessing OP-induced brain injury and its rescue by these therapies. In this study we used a rat model of acute intoxication with the OP, diisopropylfluorophosphate (DFP), to test the hypothesis that T<sub>2</sub> measures obtained from brain magnetic resonance imaging (MRI) scans provide quantitative metrics of brain injury and therapeutic efficacy. Adult male Sprague Dawley rats were imaged on a 7T MRI scanner at 3, 7 and 28 days post-exposure to DFP or vehicle (VEH) with or without treatment with the standard of care antiseizure drug, midazolam (MDZ); a novel antiseizure medication, allopregnanolone (ALLO); or combination therapy with MDZ and ALLO (DUO). Our results show that mean T<sub>2</sub> values in DFP-exposed animals were: (1) higher than VEH in all volumes of interest (VOIs) at day 3; (2) decreased with time; and (3) decreased in the thalamus at day 28. Treatment with ALLO or DUO, but not MDZ alone, significantly decreased mean T<sub>2</sub> values relative to untreated DFP animals in the piriform cortex at day 3. On day 28, the DUO group showed the most favorable T<sub>2</sub> characteristics. This study supports the utility of T<sub>2</sub> mapping for longitudinally monitoring brain injury and highlights the therapeutic potential of ALLO as an adjunct therapy to mitigate chronic morbidity associated with acute OP intoxication.

## 1. Introduction

Organophosphate (OP) anticholinesterases are among the most commonly used pesticides worldwide (Bouchard et al., 2010; Voorhees et al., 2017; Kumar et al., 2018), and self-poisoning with OP pesticides is one of the most prevalent means of suicide in low- and middle-income countries, with an estimated hundreds of thousands of deaths globally per year between 2010 and 2014 (Gunnell et al., 2007; Mew et al., 2017). Additionally, OP nerve agents are a significant threat to public

health due to their ongoing use by militaries and terrorist groups (Tu, 2007; John et al., 2018; Nepovimova and Kuca, 2020; Young and Watson, 2020).

Acute intoxication with OPs inhibits acetylcholinesterase, which normally functions to terminate acetylcholine-mediated neurotransmission, resulting in hyperactivation of cholinergic receptors at central and peripheral synapses throughout the body (Sidell and Borak, 1992; Marrs, 1993). This manifests clinically as autonomic, somatic, and central nervous system dysfunction, including *status epilepticus* (SE), defined as a life-threatening seizure lasting more than 5 min (Todorovic

\* Corresponding author. Department of Molecular Biosciences, University of California-Davis School of Veterinary Medicine, Davis, CA, 95616, USA.

\*\* Corresponding author. Department of Radiology, UC Davis School of Medicine, 4860 Y Street, Suite 3100, Sacramento, CA, 95817, USA.

E-mail addresses: [ajdalmeida@ucdavis.edu](mailto:ajdalmeida@ucdavis.edu) (A.J.D. Almeida), [bahobson@ucdavis.edu](mailto:bahobson@ucdavis.edu) (B.A. Hobson), [dabruun@ucdavis.edu](mailto:dabruun@ucdavis.edu) (D.A. Bruun), [vporter@ucdavis.edu](mailto:vporter@ucdavis.edu) (V.A. Porter), [djharvey@ucdavis.edu](mailto:djharvey@ucdavis.edu) (D.J. Harvey), [garbow@wustl.edu](mailto:garbow@wustl.edu) (J.R. Garbow), [ajchaudhari@ucdavis.edu](mailto:ajchaudhari@ucdavis.edu) (A.J. Chaudhari), [pjlein@ucdavis.edu](mailto:pjlein@ucdavis.edu) (P.J. Lein).

<sup>1</sup> These authors contributed equally to this work.

### Abbreviations

(AIC)	Akaike information criterion
(ALLO)	allopregnanolone
(CI)	confidence interval
(DFP)	diisopropylfluorophosphate
(TE)	echo time
(FDR)	false discovery rate
(FOV)	field of view
(GMR)	geometric mean ratio
(H&E)	hematoxylin-eosin
(MRI)	magnetic resonance imaging
(MDZ)	midazolam
(MSME)	multi-slice multi-echo
(OP)	organophosphate
(SE)	status epilepticus
(VEH)	vehicle
(VOIs)	volumes of interest
(RARE)	rapid acquisition with relaxation enhancement
(SLUD)	Salivation, Lacrimation, Urination and Defecation
(TR)	repetition time

et al., 2012). The long-term neurologic consequences observed in individuals who survive acute OP intoxication include acquired epilepsy, cognitive dysfunction, and changes in brain structure that develop in weeks to months following intoxication (Yamasue et al., 2007; Chen, 2012; Tsai and Lein, 2021).

Studies of rat models of acute intoxication with the OP diisopropylfluorophosphate (DFP) have demonstrated neurotoxic effects similar to those observed in humans acutely intoxicated with OP nerve agents (Pouliot et al., 2016), including parasympathetic signs, muscle twitching, acute seizure activity that rapidly progresses to SE, and subsequent chronic morbidities. These morbidities include progressive neuropathology, spontaneous recurrent seizures, and impaired cognitive function (Sisó et al., 2017; Guignet et al., 2020). Improved control of acute seizure activity has shown promise for reducing the severity of the long-term consequences of OP intoxication in rat models (Tattersall, 2009). However, current standard of care for OP-induced SE, the benzodiazepine, midazolam (MDZ) (Lowenstein and Alldredge, 1998), shows minimal benefit in this regard even when administered within an optimal therapeutic window (e.g., within minutes of exposure) (Figueiredo et al., 2018; Wu et al., 2018; Dhir et al., 2020; Supasai et al., 2020). Adjunct therapy with the neurosteroid, allopregnanolone (ALLO), in combination with MDZ, has shown promise, demonstrating attenuation of brain injury following acute OP intoxication as assessed by histology (Lumley et al., 2019; Reddy, 2019; Dhir et al., 2020). Histological methods for evaluating brain injury, however, do not allow for repeated measurements within the same animal longitudinally, thereby precluding the study of pathologic progression. In contrast, *in vivo* imaging methods permit longitudinal studies to evaluate the evolution of neuropathology and the efficacy of therapeutic candidates.

Magnetic resonance imaging (MRI) permits non-invasive monitoring of brain pathology. The spin-spin relaxation time ( $T_2$ ), a fundamental property of hydrogen nuclei, measures the decay of net magnetization in the transverse plane. Tissue  $T_2$  values are sensitive to changes in free water and absolute quantification of  $T_2$ , represented as parametric  $T_2$  maps, has been shown to be a reliable and reproducible metric across a range of scanning environments (Deoni, 2010; Liachenko and Ramu, 2017). To date,  $T_2$  mapping has been utilized in rodent models of acute OP intoxication (Bhagat, Yusuf A.; Obenaus, André; Hamilton, Murray G.; Kendall, 2001; Bhagat et al., 2005; Gullapalli et al., 2010; Shrot et al., 2012, 2015; Lee et al., 2020); however, these studies primarily focused on identifying initial pathology. There is a paucity of data characterizing

changes in  $T_2$  beyond the first few days of OP intoxication or evaluating its efficacy for monitoring long-term therapeutic response.

In this paper we evaluated longitudinal  $T_2$  mapping as a noninvasive means to assess the impact of acute DFP intoxication over the first month in a rat model. Furthermore, we assessed the ability of  $T_2$  mapping to quantify the response to anti-seizure treatments, namely MDZ, ALLO, and the combination of MDZ and ALLO (DUO).

## 2. Materials and methods

### 2.1. Animals and acute DFP intoxication paradigm

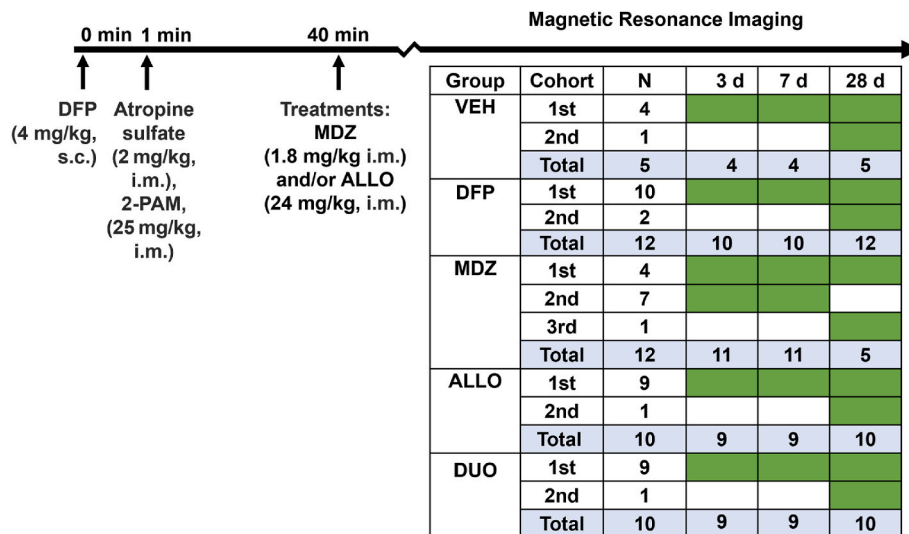
Adult male Sprague Dawley rats (250–280 g;  $n = 49$ ; Charles River Laboratories, Hollister, California) were housed individually in standard plastic cages under controlled environmental conditions ( $22 \pm 2^\circ\text{C}$ , 40%–50% humidity) with a normal 12 h light/dark cycle and *ad libitum* access to food and water. Since these studies were based on previous observations (Dhir et al., 2020) that examined the effects of MDZ and ALLO alone or in combination on acute seizure activity in male rats following acute intoxication with DFP, only male rats were used in the current study. All animal experiments were approved by the University of California at Davis Institutional Animal Care and Use Committee with attention to minimizing distress and discomfort.

Prior to any exposures, a random number generator was used to randomize rats into vehicle control (VEH) vs. DFP-intoxicated groups with the software designed to yield a greater number of DFP animals than VEH animals, in part to account for historical observations that 10–20% of DFP rats in any given cohort will either not develop SE or will die during the 24 h post-exposure. A smaller number of VEH relative to DFP animals were also selected based on the lower intra- and inter-animal variation in imaging metrics observed within VEH controls (Supplemental Fig. S9) as previously demonstrated in this animal model (Hobson et al., 2017). Rats were injected with DFP (4 mg/kg, s. c., purity  $90\% \pm 7\%$  as determined by [ $^1\text{H}$ ]-NMR (Sigma Chemical Company, St Louis, MO)), diluted in 300  $\mu\text{l}$  of sterile, ice-cold, phosphate-buffered saline (PBS, 150 mM NaCl, pH 7.2); VEH rats were injected with an equivalent volume of PBS alone. Approximately 1 min following the first injection, all animals received atropine sulfate (2 mg/kg, i. m., purity  $\geq 97\%$ , Sigma Chemical Company) and pralidoxime (2-PAM, 25 mg/kg, i. m., purity  $\geq 97\%$ , Sigma Chemical Company) (Sisó et al., 2017; Guignet et al., 2020). This protocol was shown to increase survival rate in rats acutely intoxicated with DFP (Li et al., 2011; Bruun et al., 2019). Following DFP or vehicle administration, animals were scored for seizure behavior at 5 min intervals for the first 2 h and 20 min intervals between 2 and 4 h post-DFP intoxication using a previously described modified Racine scale (Deshpande et al., 2010). This scale ranges from 1 to 5, with 5 indicating severe seizure behavior. Criterion for each score were: 1- Salivation, Lacrimation, Urination and Defecation (SLUD); 2- Tremors/Wet dog shakes; 3- Forelimb Clonus; 4- Rearing of torso; and 5- Forelimb and Hindlimb Clonus.

At 40 min post-DFP intoxication, DFP rats that achieved consecutive seizure scores  $\geq 3$  during the first 40 min post-exposure were randomized using a random number generator into one of three treatment groups: 1) DFP rats that did not receive any anti-seizure medication, 2) DFP rats treated with MDZ (1.8 mg/kg, i. m.), 3) DFP rats treated with ALLO (24 mg/kg, i. m.), and 4) DFP rats treated with combined MDZ (1.8 mg/kg, i. m.) + ALLO (24 mg/kg, i. m.) hereafter referred to as the DUO group. A schematic of the injection protocol and treatment groups is shown in Fig. 1.

### 2.2. Magnetic resonance imaging

Magnetic resonance imaging (MRI) scans were performed 3, 7 and 28 days after DFP intoxication or vehicle injection. Scans were acquired on a 7T preclinical MR scanner (Biospec 70/30, Bruker Biospin MRI, Ettlingen, Germany) equipped with a 116-mm internal diameter B-



**Fig. 1.** Schematic showing the diisopropylfluorophosphate (DFP) intoxication paradigm design and study timeline, with the number of animals imaged from each group: vehicle (VEH), DFP with no anti-seizure treatment, midazolam (MDZ), allopregnanolone (ALLO) and the MDZ + ALLO combination (DUO) at each timepoint (days 3, 7 and 28) (bottom right). The table indicates the cohorts of imaged animals per treatment group (rows), and the timepoints at which data was acquired (green shaded boxes). The final row for each treatment (light blue shading) indicates the total number of animals imaged for each treatment group at each timepoint. Note that not all animals were imaged at every timepoint.

GA12S gradient (450 mT/m, 4500 T/m/s), a 72-mm internal diameter linear transmit coil, and a four-channel, rat-brain phased array in cross coil configuration for signal reception. Animals were anesthetized with isoflurane/O<sub>2</sub> (Piramal Healthcare, Bethlehem PA), 2.0–3.0% vol/vol for 3–5 min, and then placed head-first prone in the MR scanner where they received 1.0–2.0%vol/vol of anesthesia. A stereotactic head holder consisting of a bite bar and ear canal bars was used to minimize motion during scanning. For the duration of the scanning, temperature and respiration rate were monitored (Small Animal Instruments, Inc., Stony Brook, NY), and anesthesia and delivery of warm air to the scanner bore were modulated to ensure that a body temperature of 37 °C and respiration rate of 50–70 breaths/min were maintained. Images were acquired using Paravision 5.1 (Bruker BioSpin MRI, Ettlingen, Germany).

The scanning paradigm included a localizer, followed by a multi-slice T<sub>2</sub>-weighted, rapid acquisition with relaxation enhancement (RARE) sequence for anatomic reference. Axial images were collected using the following parameters: repetition time (TR) = 6250 ms, echo time (TE) = 33 ms; RARE factor = 8; averages = 4; field-of-view (FOV) = 35.0 × 25.0 × 29.5 mm<sup>3</sup>, with a matrix size of 280 × 200 in plane and 59 slices in total, resulting in a resolution of 0.125 × 0.125 mm<sup>2</sup> in-plane and a slice thickness of 0.500 mm, spanning approximately –11.25 mm–18.25 mm Bregma. Lastly, a multi-slice multi-echo (MSME) pulse sequence was used to generate T<sub>2</sub> maps with 15 echoes equally spaced by 10 ms from 10 to 150 ms. The MSME sequence had a TR of 1500 ms; FOV of 35.0 × 25.0 × 13.0 mm<sup>3</sup>; in-plane matrix size of 140 × 100; and 9 slices along the z direction; in-plane resolution of 0.25 × 0.25 mm<sup>2</sup>, slice thickness of 1.00 mm and slice gap of 0.50 mm, spanning approximately –10 mm–3 mm Bregma. At the time of data acquisition, scan geometry was set to align the slice coordinates of the T<sub>2</sub>-weighted data with the slice coordinates of the T<sub>2</sub> maps. As a result, the T<sub>2</sub>-weighted and MSME data were inherently co-registered. Total imaging time for each rat was approximately 40 min when also accounting for animal setup and data acquisition. Positron emission tomography (PET) data were also acquired on the same days as MRI. PET scans were 30 min in duration and those data are the subject of a separate manuscript and, therefore, will not be discussed further in this article.

### 2.3. Regional delineation

Five volumes of interest (VOIs) that encompassed the piriform

cortex, amygdala, hippocampus, dorsolateral thalamus and medial thalamus were delineated from the brain scans, as these brain regions have previously been demonstrated to exhibit pathological changes following acute DFP intoxication (Hobson et al., 2017; Guignet et al., 2020). The VOIs were manually segmented on T<sub>2</sub>-weighted, axial images by individuals with experience in analyzing rat brain images and then reviewed and confirmed by an expert in rat brain anatomy using AMIRA 6.0 software (Thermo Fisher Scientific, Waltham, MA). The VOIs are illustrated in Fig. 2. Atlas-based segmentation, including non-linear warping, was not performed, as the marked brain atrophy observed in this model (Hobson et al., 2017) produces sub-optimal results. Segmentation was guided by the Paxinos and Watson's atlas of the rat brain (Paxinos, George; Watson, 2007). Where possible, image derived anatomic boundaries such as the skull, white matter tracts, and vasculature were utilized as anatomic landmarks. When unavailable, such as between portions of the piriform cortex and amygdala, region boundaries were approximated using the atlas. The axial extents of VOIs, listed as "start/end distance from Bregma" in mm are as follows: amygdala, –2.0/–4.0; hippocampus, –2.0/6.0; piriform cortex, 0.0/–3.5; thalamus (medial and dorsolateral) –1.5/5.0, and were selected to overlap known areas of OP-induced neuropathology (Sisó et al., 2017), and avoid any susceptibility or motion artifacts.

### 2.4. Creation of T<sub>2</sub> maps

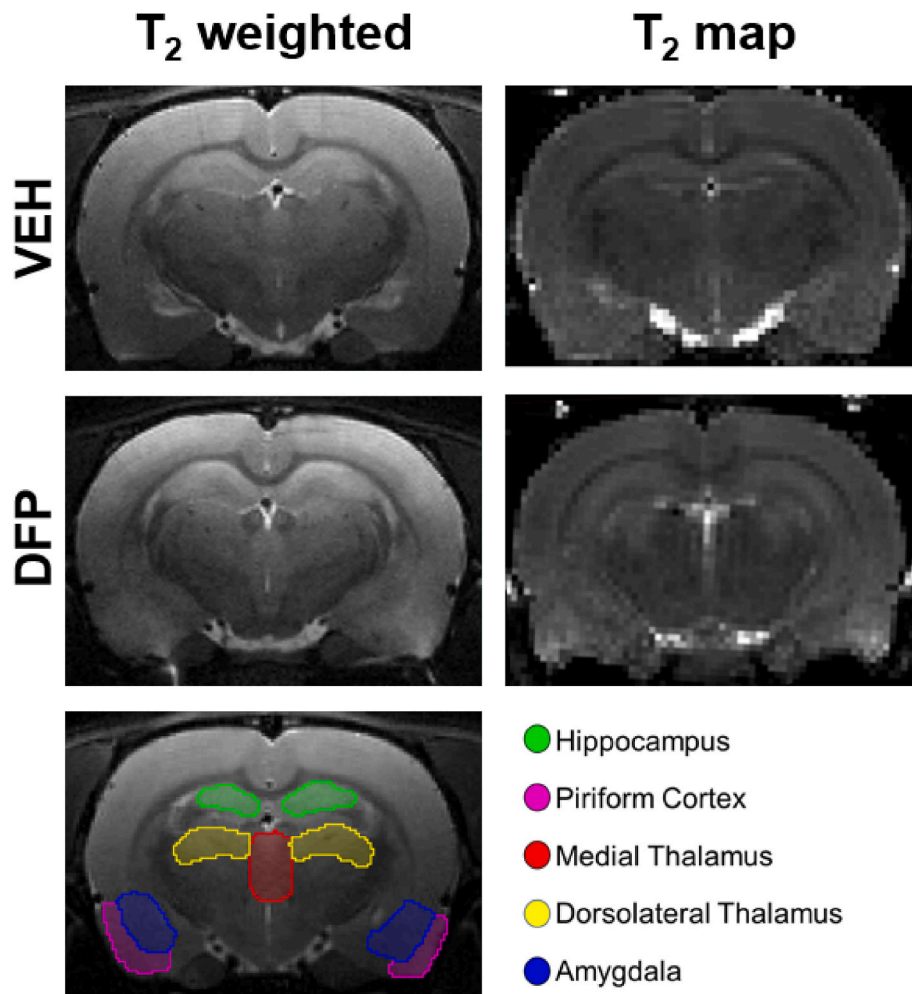
The MRI scans were cropped using a center-of-mass algorithm to remove extracerebral anatomy and reduce computation time (MATLAB and Statistics Toolbox Release, 2012b; The MathWorks Inc., Natick, MA). The software code for image cropping and T<sub>2</sub> map calculation is available on GitHub (<https://github.com/ajchaudhari/T2-mapping>).

To generate the T<sub>2</sub> maps, a mono-exponential curve fit was performed voxel-wise on the intensity vs. TE curves over the 15 TEs according to the equation:

$$M = M_0 e^{-\frac{TE}{T_2}} + f$$

where M is the signal intensity, M<sub>0</sub> is the equilibrium magnetization, and f is a constant offset/y-intercept. From the T<sub>2</sub> map, the mean, standard deviation, skewness, and kurtosis of T<sub>2</sub> values in each VOI were calculated. This voxel-wise method is used as the primary method of analysis.





**Fig. 2.** Exemplary  $T_2$ -weighted anatomical scans (left) and corresponding  $T_2$  maps (right) of the vehicle control (VEH) (row 1) and diisopropylfluorophosphate (DFP) groups (row 2) imaged at day 3 post-DFP intoxication. The volumes of interest (VOIs) are overlaid on a representative  $T_2$ -weighted MRI axial section (row 3). The bilateral VOIs evaluated from the magnetic resonance images (MRI) were the hippocampus (green), dorsolateral thalamus (yellow), medial thalamus (red), amygdala (blue) and piriform cortex (pink).

Another method of regional quantification, whereby voxel intensity values in each VOI were averaged at each TE, was also explored. A single curve fit was then performed on these averaged  $T_2$  values over 15 TEs for each VOI. This method, referred to as *Simplified Regional Quantification*, was investigated due to its reduced computation time and the results of this analysis method are presented in the supplementary material (Supplemental Figs. S1 and 2).

## 2.5. Statistical analysis

For all outcomes, which were measured in multiple brain regions and across time for the same animals, mixed effects models with robust standard errors, including animal-specific random effects, were fit to assess differences between groups. These models allow for the inclusion of between- (group) and within-animal (timepoint and brain VOI) factors and can accommodate missing data so that not all animals need to have outcomes at all timepoints. Primary factors of interest included exposure (VEH, DFP, DFP + MDZ, DFP + ALLO, DFP + DUO), brain VOI (piriform cortex, amygdala, hippocampus, dorsolateral thalamus, and medial thalamus) and timepoint (3, 7 and 28 days post-exposure). For all VOIs, except the medial thalamus, outcomes for the left and right hemisphere were averaged. For the mean metric, a log transformation after a shift in all values (subtracting 40 from the  $T_2$  mean to better meet the assumptions of the model without changing the conclusions)

resulted in better meeting the assumptions of the model. Kurtosis (a metric that encompasses the extent of outliers in the  $T_2$  value distribution) and standard deviation of  $T_2$  values were transformed using the natural logarithm to better meet the assumptions of the model (Groeneweld and Meeden, 1984). Interactions between the factors (acute-intoxication, VOI, timepoint) were considered and the best model was chosen using Akaike Information Criterion (AIC) (Akaike, 1974). Contrasts for differences between groups, either overall or by VOI or timepoint, were constructed and tested using a Wald test. Comparison of the classical likelihood-based standard errors to sandwich standard errors that are robust to heteroskedasticity revealed substantial differences for one of our outcomes (the transformed  $T_2$  mean), indicating the likelihood-based AIC method was not trustworthy unless adjustments are made for the heteroskedasticity (White, 1980; Ronchetti, 1985). Hence, for that model, we relied instead on statistical significance testing based on the robust standard errors to decide whether or not to include the interaction term. The Benjamini-Hochberg false discovery rate (FDR) was used within an outcome measure to account for multiple comparisons. Results for mean, standard deviation and kurtosis (log-transformed outcomes) across VOIs or timepoints are presented as geometric mean ratios (GMRs) between groups, while those for skew (a measure of asymmetry given by the amount of deviation from a Gaussian curve) and Simplified Regional Quantification mean (non-transformed outcomes) are presented as average differences between

groups. All analyses were performed using SAS software, version 9.4 and alpha was set at 0.05. A nonparametric one-way ANOVA was run to compare averaged seizure scores across groups at 40 min and 4 h. The  $T_2$  values were correlated with seizure severity using the nonparametric Spearman correlation test with a two-tailed p-value; multiple tests were accounted for using FDR. Comparisons remained statistically significant after the FDR correction, unless otherwise stated.

### 3. Results

#### 3.1. Acute DFP intoxication elicited robust seizure behavior and seizure score was reduced by antiseizure therapy

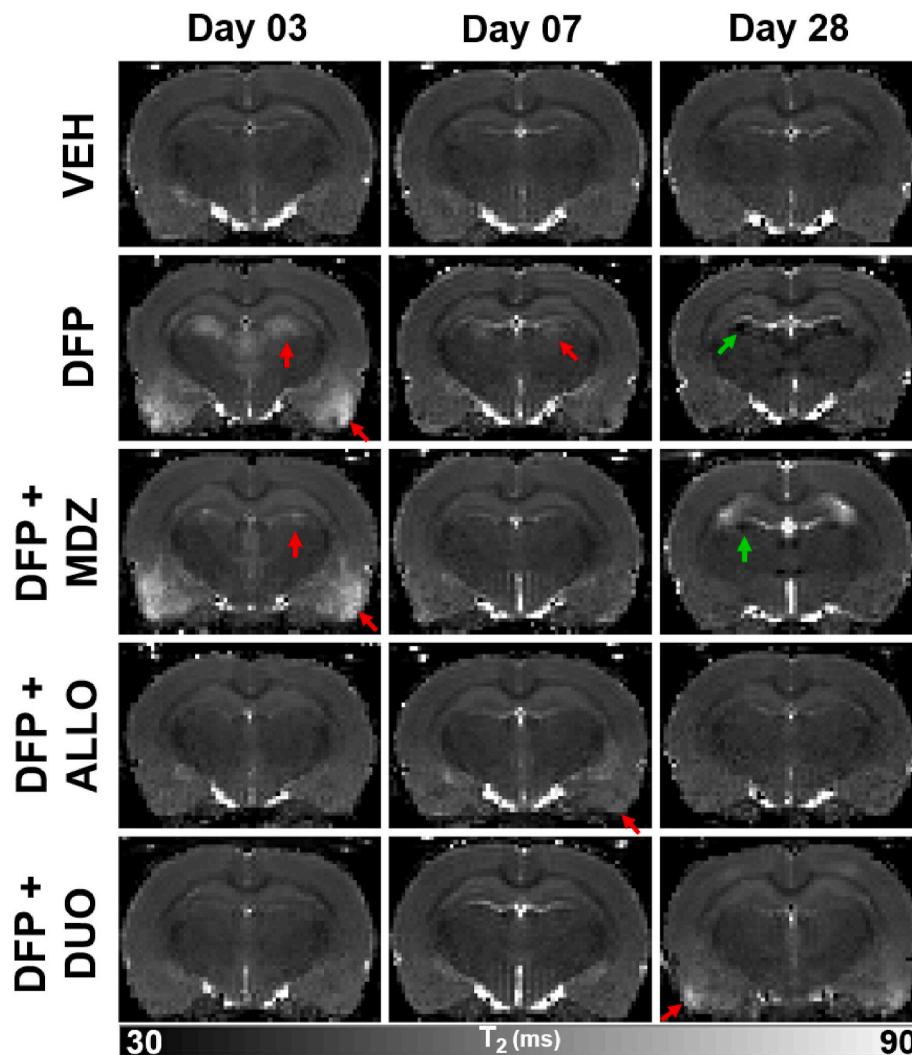
The behavioral seizure score of all DFP-intoxicated animals included in the study averaged 2.83 over the first 40 min ( $n = 44$ , although note that not all animals were imaged at all timepoints). All animals achieved consecutive seizure severity scores of 3 or above during the first 40 min, indicative of SE (Pouliot et al., 2016; Guignet et al., 2020). At 40 min post-DFP, DFP animals were randomized into treatment groups, and immediately prior to administration of treatments, no significant differences were detected in average seizure scores across treatment groups. The mean and 95% confidence intervals (CI) of behavioral seizure scores at 40 min for the groups were: DFP: 2.80 [2.63, 2.96],

MDZ: 2.86 [2.76, 2.97], ALLO: 2.81 [2.62, 3.00], DUO 2.87 [2.72, 3.02].

Rats in the ALLO and DUO anti-seizure treatment group exhibited mild-to-moderate sedation, including brief loss of righting reflex, but remained responsive to tail or toe pinch as reported in previous studies (Dhir et al., 2020). Relative to the DFP group that did not receive anti-seizure medication, administration of anti-seizure treatments significantly reduced average seizure scores from 45 min until 4 h post-exposure (MDZ,  $p < 0.001$ ; ALLO,  $p = 0.001$ ; DUO,  $p < 0.001$ ). The MDZ group had significantly higher averaged seizure score at 4 h post-DFP compared to the DUO ( $p < 0.001$ ) and ALLO ( $p = 0.01$ ) groups. No significant differences were detected between ALLO and DUO groups. The mean and 95% CI for seizure scores measured from 45 min to 4 h were: DFP: 2.51 [2.19, 2.82], MDZ: 1.78 [1.43, 2.14], ALLO: 0.76 [-0.003, 1.52], DUO: 0.13 [0.07, 0.19].

#### 3.2. $T_2$ maps revealed lesions following acute DFP intoxication

Acute DFP intoxication produced hyperintense lesions on  $T_2$  maps in locations that included, but were not limited to, the piriform cortex, thalamus, amygdala, and hippocampus (Fig. 3). Qualitatively, these lesions appeared to be more distinguishable (Fig. 2) on  $T_2$  maps compared to  $T_2$ -weighted images, and were most prominent on day 3 post-intoxication. Relative to day 3, lesions detected on days 7 and 28



**Fig. 3.** Reconstructed voxel-wise  $T_2$  parametric maps depicted brain injury following acute diisopropylfluorophosphate (DFP) intoxication that is modified by intervention with anti-seizure medication. Red arrows indicate hyperintensities consistent with edema, while green arrows indicate hypointensities suggestive of magnetic susceptibility consistent with previously reported mineralization at later timepoints.

generally demonstrated reduced intensity, extent, and prevalence. At the 28-day timepoint, hypointensities were noted in the thalami and hippocampi of animals in the DFP group, consistent with signal loss due to magnetic susceptibility (Fig. 3).

### 3.3. Region- and time-dependent changes were observed in mean $T_2$ values following acute DFP intoxication

Acute DFP intoxication resulted in a significant increase in mean  $T_2$  values (Table 1) when compared to the VEH group across all VOIs on day 3 ( $p < 0.001$ ) (Fig. 4, Table 2). This increase diminished with time, consistent with the visual appearance of lesions. The increased mean  $T_2$  values observed in the DFP group on day 3 persisted in the piriform cortex ( $p = 0.003$ ) until day 7 (Fig. 4). However, mean  $T_2$  values in the medial and dorsolateral thalamus ( $p < 0.001$ ) were significantly reduced when compared to the VEH group on day 28 (Fig. 4).

### 3.4. MDZ significantly altered DFP-induced changes in mean $T_2$ values on day 3 but not on days 7 and 28

The MDZ group displayed significantly lower mean  $T_2$  values compared to the DFP group in the hippocampus ( $p = 0.003$ ), medial thalamus ( $p = 0.004$ ) and dorsolateral thalamus ( $p = 0.001$ ) on day 3. There were no significant differences compared to the DFP group on day 7. At day 28, the MDZ group had lower mean  $T_2$  values compared to the DFP group in the amygdala ( $p = 0.002$ ) (Fig. 5, Table 2).

### 3.5. $T_2$ mapping demonstrated therapeutic rescue by DUO

On day 3, mean  $T_2$  values were significantly lower in the DUO ( $p < 0.002$ ) group across all VOIs compared to the DFP group (Table 1). The ALLO group displayed a similar trend of significantly lower  $T_2$  values on day 3 ( $p < 0.016$ ), although the difference in the amygdala was no longer significant after FDR correction (Fig. 5, Table 2). Mean  $T_2$  values remained significantly lower in the piriform cortex of the ALLO group on day 7 ( $p = 0.023$ ), but this did not remain significant after FDR correction. By contrast, mean  $T_2$  values were significantly higher in the medial thalamus of the DUO group on days 7 ( $p = 0.005$ ) and 28 ( $p = 0.001$ ) and the dorsolateral thalamus ( $p = 0.001$ ) on day 28 (Fig. 5).

### 3.6. The skewness, standard deviation, and kurtosis of $T_2$ values varied spatiotemporally

When comparing the DFP and VEH group, the standard deviation of  $T_2$  values was significantly higher in the former timepoint on days 3 and 28 (day 3;  $p < 0.001$  and day 28;  $p = 0.005$ ); this difference did not significantly vary by VOI. Standard deviation of  $T_2$  values was lower in ALLO and DUO groups in comparison to the DFP group on day 3 ( $p <$

0.001) and day 7 for the DUO group ( $p = 0.008$ ) (Supplementary Figs. S3 and 4).

Skewness of the fitted  $T_2$  values varied by VOI and was higher in the DFP group compared to the VEH group in the amygdala ( $p = 0.005$ ) and hippocampus ( $p = 0.007$ ), but lower in the piriform cortex ( $p = 0.002$ ). Skewness of  $T_2$  was significantly lower for the MDZ group in the hippocampus ( $p = 0.003$ ), for the ALLO group in the amygdala ( $p = 0.004$ ) and medial thalamus ( $p = 0.007$ ) and for the DUO group in the hippocampus ( $p < 0.001$ ) when compared to the DFP group (Supplemental Figs. S5 and 6).

Kurtosis of the  $T_2$  values was significantly lower in the piriform cortex ( $p < 0.001$ ) of the DFP group compared to the VEH group. Kurtosis of  $T_2$  values was also significantly lower in the hippocampus of the DUO group ( $p < 0.001$ ), but significantly higher in the piriform cortex ( $p = 0.002$ ) when compared to the DFP group (Supplemental Figs. S7 and 8).

### 3.7. Seizure scores and mean $T_2$ values

Seizure scores averaged at 4 h post-DFP intoxication were significantly and positively correlated with mean  $T_2$  values at day 3 across all analyzed brain VOIs ( $p < 0.02$ ) (Table 3). There was a significant negative correlation between this averaged seizure score and mean  $T_2$  values at day 7 in the medial thalamus ( $r_s = -0.41$ ,  $p = 0.008$ ) (Supplementary Table S1). A similar significantly negative correlation in  $T_2$  was observed at day 28 in the medial thalamus ( $r_s = -0.62$ ,  $p < 0.0001$ ) and the dorsolateral thalamus ( $r_s = -0.62$ ,  $p < 0.0001$ ). However, a positive correlation was observed between the average seizure score and the mean  $T_2$  value in the hippocampus at day 28 ( $r_s = 0.39$ ,  $p = 0.02$ ) (Table 4) (Supplemental Figs. S9 and 10).

## 4. Discussion

Our study utilized  $T_2$  mapping to quantify MR scans collected longitudinally in a rat model of acute DFP intoxication. Our results revealed patterns of region-specific and time-dependent changes in  $T_2$  values in brain regions known to be impacted by acute OP intoxication (Lemerrier et al., 1983; Chen, 2012). Treatment with ALLO, in combination with MDZ (DUO), significantly attenuated DFP-induced changes in  $T_2$ , highlighting its potential in mitigating brain injury following acute DFP intoxication. Significant differences in regional mean  $T_2$  values between treatment groups evolved with time, demonstrating the advantages of  $T_2$  mapping as a non-invasive biomarker for longitudinal monitoring of dynamic brain injury from acute OP intoxication and its mitigation by therapy.

Mean  $T_2$  values were significantly elevated in the DFP group on day 3 in all assessed brain VOIs compared to VEH controls. This increase was corroborated by the appearance of hyperintense lesions in those VOIs on

**Table 1**

Mean  $T_2$  values in milliseconds (ms) along with standard deviation in vehicle controls (VEH), diisopropylfluorophosphate (DFP) animals that received no anti-seizure medication and DFP animals that received combined antiseizure medication of midazolam (MDZ) and allopregnanolone (ALLO), referred to as DUO) across the 5 volumes of interest (VOIs) at the three timepoints; days 3, 7 and 28 post-acute intoxication.

VOI	$T_2$ values (ms)								
	VEH			DFP			DUO		
	Day 03	Day 07	Day 28	Day 03	Day 07	Day 28	Day 03	Day 07	Day 28
Piriform Cortex	58.0 ±4.5	57.6 ±4.0	57.1 ±4.1	71.9 ±9.0	59.7 ±4.8	57.1 ±4.8	58.0 ±3.8	58.3 ±3.3	57.2 ±4.6
Hippocampus	54.1 ±2.8	54.0 ±3.1	53.5 ±3.0	57.3 ±4.7	54.3 ±3.9	54.1 ±4.2	53.8 ±2.9	54.1 ±3.1	53.6 ±3.1
Amygdala	55.4 ±2.9	55.4 ±2.9	54.5 ±3.2	63.6 ±6.5	55.8 ±3.3	55.3 ±3.8	56.5 ±4.1	56.2 ±3.2	55.3 ±3.4
Medial Thalamus	51.6 ±3.5	51.6 ±3.5	50.9 ±3.8	56.8 ±4.9	51.0 ±3.8	47.8 ±4.8	51.7 ±3.7	52.3 ±3.6	50.5 ±3.8
Dorsolateral Thalamus	48.2 ±2.3	48.0 ±2.5	47.4 ±2.3	50.8 ±4.2	47.8 ±3.0	46.0 ±3.6	47.9 ±2.3	48.0 ±2.2	47.4 ±2.4

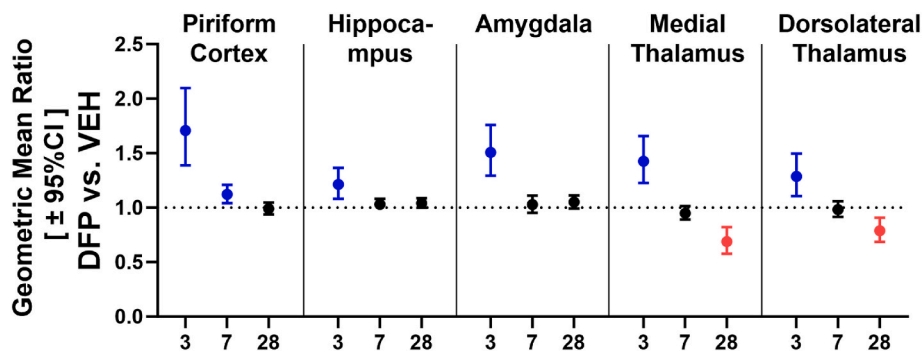


Fig. 4. Geometric mean ratios (GMRs) of mean  $T_2$  values in diisopropylfluorophosphate (DFP) vs. vehicle (VEH) groups across all timepoints (in days) and in all 5 volumes of interest (VOIs). Each dot represents the GMR; the bars indicate 95% confidence intervals (VEH:  $n = 5$  and DFP:  $n = 12$ ). Blue indicates significantly higher  $T_2$  values in DFP vs. VEH; red, significantly lower  $T_2$  values in DFP vs. VEH; and black, no significant difference in  $T_2$  values between the DFP and VEH groups. (When the confidence interval for the GMR includes 1, there is no statistical evidence of a difference between groups).

Table 2

Summary of the differences observed in mean  $T_2$  values (ms) in vehicles (VEH) diisopropylfluorophosphate (DFP) animals that received one of the antiseizure medication (midazolam (MDZ), allopregnanolone (ALLO) or combined MDZ and ALLO, referred to as DUO) vs. DFP animals that received no anti-seizure medication across all timepoints (in days) and in all 5 volumes of interest (VOIs).

VOI	VEH vs DFP			MDZ vs DFP			ALLO vs DFP			DUO vs DFP		
	Day 03	Day 07	Day 28	Day 03	Day 07	Day 28	Day 03	Day 07	Day 28	Day 03	Day 07	Day 28
Piriform Cortex	↓*	↓*					↓*	↓		↓*		
Hippocampus	↓*			↓*			↓*			↓*		
Amygdala	↓*					↓*	↓			↓*		
Dorsolateral Thalamus	↓*		↑*	↓*			↓*			↓*		↑*
Medial Thalamus	↓*		↑*	↓*			↓*			↓*	↑*	↑*

\* indicates  $p < .05$  after FDR.

Note: Upward arrows indicate increase, while downward arrows indicate decrease. \* $p < .05$  after False Discovery Rate (FDR) correction.

the corresponding  $T_2$  parametric maps, and suggest that the observed  $T_2$  hyperintensity is the result of cerebral edema (Loubinoux et al., 1997; Testylier et al., 2007). Our results are consistent with previous studies of rodent models of acute intoxication with the nerve agent soman or the pesticide paraoxon (Gullapalli et al., 2010; Shrot et al., 2012, 2015; Lee et al., 2020), as well as SE induced by kainic acid (Liachenko et al., 2015) or lithium-pilocarpine intoxication (Roch et al., 2002; Fabene et al., 2003; Duffy et al., 2014; Suleymanova et al., 2016). Specifically, consistent with our observations, these published studies noted increased  $T_2$  values at early stages post-intoxication, which were similarly attributed to edema.

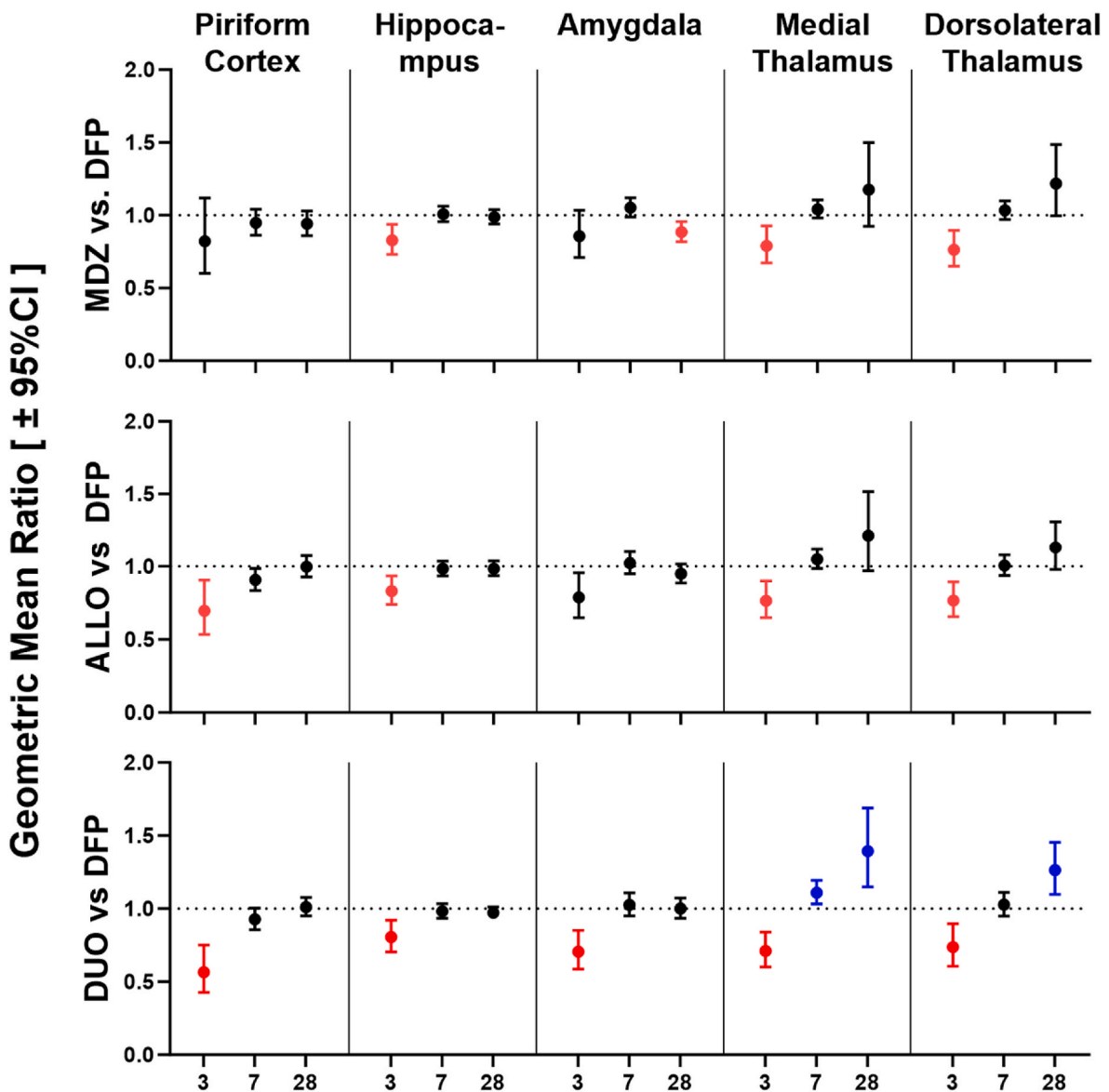
While correlative histology was not conducted in this study, and is thus a limitation, the hyperintense lesions apparent on  $T_2$ -weighted images and  $T_2$  maps were highly consistent with regards to visual appearance and spatial distribution to lesions detected in previous MRI-based assessments of acute OP intoxicated rodent models in mice (Testylier et al., 2007), rats (Bar-Klein et al., 2017; Lee et al., 2020; Reddy et al., 2020), and guinea pigs (Gullapalli et al., 2010). In these studies,  $T_2$  or  $T_2$ -weighted abnormalities on MR images were mapped onto, or show high levels of correlation with, areas of neuronal necrosis (Hobson et al., 2017; Reddy et al., 2020), neurodegeneration (Gullapalli et al., 2010), and neuroinflammation (Hobson et al., 2017) as identified histologically. This association between lesions on  $T_2$  MRI and underlying edema, cell death, and inflammation is consistent with the interpretation of  $T_2$  MRI in the broader field of preclinical seizure and epilepsy models (Gröhn and Pitkänen, 2007; Liachenko et al., 2015; Reddy et al., 2019). Additionally, histological and biochemical assessments of these same brain areas in DFP intoxicated rats (Sisó et al., 2017; Kuruba et al., 2018; Lee et al., 2020; Supasai et al., 2020) that used hematoxylin-eosin (H&E) staining or FluoroJade C staining and immunohistochemical analyses detected severe neuronal necrosis, microgliosis, reactive astrogliosis. Thus, while the present study utilized longitudinal *in vivo* MRI rather than histological assessment, the data suggest significant underlying

neuropathology resultant to DFP intoxication.

At the 7 day timepoint, we observed a reduction in mean  $T_2$  values relative to day 3 across all VOIs in the DFP group. While the most likely explanation of this effect is the attenuation of edema, we acknowledge that this normalization may be the result of concurrent mechanisms with differential effects on  $T_2$ . Specifically, increased  $T_2$  due to edematous processes may coincide with mechanisms driving the decreased  $T_2$  values seen at later timepoints, such as mineralization (discussed below). In this regard, we observed a reduction in the standard deviation of  $T_2$  values at the 7 day timepoint across regions, suggesting the attenuation of edema rather than a cancellation of  $T_2$  effects due to competing mechanisms. This pattern of increased and subsequently decreased mean  $T_2$  during the first week post-OP intoxication has been previously reported (Gullapalli et al., 2010; Shrot et al., 2015) and is consistent with the slow restoration of the blood brain barrier (Bar-Klein et al., 2017; Bernardino et al., 2023) and may be associated with subsiding of vasogenic edema (Carpentier et al., 1990). However, mean  $T_2$  values remained elevated in the piriform cortex on day 7. The persistent injury in the piriform cortex may be the result of more severe underlying neuropathology that has been previously reported in the piriform cortex in animal models of acute OP intoxication (Baille et al., 2005; de Araujo Furtado et al., 2012; Sisó et al., 2017) and has been characterized by marked reduction in tissue cellularity following pan necrosis and tissue liquefaction. In terms of regional progression of injury beyond 7 days, at the 28 day timepoint,  $T_2$  values in the amygdala further decreased to levels observed in VEH group, suggesting near complete resolution of the edematous process.

The spatiotemporal progression of edema detected by  $T_2$  imaging following OP intoxication has varied significantly across studies. In a guinea pig model of soman intoxication, mean  $T_2$  values remained elevated at day 7 in the piriform cortex, amygdala and thalami, but returned to VEH levels in the hippocampus (Gullapalli et al., 2010). This is in partial agreement with our study where mean  $T_2$  values at day 7





**Fig. 5.** Geometric mean ratios (GMRs) of mean  $T_2$  values between diisopropylfluorophosphate (DFP) animals that received no anti-seizure medication vs. DFP animals that received midazolam (MDZ), allopregnanolone (ALLO) or combined MDZ and ALLO, referred to as DUO across all timepoints (in days) and in all 5 volumes of interest (VOIs). Each dot represents the GMR between DFP and VEH; bars, the 95% confidence intervals (DFP:  $n = 12$ , MDZ:  $n = 12$ , ALLO:  $n = 10$ , DUO:  $n = 10$  animals). Blue indicates significantly higher  $T_2$  values in the treated group vs. DFP alone; red, significantly lower  $T_2$  values in the treated group vs. DFP; black, no significant difference in  $T_2$  values between anti-seizure treatment groups (MDZ, ALLO, DUO) and the DFP group. (When the confidence interval for the GMR includes 1, there is no statistical evidence of a difference between groups).

**Table 3**

Spearman correlation coefficients between averaged seizure scores and mean  $T_2$  values at day 3 across all 5 volumes of interest (VOIs). ( $r_s$ : Spearman correlation coefficient, CI: confidence interval).

VOI	Correlation between averaged seizure score at 4 h vs $T_2$ value at day 3		
	$r_s$	95% CI	p
Piriform Cortex	0.59	0.33 to 0.76	<0.0001
Hippocampus	0.38	0.07 to 0.62	0.02
Amygdala	0.65	0.41 to 0.80	<0.0001
Medial Thalamus	0.64	0.40 to 0.79	<0.0001
Dorsolateral Thalamus	0.46	0.16 to 0.67	0.003

remained elevated in the piriform cortex. However, another study of acute paraoxon intoxication reported no changes in  $T_2$  in any VOI of the rat brain at 7 days post-exposure (Shrot et al., 2012). Elevated  $T_2$  values

**Table 4**

Spearman correlation coefficients between averaged seizure scores and mean  $T_2$  values at day 28 across all 5 volumes of interest (VOIs). ( $r_s$ : Spearman correlation coefficient, CI: confidence interval).

VOI	Correlation between averaged seizure score at 4 h vs $T_2$ value at day 28		
	$r_s$	95% CI	p
Piriform Cortex	-0.07	-0.38 to 0.26	0.7
Hippocampus	0.39	0.07 to 0.63	0.02
Amygdala	0.08	-0.25 to 0.39	0.6
Medial Thalamus	-0.62	-0.78 to -0.37	<0.0001
Dorsolateral Thalamus	-0.62	-0.78 to -0.37	<0.0001

at timepoints beyond 7 days have also been reported in the lithium pilocarpine model of SE that reported elevated  $T_2$  values persisting until the 30-day timepoint (Suleymanova et al., 2016), specifically in the

parietal and prefrontal cortices. These discrepancies in patterns of  $T_2$  across studies could be due to differences in the specific mechanism by which the intoxicating agent induces SE (Banks and Lein, 2012), intra-species variability (Calsbeek et al., 2021), or the variation in MR image quality resulting from the tradeoff between acquisition time, spatial resolution, and signal to noise, as well as image analysis methods.

A unique finding of the present work is the significant reduction in  $T_2$  values in the medial and dorsolateral thalamus at 28 days post-acute OP intoxication. Visual inspection of  $T_2$  maps (Fig. 3) revealed small hypointense lesions in both these VOIs. The stark drop-out in signal and reduction of estimated  $T_2$  values is suggestive of magnetic susceptibility effects, often associated with the presence of mineral deposits (Langkammer et al., 2012), which can alter the local magnetic field, thereby reducing the  $T_2$  values (Borradaile, 1988; Schenck, 1996). Dystrophic mineralization in the thalamus is a common finding in preclinical models of acute OP intoxication (Kadar et al., 1995; Pouliot et al., 2016; Sisó et al., 2017) and preclinical models of epilepsy (Gramsbergen and van den Berg, 1994; Gayoso et al., 2003; Aggarwal et al., 2018). Previous *ex vivo* studies using CT (González et al., 2020) and histology (Kadar et al., 1995; Sisó et al., 2017) have described mineralized lesions with a similar spatial pattern within the hippocampus and thalami as a long-term consequence of DFP intoxication. While the appearance of these lesions is consistent with mineralization, their exact origin remains unclear, and may be the result of multiple pathological processes associated with OP intoxication including, but not limited, to neuroinflammation (Maheshwari et al., 2022), necrotic (Gezercan et al., 2016) or non-necrotic (Deshpande et al., 2010) calcium dysregulation (Deshpande et al., 2010), and slow accumulation of iron following the development of spontaneous recurrent seizures (Zimmer et al., 2021). Nonetheless, their development and effect on  $T_2$  following DFP intoxication appears to have been significantly attenuated following administration of MDZ and ALLO.

Administration of 1.8 mg/kg of MDZ reduced  $T_2$  values in the hippocampus and thalamic VOIs at day 3 compared to DFP alone. In addition to these VOIs, a single 24 mg/kg dose of ALLO administered 40 min post-DFP exposure consistently reduced  $T_2$  hyperintense lesions in the piriform cortex as well. Furthermore, combined treatment with ALLO and MDZ (DUO) not only diminished the early increase in mean  $T_2$  values across all VOIs, but also attenuated the  $T_2$  hypointensity observed in both the thalamic VOIs at the 28 day timepoint in the DFP group. By comparison, MDZ or ALLO alone did not significantly alter  $T_2$  values in any VOI at the 7 or 28 day timepoint when compared to the DFP group. These observations strongly suggest that ALLO, particularly administered alongside MDZ, may be a more effective strategy than MDZ alone in attenuating OP-related neuropathology.

The ability of a one-time combined dose of ALLO and MDZ (DUO therapy) to attenuate both the initial hyperintense lesions and their later hypointense counterparts is likely attributable to an improvement in anti-seizure performance compared to MDZ or ALLO alone. Administration of DUO significantly reduced seizure behavior when compared to MDZ alone. We observed significant positive correlations between the average seizure score over the 4-h monitoring window after DFP intoxication and mean  $T_2$  values in all VOIs assessed at the 3-day timepoint. Additionally, the significant negative correlation between averaged seizure score at 4 h and  $T_2$  measured at day 28 in the medial and dorsolateral thalamus likely reflect an association between seizure severity and mineralization. The DUO group had the lowest seizure score at 4 h, followed by the ALLO and then the MDZ groups, which was in agreement with the spatiotemporal patterns of  $T_2$ , where administration of the DUO treatment appeared to be most effective in attenuating both hyper- and hypointense lesions, followed by ALLO. While both ALLO and MDZ are positive allosteric modifiers of the GABA<sub>A</sub> receptor, each binds to a distinct subset of GABA<sub>A</sub> receptor subtypes (Rogawski et al., 2013). Benzodiazepine-sensitive GABA<sub>A</sub> receptors are synaptic receptors that function to decrease neuronal excitability (Masiulis et al., 2019). These synaptic receptors are rapidly internalized

during SE (Naylor et al., 2005), which diminishes the therapeutic effectiveness of benzodiazepines, such as MDZ. By contrast, ALLO binds not only synaptic but also extrasynaptic  $\alpha_4\delta$  GABA<sub>A</sub> receptors (Mottalino et al., 2006) that are not internalized during seizure, and may mediate tonic inhibition. The broader target specificity of ALLO likely explains the improved performance of the DUO therapy. While it is difficult to speculate on the impact of the DUO therapy with regards to long-term outcomes of OP intoxication such as SRS and cognition, there is some evidence that attenuation of  $T_2$ -related lesions on MRI is associated with a reduction in long-term brain atrophy. Guinea pigs intoxicated with the OP nerve agent soman exhibited a similar pattern of  $T_2$  elevation in the hippocampus followed by an expansion of ventricular CSF highly suggestive of brain atrophy (Gullapalli et al., 2010). Similarly, in a rat model of lithium-pilocarpine-induced SE, initial increases in  $T_2$  in the hippocampus were replaced by expansion of the lateral ventricles and a marked reduction in hippocampal volume (Suleymanova et al., 2014). In both models, intervention with candidate therapies attenuated both the increased  $T_2$  and the subsequent atrophy. In the guinea pig model this attenuation associated in reduced histopathological evidence of neuronal cell death in the same spatial location as the  $T_2$ -weighted lesions (Gullapalli et al., 2010). Notably, in a mouse model of SE using pilocarpine, animals that experienced SE developed  $T_2$  increases in the hippocampus, piriform cortex, and amygdala similar to the present study, and subsequent hippocampal atrophy. When a subset of animals was further instrumented with EEG, no association was observed between these changes and spike frequency when assessed during a randomly selected 96 h window (Kharatishvili et al., 2014). Such data indicate  $T_2$  may not be predictive of long-term electrographic abnormalities; however, a more comprehensive assessment by EEG following  $T_2$  MRI is warranted.

In addition to mean  $T_2$  values, our study analyzed additional  $T_2$  metrics, including standard deviation, skewness, and kurtosis. Across regions, the standard deviation of  $T_2$  values was increased in the DFP group relative to controls. This higher level of intra-tissue variance is in line with our previous work demonstrating a similar increase in intra-regional variance in tissue diffusion following DFP intoxication (Hobson et al., 2017). These results suggest subregional structures, such as neuronal cell layers, are experiencing varying levels of underlying neuropathological processes such as edema, necrosis, or inflammation (Hobson et al., 2017, 2018; Sisó et al., 2017). The increased skewness, in the hippocampus, would further suggest that the observed changes in mean  $T_2$  may be similarly driven by localized regions of severe neuropathology. Kurtosis assesses the degree by which the mean is influenced by extreme values in the tails of the distribution. Thus, increases in kurtosis may indicate spatially heterogeneous effects of DFP intoxication on the underlying tissue while reduced kurtosis may indicate that a given brain area experienced a more spatially homogenous effect. While some patterns exhibited by these novel metrics were unique, differences were not as striking as those detected by the mean metric. Further, it is important to note that while the subject numbers in the present study demonstrated robust effects of DFP intoxication and therapeutic rescue as determined by analysis of mean values, assessments of skewness and kurtosis values are typically much more meaningful when significantly larger datasets are evaluated (Cain et al., 2017). Finally, methods for texture analysis of images have recently been developed that capture even higher order information from image intensity features (Gillies et al., 2016). It is possible that such analyses may aid in the detection of subtle features not detected by current metrics (Crombé et al., 2021).

Traditional  $T_2$ -weighted images were limited in their sensitivity to detect lesions one-week post-acute DFP intoxication (Hobson et al., 2017). In contrast,  $T_2$  mapping enabled absolute quantification and further eliminated dependencies on extraneous factors such as acquisition parameters, signal amplifier gains and geometry of receiver coil (Deoni, 2010), thereby resulting in better reproducibility of intensity values across scans and timepoints. Previous methods have evaluated diffusion weighted imaging (DWI) for tracking pathology in the brain of

rats acutely intoxicated with DFP (Testylier et al., 2007; Rosman et al., 2012; Shrot et al., 2015; Hobson et al., 2017). These studies showed heterogenous diffusion in regions that align with those exhibiting T<sub>2</sub> alterations in our study. It is currently unclear if DWI metrics could provide complementary information compared to that from T<sub>2</sub> maps. Additionally, there could be value in comparing assessments of neuroinflammation through positron emission tomography (PET) imaging (Sisó et al., 2017; Hobson et al., 2019) to evaluate spatiotemporal changes in neuroinflammation in the brain relative to lesions identified on the T<sub>2</sub> maps (Shrot et al., 2015; Hobson et al., 2017, 2018). While beyond the scope of this study, such a head-to-head comparison of other *in vivo* imaging metrics and those derived from T<sub>2</sub> maps is warranted in future studies.

Our study focused on the effects of ALLO and MDZ on animals undergoing acute OP intoxication and did not include separate assessments of ALLO and MDZ on animals not intoxicated with DFP. Similar dosing of ALLO was well tolerated in rats (Irwin et al., 2015; Dhir et al., 2020) and MDZ is canonically used at similar doses in preclinical models of OP intoxication and epilepsy (Schläppi, 1983; Reddy and Reddy, 2015) without adverse outcomes. Thus, while we do not anticipate any meaningful effect of these agents on T<sub>2</sub> in healthy animals, we are unable to decisively conclude this based on the data from the present study. We analyzed five brain VOIs chosen because of their known significance as targets of acute OP intoxication. A comparative analysis of T<sub>2</sub> signals in these VOIs with other brain regions could provide comprehensive insight into the whole-brain impact of acute OP intoxication. While manual delineation of MRI scans is tedious, these analyses could be facilitated by recent advancements in machine learning techniques enabling more efficient and accurate brain regional analysis (Feo and Giove, 2019). While all the animals included in the present study were males, we are currently performing experiments with male and female rats to account for gender differences. Lastly, a pragmatic choice was made at the beginning of the study to balance factors such as spatial resolution, signal-to-noise ratio, number of echoes and acquisition time to tailor the imaging protocol to the specific goals of this project. A different trade-off between imaging parameters may be needed when implementing our method for other strains or species or when using an MRI scanner operating at a different field strength or having a different hardware configuration.

In conclusion our study demonstrates the potential for T<sub>2</sub> mapping as a valuable, quantitative tool to track brain injury resulting from OP-intoxication and to assess the efficacy of neuroprotective therapies. It also highlights the potential of the neurosteroid ALLO as an adjunct therapy for mitigating OP-induced damage. The availability of analogous methods for T<sub>2</sub> mapping for human brain imaging (Carr and Purcell, 1954; Meiboom and Gill, 1958; Peters et al., 2007) suggests that the findings from our study could be translated to clinical settings (Peters et al., 2007). With the increasing availability of MRI, T<sub>2</sub> mapping could serve as a useful tool for developing broad-spectrum medical countermeasures against various seizure-inducing chemicals, where longitudinal monitoring of brain injury and therapeutic response is crucial.

## Funding

This work was supported by the National Institutes of Health [grant numbers U54 NS079202, U54 NS127758 and R21 AG064599] and the National Center for Advancing Translational Sciences [grant number UL1 TR001860 and linked award TL1 TR001861].

## CRediT authorship contribution statement

**Alita Jesal D. Almeida:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Writing – original draft, Writing – review & editing. **Brad A. Hobson:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Supervision, Validation, Writing – original draft, Writing – review &

editing. **Naomi Saito:** Data curation, Formal analysis, Investigation, Software, Writing – review & editing. **Donald A. Bruun:** Data curation, Investigation, Methodology, Validation, Writing – review & editing. **Valerie A. Porter:** Data curation, Investigation, Writing – review & editing. **Danielle J. Harvey:** Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Supervision, Validation, Writing – original draft, Writing – review & editing. **Joel R. Garbow:** Conceptualization, Investigation, Methodology, Validation, Writing – review & editing. **Abhijit J. Chaudhari:** Conceptualization, Investigation, Methodology, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing. **Pamela J. Lein:** Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – review & editing.

## Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Pamela J. Lein has patent Mitigation of epileptic seizures by combination therapy using benzodiazepines and neurosteroids issued to The Regents of the University of California.

## Data availability

Data will be made available on request.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.neuropharm.2024.109895>.

## References

- Aggarwal, M., et al., 2018. Nuclei-specific deposits of iron and calcium in the rat thalamus after status epilepticus revealed with quantitative susceptibility mapping (QSM). *J. Magn. Reson. Imag. : JMIRI* 47 (2), 554–564. <https://doi.org/10.1002/jmri.25777>.
- Akaike, H., 1974. A new look at the statistical model identification. *IEEE Trans. Automat. Control* 19 (6), 716–723. <https://doi.org/10.1109/TAC.1974.1100705>.
- de Araujo Furtado, M., et al., 2012. Exposure to nerve agents: from status epilepticus to neuroinflammation, brain damage, neurogenesis and epilepsy. *Neurotoxicology* 33 (6), 1476–1490. <https://doi.org/10.1016/j.neuro.2012.09.001>.
- Baille, V., et al., 2005. Soman-induced convulsions: the neuropathology revisited. *Toxicology* 215 (1–2), 1–24. <https://doi.org/10.1016/j.tox.2005.05.028>.
- Banks, C.N., Lein, P.J., 2012. A review of experimental evidence linking neurotoxic organophosphorus compounds and inflammation. *Neurotoxicology* 33 (3), 575–584. <https://doi.org/10.1016/j.neuro.2012.02.002>.
- Bar-Klein, G., et al., 2017. Imaging blood–brain barrier dysfunction as a biomarker for epileptogenesis. *Brain* 140 (6), 1692–1705. <https://doi.org/10.1093/brain/awx073>.
- Bernardino, P.N., et al., 2023. Time- and region-dependent blood-brain barrier impairment in a rat model of organophosphate-induced status epilepticus. *Neurobiol. Dis.* 187, 106316. <https://doi.org/10.1016/j.nbd.2023.106316>.
- Bhagat, Y.A., et al., 2005. Neuroprotection from soman-induced seizures in the rodent: evaluation with diffusion- and T2-weighted magnetic resonance imaging. *Neurotoxicology* 26 (6), 1001–1013. <https://doi.org/10.1016/j.neuro.2005.04.006>.
- Bhagat, Yusuf A., Obenaus, André, Hamilton, Murray G., Kendall, E.J., 2001. Magnetic resonance imaging predicts neuropathology from soman-mediated seizures in the rodent. *Neuroreport* 12 (7), 1481–1487.
- Borradaile, G.J., 1988. Magnetic susceptibility, petrofibrils and strain. *Tectonophysics* 156 (1–2), 1–20. [https://doi.org/10.1016/0040-1951\(88\)90279-X](https://doi.org/10.1016/0040-1951(88)90279-X).
- Bouchard, M.F., et al., 2010. Attention-deficit/hyperactivity disorder and urinary metabolites of organophosphate pesticides. *Pediatrics* 125 (6), e1270–e1277. <https://doi.org/10.1542/peds.2009-3058>.
- Bruun, D.A., et al., 2019. Pretreatment with pyridostigmine bromide has no effect on seizure behavior or 24 hour survival in the rat model of acute diisopropylfluorophosphate intoxication. *Neurotoxicology* 73, 81–84. <https://doi.org/10.1016/j.neuro.2019.03.001>.
- Cain, M.K., Zhang, Z., Yuan, K.-H., 2017. Univariate and multivariate skewness and kurtosis for measuring nonnormality: prevalence, influence and estimation. *Behav. Res. Methods* 49 (5), 1716–1735. <https://doi.org/10.3758/s13428-016-0814-1>.
- Calsbeek, J.J., et al., 2021. Strain differences in the extent of brain injury in mice after tetramethylenedisulfotetramine-induced status epilepticus. *Neurotoxicology* 87, 43–50. <https://doi.org/10.1016/j.neuro.2021.08.011>.
- Carpentier, P., et al., 1990. Seizure-related opening of the blood-brain barrier induced by soman: possible correlation with the acute neuropathology observed in poisoned

- rats. *Neurotoxicology* 11 (3), 493–508. <http://www.ncbi.nlm.nih.gov/pubmed/2284054>.
- Carr, H.Y., Purcell, E.M., 1954. Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Phys. Rev.* 94 (3), 630–638. <https://doi.org/10.1103/PhysRev.94.630>.
- Chen, Y., 2012. Organophosphate-induced brain damage: mechanisms, neuropsychiatric and neurological consequences, and potential therapeutic strategies. *Neurotoxicology* 33 (3), 391–400. <https://doi.org/10.1016/j.neuro.2012.03.011>.
- Cromb , A., et al., 2021. Assessment of repeatability, reproducibility, and performances of <sc>T2</sc> mapping-based radiomics features: a comparative study. *J. Magn. Reson. Imag.* 54 (2), 537–548. <https://doi.org/10.1002/jmri.27558>.
- Deoni, S.C.L., 2010. Quantitative relaxometry of the brain. *Top. Magn. Reson. Imag.* 21 (2), 101–113. <https://doi.org/10.1097/RMR.0b013e31821e56d8>.
- Deshpande, L.S., et al., 2010. Development of a prolonged calcium plateau in hippocampal neurons in rats surviving status epilepticus induced by the organophosphate diisopropylfluorophosphate. *Toxicol. Sci. : an official journal of the Society of Toxicology* 116 (2), 623–631. <https://doi.org/10.1093/toxsci/kfq157>.
- Dhir, A., et al., 2020. Allopregnanolone and perampanel as adjuncts to midazolam for treating diisopropylfluorophosphate-induced status epilepticus in rats. *Ann. N. Y. Acad. Sci.* 1480 (1), 183–206. <https://doi.org/10.1111/nyas.14479>.
- Duffy, B.A., et al., 2014. Dexamethasone exacerbates cerebral edema and brain injury following lithium-pilocarpine induced status epilepticus. *Neurobiol. Dis.* 63, 229–236. <https://doi.org/10.1016/j.nbd.2013.12.001>.
- Fabene, P.F., et al., 2003. Magnetic resonance imaging of changes elicited by status epilepticus in the rat brain: diffusion-weighted and T2-weighted images, regional blood volume maps, and direct correlation with tissue and cell damage. *Neuroimage* 18 (2), 375–389. [https://doi.org/10.1016/S1053-8119\(02\)00025-3](https://doi.org/10.1016/S1053-8119(02)00025-3).
- Feo, R., Giove, F., 2019. Towards an efficient segmentation of small rodents brain: a short critical review. *J. Neurosci. Methods* 323, 82–89. <https://doi.org/10.1016/j.jneumeth.2019.05.003>.
- Figueiredo, T.H., et al., 2018. Acute and long-term consequences of exposure to organophosphate nerve agents in humans. *Epilepsia* 59, 92–99. <https://doi.org/10.1111/epi.14500>.
- Gayoso, M.J., et al., 2003. Selective calcification of rat brain lesions caused by systemic administration of kainic acid. *Histol. Histopathol.* 18 (3), 855–869. <https://doi.org/10.14670/HH-18.855>.
- Gezercan, Y., et al., 2016. Six different extremely calcified lesions of the brain: brain stones. *SpringerPlus* 5 (1), 1941. <https://doi.org/10.1186/s40064-016-3621-3>.
- Gillies, R.J., Kinahan, P.E., Hricak, H., 2016. Radiomics: images are more than pictures, they are data. *Radiology* 278 (2), 563–577. <https://doi.org/10.1148/radiol.2015151169>.
- Gonz lez, E.A., et al., 2020. The chemical convulsant diisopropylfluorophosphate (DFP) causes persistent neuropathology in adult male rats independent of seizure activity. *Arch. Toxicol.* 94 (6), 2149–2162. <https://doi.org/10.1007/s00204-020-02747-w>.
- Gramsbergen, J.B., van den Berg, K.J., 1994. Regional and temporal profiles of calcium accumulation and glial fibrillary acidic protein levels in rat brain after systemic injection of kainic acid. *Brain Res.* 667 (2), 216–228. [https://doi.org/10.1016/0006-8993\(94\)91499-0](https://doi.org/10.1016/0006-8993(94)91499-0).
- Groeneveld, R.A., Meeden, G., 1984. Measuring skewness and kurtosis. *The Statistician* 33 (4), 391. <https://doi.org/10.2307/2987742>.
- Gr hn, O., Pitk nen, A., 2007. Magnetic resonance imaging in animal models of epilepsy—noninvasive detection of structural alterations. *Epilepsia* 48 (s4), 3–10. <https://doi.org/10.1111/j.1528-1167.2007.01236.x>.
- Guignat, M., et al., 2020. Persistent behavior deficits, neuroinflammation, and oxidative stress in a rat model of acute organophosphate intoxication. *Neurobiol. Dis.* 133, 104431. <https://doi.org/10.1016/j.nbd.2019.03.019>.
- Gullapalli, R.P., et al., 2010. Magnetic resonance imaging reveals that galantamine prevents structural brain damage induced by an acute exposure of Guinea pigs to soman. *Neurotoxicology* 31 (1), 67–76. <https://doi.org/10.1016/j.neuro.2009.09.004>.
- Gunnell, D., et al., 2007. The global distribution of fatal pesticide self-poisoning: systematic review. *BMC Public Health.* *BioMed Central* 1–15. <https://doi.org/10.1186/1471-2458-7-357>.
- Hobson, B.A., et al., 2017. From the cover: magnetic resonance imaging reveals progressive brain injury in rats acutely intoxicated with diisopropylfluorophosphate. *Toxicol. Sci.* 157 (2), 342–353. <https://doi.org/10.1093/toxsci/kfx049>.
- Hobson, B.A., et al., 2018. A magnetic resonance imaging study of early brain injury in a rat model of acute DFP intoxication. *Neurotoxicology* 66, 170–178. <https://doi.org/10.1016/j.neuro.2017.11.009>.
- Hobson, B.A., et al., 2019. TSPO PET using [18F]PBR111 reveals persistent neuroinflammation following acute diisopropylfluorophosphate intoxication in the rat. *Toxicol. Sci.* 170 (2), 330–344. <https://doi.org/10.1093/toxsci/kfz096>.
- Irwin, R.W., et al., 2015. Allopregnanolone preclinical acute pharmacokinetic and pharmacodynamic studies to predict tolerability and efficacy for Alzheimer’s disease. *PLoS One* 10 (6), e0128313. <https://doi.org/10.1371/journal.pone.0128313>.
- John, H., et al., 2018. Fatal sarin poisoning in Syria 2013: forensic verification within an international laboratory network. *Forensic Toxicol.* 36 (1), 61–71. <https://doi.org/10.1007/s11419-017-0376-7>.
- Kadar, T., et al., 1995. Sarin-induced neuropathology in rats. *Hum. Exp. Toxicol.* 14 (3), 252–259. <https://doi.org/10.1177/096032719501400304>.
- Kharatishvili, I., et al., 2014. MRI changes and complement activation correlate with epileptogenicity in a mouse model of temporal lobe epilepsy. *Brain Struct. Funct.* 219 (2), 683–706. <https://doi.org/10.1007/s00429-013-0528-4>.
- Kumar, S., et al., 2018. Microbial degradation of organophosphate pesticides: a review. *Pedosphere* 28 (2), 190–208. [https://doi.org/10.1016/S1002-0160\(18\)60017-7](https://doi.org/10.1016/S1002-0160(18)60017-7).
- Kuruba, R., Wu, X., Reddy, D.S., 2018. Benzodiazepine-refractory status epilepticus, neuroinflammation, and interneuron neurodegeneration after acute organophosphate intoxication. *Biochim. Biophys. Acta (BBA) - Mol. Basis Dis.* 1864 (9 Pt B), 2845–2858. <https://doi.org/10.1016/j.bbadis.2018.05.016>.
- Langkammer, C., et al., 2012. Quantitative susceptibility mapping (QSM) as a means to measure brain iron? A post mortem validation study. *Neuroimage* 62 (3), 1593–1599. <https://doi.org/10.1016/j.neuroimage.2012.05.049>.
- Lee, K., et al., 2020. Quantitative T2 MRI is predictive of neurodegeneration following organophosphate exposure in a rat model. *Sci. Rep.* 10 (1) <https://doi.org/10.1038/s41598-020-69991-z>.
- Lemercier, G., et al., 1983. Histological and histochemical changes in the central nervous system of the rat poisoned by an irreversible anticholinesterase organophosphorus compound. *Acta Neuropathol.* 61 (2), 123–129. <https://doi.org/10.1007/BF00697391>.
- Li, Y., et al., 2011. Spatiotemporal pattern of neuronal injury induced by DFP in rats: a model for delayed neuronal cell death following acute OP intoxication. *Toxicol. Appl. Pharmacol.* 253 (3), 261–269. <https://doi.org/10.1016/j.taap.2011.03.026>.
- Liachenko, S., et al., 2015. Quantitative assessment of MRI T2 response to kainic acid neurotoxicity in rats in vivo. *Toxicol. Sci.* 146 (1), 183–191. <https://doi.org/10.1093/toxsci/kfv083>.
- Liachenko, S., Ramu, J., 2017. Quantification and reproducibility assessment of the regional brain T2 relaxation in naive rats at 7T. *J. Magn. Reson. Imag.* 45 (3), 700–709. <https://doi.org/10.1002/jmri.25378>.
- Loubinoux, I., et al., 1997. Spreading of vasogenic edema and cytotoxic edema assessed by quantitative diffusion and T2 magnetic resonance imaging. *Stroke* 28 (2), 419–427. <https://doi.org/10.1161/01.STR.28.2.419>.
- Lowenstein, D.H., Allredge, B.K., 1998. Status epilepticus. *N. Engl. J. Med.* 338 (14), 970–976. <https://doi.org/10.1056/NEJM199804023381407>.
- Lumley, L., et al., 2019. Neurosteroid and benzodiazepine combination therapy reduces status epilepticus and long-term effects of whole-body sarin exposure in rats. *Epilepsia* open 4 (3), 382–396. <https://doi.org/10.1002/epi4.12344>.
- Maheshwari, U., et al., 2022. The interplay between brain vascular calcification and microglia. *Front. Aging Neurosci.* 14 <https://doi.org/10.3389/fnagi.2022.848495>.
- Marrs, T.C., 1993. Organophosphate poisoning. *Pharmacol. Therapeut.* 58 (1), 51–66. [https://doi.org/10.1016/0163-7258\(93\)90066-M](https://doi.org/10.1016/0163-7258(93)90066-M).
- Masiulis, S., et al., 2019. GABA<sub>A</sub> receptor signalling mechanisms revealed by structural pharmacology. *Nature* 565 (7740), 454–459. <https://doi.org/10.1038/s41586-018-0832-5>.
- Meiboom, S., Gill, D., 1958. Modified spin-echo method for measuring nuclear relaxation times. *Rev. Sci. Instrum.* 29 (8), 688–691. <https://doi.org/10.1063/1.1716296>.
- Mew, E.J., et al., 2017. The global burden of fatal self-poisoning with pesticides 2006–15: systematic review. *J. Affect. Disord.* 219, 93–104. <https://doi.org/10.1016/j.jad.2017.05.002>.
- Mostallino, M.C., et al., 2006. Changes in expression of the  $\delta$  subunit of the GABA A receptor and in receptor function induced by progesterone exposure and withdrawal. *J. Neurochem.* 99 (1), 321–332. <https://doi.org/10.1111/j.1471-4159.2006.04055.x>.
- Naylor, D.E., Liu, H., Wasterlain, C.G., 2005. Trafficking of GABA<sub>A</sub> receptors, loss of inhibition, and a mechanism for pharmacoresistance in status epilepticus. *J. Neurosci.* 25 (34), 7724–7733. <https://doi.org/10.1523/JNEUROSCI.4944-04.2005>.
- Nepovimova, E., Kuca, K., 2020. Novichoks. In: *Handbook of Toxicology of Chemical Warfare Agents*. Elsevier, pp. 143–148. <https://doi.org/10.1016/B978-0-12-819090-6.00010-6>.
- Paxinos, George, Watson, C., 2007. *The Rat Brain in Stereotaxic Coordinates*. Academic Press [Preprint], London.
- Peters, A.M., et al., 2007. T2\* measurements in human brain at 1.5, 3 and 7 T. *Magn. Reson. Imag.* 25 (6), 748–753. <https://doi.org/10.1016/j.mri.2007.02.014>.
- Pouliot, W., et al., 2016. A rodent model of human organophosphate exposure producing status epilepticus and neuropathology. *Neurotoxicology* 56, 196–203. <https://doi.org/10.1016/j.neuro.2016.08.002>.
- Reddy, D.S., 2019. Mechanism-based novel antidotes for organophosphate neurotoxicity. *Current Opinion in Toxicology* 14, 35–45. <https://doi.org/10.1016/j.cotox.2019.08.001>.
- Reddy, S., et al., 2019. Neuroimaging biomarkers of experimental epileptogenesis and refractory epilepsy. *Int. J. Mol. Sci.* 20 (1), 220. <https://doi.org/10.3390/ijms20010220>.
- Reddy, S.D., et al., 2020. Magnetic resonance imaging analysis of long-term neuropathology after exposure to the nerve agent soman: correlation with histopathology and neurological dysfunction. *Ann. N. Y. Acad. Sci.* 1480 (1), 116–135. <https://doi.org/10.1111/nyas.14431>.
- Reddy, S.D., Reddy, D.S., 2015. Midazolam as an anticonvulsant antidote for organophosphate intoxication—A pharmacotherapeutic appraisal. *Epilepsia* 56 (6), 813–821. <https://doi.org/10.1111/epi.12989>.
- Roch, C., et al., 2002. Magnetic resonance imaging in the study of the lithium-pilocarpine model of temporal lobe epilepsy in adult rats. *Epilepsia* 43 (4), 325–335. <https://doi.org/10.1046/j.1528-1157.2002.11301.x>.
- Rogawski, M.A., et al., 2013. Neuroactive steroids for the treatment of status epilepticus. *Epilepsia* 54, 93–98. <https://doi.org/10.1111/epi.12289>.
- Ronchetti, E., 1985. Robust model selection in regression. *Stat. Probab. Lett.* 3 (1), 21–24.
- Rosman, Y., et al., 2012. Using MRI for the assessment of paraoxon-induced brain damage and efficacy of antidotal treatment. *J. Appl. Toxicol.* 32 (6), 409–416. <https://doi.org/10.1002/jat.1715>.



- Schenck, J.F., 1996. The role of magnetic susceptibility in magnetic resonance imaging: MRI magnetic compatibility of the first and second kinds. *Med. Phys.* 23 (6), 815–850. <https://doi.org/10.1118/1.597854>.
- Schläppi, B., 1983. Safety aspects of midazolam. *Br. J. Clin. Pharmacol.* 16 (Suppl. 1), 37S–41S. <https://doi.org/10.1111/j.1365-2125.1983.tb02269.x>.
- Shrot, S., et al., 2012. Early in vivo MR spectroscopy findings in organophosphate-induced brain damage-potential biomarkers for short-term survival. *Magn. Reson. Med.* 68 (5), 1390–1398. <https://doi.org/10.1002/mrm.24155>.
- Shrot, S., et al., 2015. Early brain magnetic resonance imaging can predict short and long-term outcomes after organophosphate poisoning in a rat model. *Neurotoxicology* 48, 206–216. <https://doi.org/10.1016/j.neuro.2015.04.003>.
- Sidell, F.R., Borak, J., 1992. Chemical warfare agents: II. nerve agents. *Ann. Emerg. Med.* 21 (7), 865–871. [https://doi.org/10.1016/S0196-0644\(05\)81036-4](https://doi.org/10.1016/S0196-0644(05)81036-4).
- Sisó, S., et al., 2017. Editor's highlight: spatiotemporal progression and remission of lesions in the rat brain following acute intoxication with diisopropylfluorophosphate. *Toxicol. Sci.* 157 (2), 330–341. <https://doi.org/10.1093/toxsci/kfx048>.
- Suleymanova, E., Gulyaev, M., Chepurnova, N., 2014. Ginseng extract attenuates early MRI changes after status epilepticus and decreases subsequent reduction of hippocampal volume in the rat brain. *Epilepsy Res.* 108 (2), 223–231. <https://doi.org/10.1016/j.eplepsyres.2013.11.018>.
- Suleymanova, E.M., Gulyaev, M.V., Abbasova, K.R., 2016. Structural alterations in the rat brain and behavioral impairment after status epilepticus: an MRI study. *Neuroscience* 315, 79–90. <https://doi.org/10.1016/j.neuroscience.2015.11.061>.
- Supasai, S., et al., 2020. Acute administration of diazepam or midazolam minimally alters long-term neuropathological effects in the rat brain following acute intoxication with diisopropylfluorophosphate. *Eur. J. Pharmacol.* 886, 173538. <https://doi.org/10.1016/j.ejphar.2020.173538>.
- Tattersall, J., 2009. Seizure activity post organophosphate exposure. *Front. Biosci.* (14), 3688. <https://doi.org/10.2741/3481>.
- Testylier, G., et al., 2007. Cerebral edema induced in mice by a convulsive dose of soman. Evaluation through diffusion-weighted magnetic resonance imaging and histology. *Toxicol. Appl. Pharmacol.* 220 (2), 125–137. <https://doi.org/10.1016/j.taap.2007.01.013>.
- Todorovic, M.S., et al., 2012. Characterization of status epilepticus induced by two organophosphates in rats. *Epilepsy Res.* 101 (3), 268–276. <https://doi.org/10.1016/j.eplepsyres.2012.04.014>.
- Tsai, Y.-H., Lein, P.J., 2021. Mechanisms of organophosphate neurotoxicity. *Current Opinion in Toxicology* 26, 49–60. <https://doi.org/10.1016/j.cotox.2021.04.002>.
- Tu, A.T., 2007. Toxicological and chemical aspects of sarin terrorism in Japan in 1994 and 1995. *Toxin Rev.* 26 (3), 231–274. <https://doi.org/10.1080/15569540701496321>.
- Voorhees, J.R., et al., 2017. Neurotoxicity in preclinical models of occupational exposure to organophosphorus compounds. *Front. Neurosci.* 10. <https://doi.org/10.3389/fnins.2016.00590>.
- White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48 (4), 817. <https://doi.org/10.2307/1912934>.
- Wu, X., Kuruba, R., Reddy, D.S., 2018. Midazolam-resistant seizures and brain injury after acute intoxication of diisopropylfluorophosphate, an organophosphate pesticide and surrogate for nerve agents. *J. Pharmacol. Exp. Therapeut.* 367 (2), 302–321. <https://doi.org/10.1124/jpet.117.247106>.
- Yamasue, H., et al., 2007. Human brain structural change related to acute single exposure to sarin. *Ann. Neurol.* 61 (1), 37–46. <https://doi.org/10.1002/ana.21024>.
- Young, R.A., Watson, A., 2020. Organophosphate nerve agents. In: *Handbook of Toxicology of Chemical Warfare Agents*. Elsevier. <https://doi.org/10.1016/B978-0-12-819090-6.00008-8>.
- Zimmer, T.S., et al., 2021. Seizure-mediated iron accumulation and dysregulated iron metabolism after status epilepticus and in temporal lobe epilepsy. *Acta Neuropathol.* 142 (4), 729–759. <https://doi.org/10.1007/s00401-021-02348-6>.

# Tumor Targeting by $\alpha_v\beta_3$ -Integrin-Specific Lipid Nanoparticles Occurs *via* Phagocyte Hitchhiking

Alexandros Marios Sofias,\* Yohana C. Toner, Anu E. Meerwaldt, Mandy M. T. van Leent, Georgios Soutanidis, Mattijs Elschot, Haruki Gonai, Kristin Grendstad, Åsmund Flobak, Ulrike Neckmann, Camilla Wolowczyk, Elizabeth L. Fisher, Thomas Reiner, Catharina de Lange Davies, Geir Bjørkøy, Abraham J. P. Teunissen, Jordi Ochando, Carlos Pérez-Medina, Willem J. M. Mulder, and Sjoerd Hak\*

Cite This: *ACS Nano* 2020, 14, 7832–7846

Read Online

ACCESS |

Metrics & More

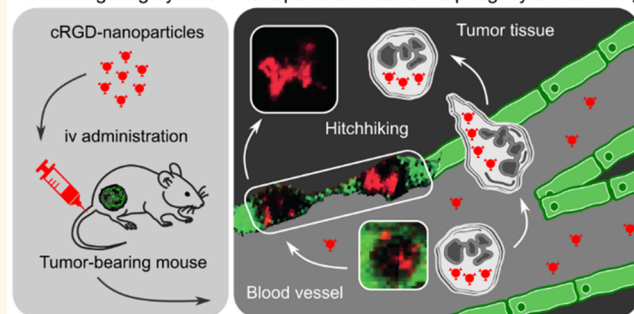
Article Recommendations

Supporting Information

**ABSTRACT:** Although the first nanomedicine was clinically approved more than two decades ago, nanoparticles' (NP) *in vivo* behavior is complex and the immune system's role in their application remains elusive. At present, only passive-targeting nanoformulations have been clinically approved, while more complicated active-targeting strategies typically fail to advance from the early clinical phase stage. This absence of clinical translation is, among others, due to the very limited understanding for *in vivo* targeting mechanisms. Dynamic *in vivo* phenomena such as NPs' real-time targeting kinetics and phagocytes' contribution to active NP targeting remain largely unexplored. To better understand *in vivo* targeting, monitoring NP accumulation and distribution at complementary levels of spatial and temporal resolution is imperative. Here, we integrate *in vivo* positron emission tomography/computed tomography imaging with intravital microscopy and flow cytometric analyses to study  $\alpha_v\beta_3$ -integrin-targeted cyclic arginine-glycine-aspartate decorated liposomes and oil-in-water nanoemulsions in tumor mouse models. We observed that ligand-mediated accumulation in cancerous lesions is multifaceted and identified "NP hitchhiking" with phagocytes to contribute considerably to this intricate process. We anticipate that this understanding can facilitate rational improvement of nanomedicine applications and that immune cell–NP interactions can be harnessed to develop clinically viable nanomedicine-based immunotherapies.

**KEYWORDS:** nanomedicine, cyclic RGD nanoparticles, immune cell hitchhiking, neutrophils, positron emission tomography/computed tomography imaging, intravital microscopy

Tumor targeting by cRGD-nanoparticles occurs *via* phagocyte hitchhiking



For medical purposes, *in vivo* nanoparticle (NP) application has become an enormous field with stakeholders ranging from individual patients to academics and large pharmaceutical corporations. Administration of drugs encapsulated in NPs can reduce systemic exposure and increase drug levels at pathological sites. Although these attractive features allowed for the first passive-targeting nanomedicine to be clinically approved more than two decades ago,<sup>1</sup> the lack of an in depth understanding of NP *in vivo* behavior remains one of the barriers for their widespread clinical use.<sup>2–4</sup>

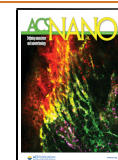
One aspect of NP *in vivo* behavior that has largely been neglected, especially for ligand-decorated formulations, is the

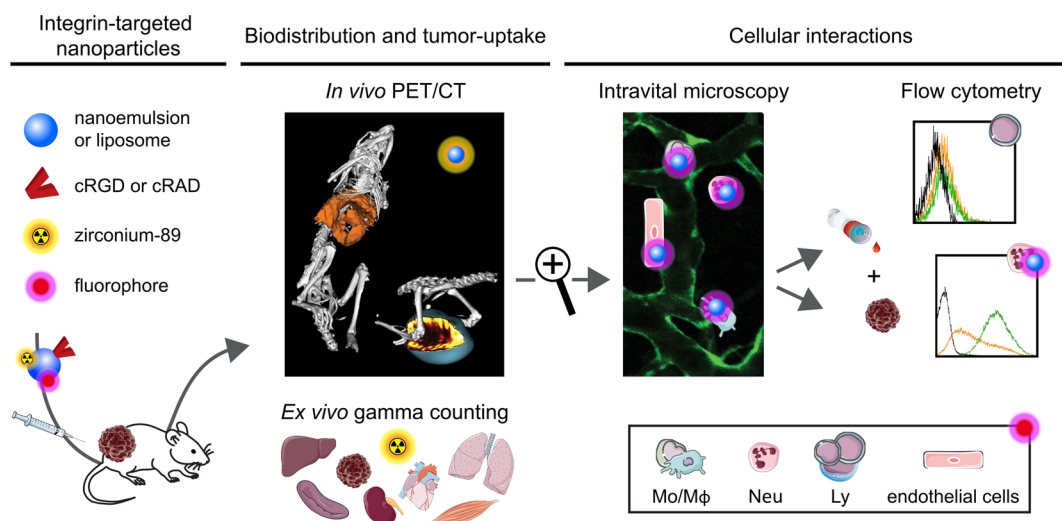
role of the immune system and its cells in NP targeting.<sup>5–7</sup> The immune system has evolved to protect its host from infections and comprises two arms: innate and adaptive immunity. Innate immunity is an ancient first-line host-defense system that primarily consists of phagocytes. These

Received: November 3, 2019

Accepted: May 15, 2020

Published: May 15, 2020





**Figure 1.** Study outline. Tumor-bearing mice were intravenously injected with  $^{89}\text{Zr}$ - or fluorophore-labeled NPs, of which the pharmacokinetics, biodistribution, and accumulation in organs were quantified using positron emission tomography imaging and *ex vivo* gamma counting. NP interactions with cells were assessed with intravital confocal microscopy of tumors and *ex vivo* flow cytometry of blood and tumor single cell suspensions. Abbreviations: Mo/M $\phi$ : monocytes/macrophages, Neu: neutrophils, Ly: lymphocytes.

cells are present in all tissues and have the capacity to rapidly, but nonspecifically, recognize and engulf foreign materials, such as nanomedicines. Upon its saturation, adaptive immunity, a more sophisticated and evolutionary modern part of the immune system, mounts a highly specific immune response and builds immunological memory. Here, we attempted to explore the immune system's roles in NP *in vivo* behavior, with a particular focus on phagocytes' contribution to active NP targeting.

As NP-targeting ligand, we selected cyclic arginine-glycine-aspartate (cRGD), which is one of the most widely used NP ligands in the field.<sup>8</sup> This is a ligand for  $\alpha_v\beta_3$ -integrin, which is upregulated on activated/angiogenic tumor vascular endothelium and several types of cancer cells.<sup>8–10</sup> Even though more than 500 preclinical studies (source: [scopus.com](https://scopus.com)) and 55 clinical trials (source: [clinicaltrials.gov](https://clinicaltrials.gov)) have attempted to establish nanoformulations containing the RGD motif, their clinical utilization remains zero. Preclinically, the targeting abilities of cRGD-NPs, as well as from other ligand-decorated nanomedicine, are mainly assessed *in vitro*, *ex vivo*, or with macroscopic *in vivo* imaging modalities.<sup>11,12</sup> This kind of data provides little insight into *in vivo* targeting dynamics, mechanisms, and potential contributions of phagocytes. Although some studies include techniques that could provide mechanistic and real-time information on the microscopic level, observations are typically limited to static “snapshots” rather than continuous dynamic tracking.<sup>13,14</sup>

These realizations were our incentive for investigating real-time the fate of two  $\alpha_v\beta_3$ -integrin-targeted lipidic NP platforms in tumor mouse models using *in vivo* positron emission tomography/computed tomography (PET/CT) imaging integrated with intravital microscopy (IVM) and flow cytometry (Figure 1 provides the study outline). Particular attention was given to the full exploration of real-time NP targeting kinetics and their specific interactions with the circulating and tumor-homing immune system. We observed that ligand-mediated NP accumulation in cancerous lesions is multifaceted and identified “NP hitchhiking” with phagocytes to contribute considerably to this intricate process.

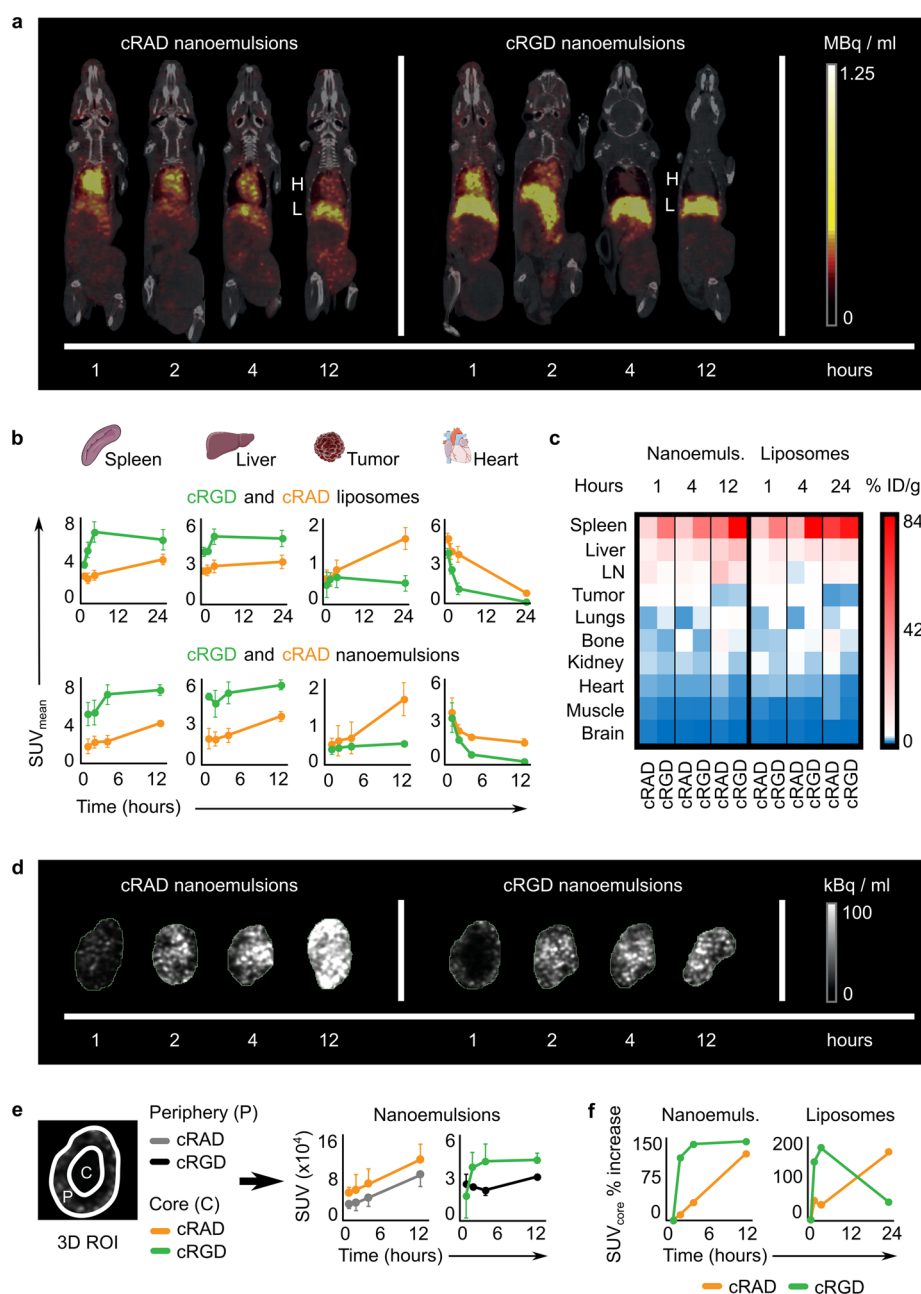
We anticipate that these insights will facilitate rational improvements of nanomedicine applications.

## RESULTS AND DISCUSSION

**Nanoparticles.** Since approximately half of the clinically approved nanomedicines are lipid-based formulations,<sup>15,16</sup> we chose to study liposomes<sup>17</sup> (100 nm) and oil-in-water nanoemulsions<sup>18</sup> (150 nm). The lipid composition is very similar to that of the clinically approved Doxil.<sup>19</sup> We surface-functionalized these NPs with  $\alpha_v\beta_3$ -integrin-specific cyclic arginine-glycine-aspartate peptides (c[RGDfK], abbreviated as cRGD).<sup>20</sup> This peptide is not only one of the most widely studied targeting ligands (illustrated by both older<sup>14,21</sup> and recent publications<sup>22–24</sup>) but also one of the few NP-targeting ligands that has been tested in patients.<sup>25–27</sup> As a nonspecific control peptide, we used cyclic arginine-alanine-aspartate peptides (c[RADfK], abbreviated as cRAD).<sup>28</sup> Figure S1a provides the NP characterization data.

**Biodistribution and Tumor Uptake.** We first quantitatively studied NP *in vivo* behavior at the macroscopic level using a combination of *in vivo* PET/CT imaging and *ex vivo* gamma counting of blood fractions and organs after intravenous administration of zirconium-89 ( $^{89}\text{Zr}$ )-labeled NPs to immunocompetent BALB/c mice bearing orthotopic 66cl4 tumors (murine mammary carcinoma). Contrary to the intuitive assumption that active targeting universally increases accumulation at target sites, cRGD-NPs' shorter blood circulation half-lives (Figure S1b) and concomitant increased liver and especially spleen uptake (as compared to cRAD-decorated analogues) diminished their targeting toward tumors (Figure 2a–c, Figures S2, S3, and S4a).

We observed that nonspecific cRAD-NPs accumulate throughout the tumors (Figure 2d), facilitated by the homogeneous vascularization of our tumor model (Figure S4b,c) and enhanced endothelial permeability. cRGD-NPs' tumor distribution pattern, especially their accumulation kinetics, was strikingly different. In the first hour following their intravenous injection,  $\alpha_v\beta_3$ -integrin-specific cRGD-NPs were found predominantly in the tumor periphery (Figure 2d,e and Figure S4d,e), a pattern we observed previously in

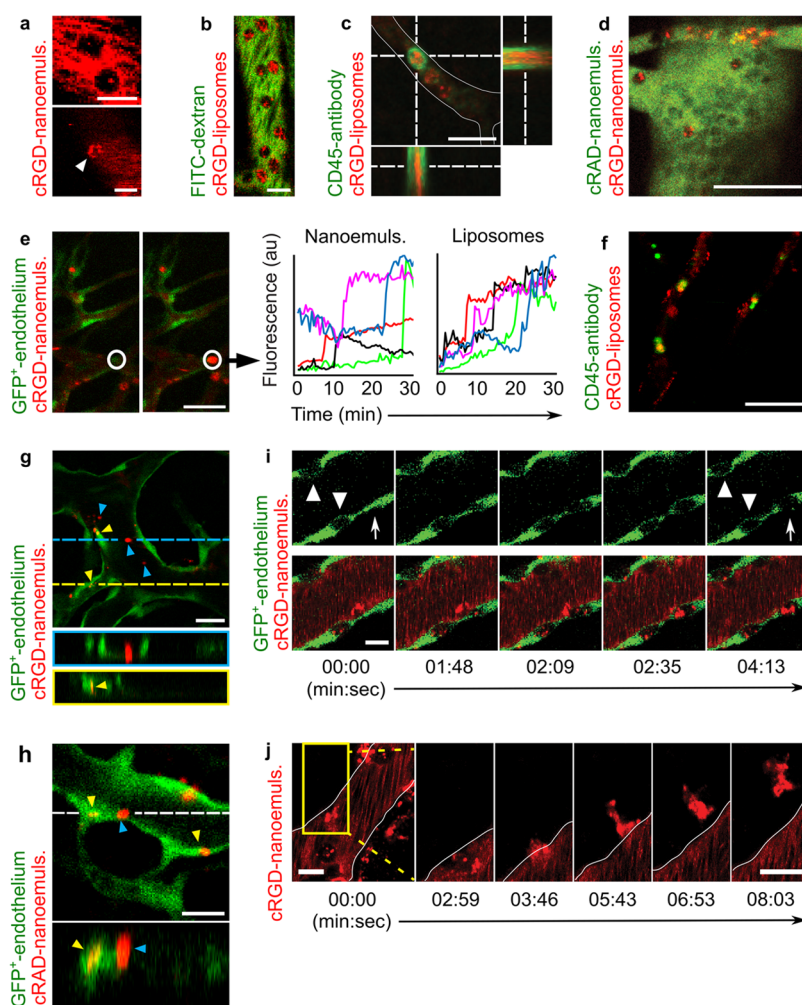


**Figure 2.** PET/CT imaging and gamma counting. (a) PET/CT images of mice injected with cRAD or cRGD nanoemulsions (liposomes in Figure S4a). Compared to cRAD-NPs, cRGD-NPs cleared faster (Figure S1b), as evidenced by more rapid signal decrease in the heart (H), and accumulated to higher extent in the liver (L). (b) Mean SUV as a function of time for spleen, liver, tumor, and heart. cRGD-NPs accumulated to higher extent in liver and spleen, whereas cRAD-NPs reached higher levels in tumors. Signal from the heart reflects the differences in circulation half-lives;  $n = 4-6$  per formulation per time point. (c) A heatmap of NP biodistribution profiles obtained with *ex vivo* gamma counting on isolated organs corroborated the PET/CT imaging (the heatmap is created based on the data in Figures S2 and S3). (d) *In vivo* PET images of tumors showing homogeneous cRAD-NP accumulation at all time-points. At 1 h post-injection, cRGD-nanoemulsions (liposomes in Figure S4d) were mainly found in the tumor periphery. (e) Tumor SUV as a function of time in the core and periphery after cRAD-NP or cRGD-nanoemulsions ( $n = 4$  per formulation per time point) administration (liposomes in Figure S4e). (f) SUV increase in the tumor core, relative to the SUV at 1 h post-injection, as a function of time. From 1 to 4 h post-injection, the cRGD-NP distribution pattern shifted from the lesion periphery to the tumor core at a much more rapid rate than cRAD-NPs. Error bars in (b) and (e): SD.

various tumor models *via* magnetic resonance imaging for cRGD-liposomes,<sup>29</sup> cRGD-nanoemulsions,<sup>30</sup> and cRGD-quantum dots.<sup>31</sup> Moreover, in a patient with liver metastasis, cRGD-conjugated silica NPs accumulated in the tumor margin as well.<sup>27</sup> Interestingly, between 1 and 4 h post-injection, the cRGD-NP distribution shifted rapidly from the tumor

periphery to its core, which we did not observe for cRAD-NPs in the same time frame (Figure 2e,f and Figure S4e). cRAD-NPs accumulated in the tumor core and periphery at similar rates (Figure 2e). cRGD-NPs however accumulated much faster in the core than in the periphery in the 1–4 h time frame (Figure 2e). Moreover, the relative accumulation



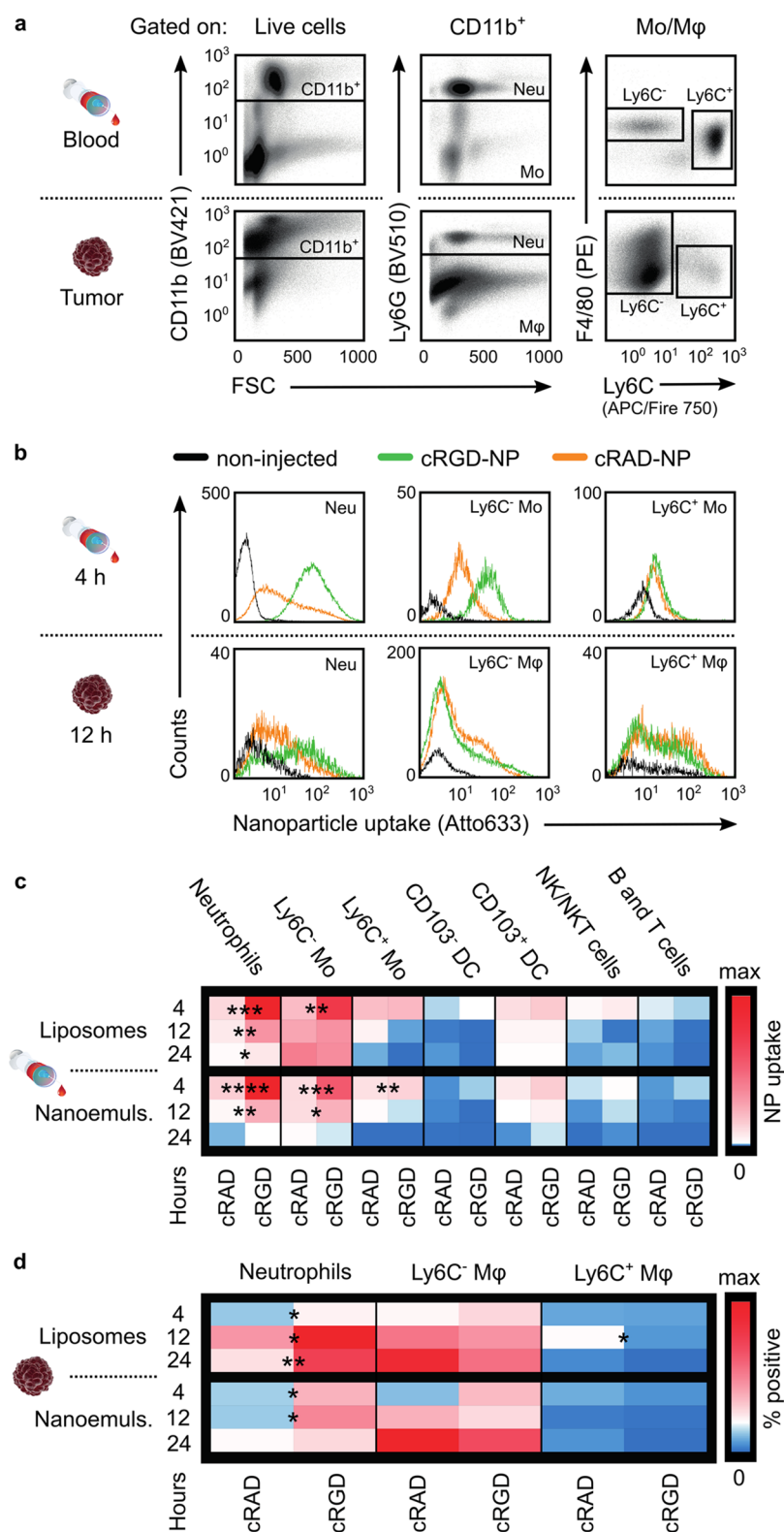


**Figure 3.** Intravital microscopy of tumors. (a) cRGD-nanoemulsions (red) agglomerates with ring-like appearances (arrowhead) and inside circulating “black holes” 15 min post-injection. (b) This phenomenon was especially apparent when FITC-Dextran (green) was co-injected with cRGD-liposomes (red, 3 h post-injection). (c) Intravital CD45-staining (green) confirmed that circulating immune cells internalize cRGD-liposomes (red, 6 h post-injection). (d) Co-injections demonstrated higher cellular uptake of cRGD-nanoemulsions (red) than cRAD-nanoemulsions (green) (35 min post-injection). (e) Frames of an imaging sequence showing a cRGD-nanoemulsion-positive cell (red) binding to tumor vasculature (green, GFP), indicated with a white region of interest (ROI). In NP fluorescence *versus* time graphs originating from such ROIs, “cell binding events” appeared as steps, further demonstrating that this binding did not result from gradual cRGD-NP accumulation. (f) A significant portion of accumulated cRGD-liposomes (red) presented in CD45<sup>+</sup> (green) cells (6 h post-injection). (g) Z-stack with *x-z* projections showing a part of cRGD-nanoemulsions (red, 24 h post-injection) to colocalize with endothelium (green, GFP), indicative of targeting (yellow arrowheads). A considerable portion of the cRGD-nanoemulsions was present in nonendothelial agglomerates (GFP negative, blue arrowheads). cRGD-liposomes in Figure S6a. (h) A portion of cRAD-nanoemulsions (red, cRAD-liposomes in Figure S6b,c) was also taken up by endothelium. (i) Frames of an imaging sequence (1.5 h post-injection) showing “black holes” (arrowheads) in the endothelium (green, GFP), positive for cRGD-nanoemulsions (red). A cRGD-nanoemulsion-positive cell entering endothelium can also be appreciated (arrow). (j) Frames of an imaging sequence showing extravasation (from an outlined vessel) of cRGD-nanoemulsion-containing cells (3 h post-injection). Scale bars: (a–c, h–j) 10  $\mu\text{m}$ , (d) 50  $\mu\text{m}$ , and (e–g) 25  $\mu\text{m}$ .

rate in the core was also faster for cRGD-NPs than for cRAD-NPs (Figure 2f). Since the cRGD-NP levels in the circulation drop rapidly in this time frame, tumors were homogeneously vascularized, and the control cRAD-NPs accumulated homogeneously throughout the tumor, this tumor accumulation pattern cannot be explained by endothelium targeting or vascular extravasation of the cRGD-NPs. Since we suspected this phenomenon was related to interactions with migrating immune cells, we next studied both lipid NPs’ *in vivo* behavior at the cellular level using a combination of real-time IVM and flow cytometry.

**Cellular Interactions Studied with Intravital Microscopy.** To study cellular interactions in real-time, we

performed high temporal and subcellular resolution IVM on 66cl4 tumors grown either in dorsal window chambers on immunodeficient TieGFP mice or orthotopically in immunocompetent Tie2GFP mice.<sup>32</sup> IVM showed that both cRGD liposomes and nanoemulsions behaved in a similar manner (Figure 3). Nevertheless, the identified patterns were strikingly different in comparison to their control-analogues cRAD liposomes and nanoemulsions. Within minutes after injection, we observed strong indications for cRGD-NP uptake by circulating immune cells (Figure 3a,b, Movies S1 and S2), which was confirmed by intravital CD45 (a membrane glycoprotein present on all leukocytes) staining (Figure 3c and Movie S3). During the first hour, we observed



**Figure 4.** Flow cytometry analysis of blood and tumor. (a) Gating strategy for myeloid cells (full gating strategy in Figure S7). (b) Representative histograms showing NP uptake by neutrophils (Neu), alternatively activated Ly6C<sup>-</sup> monocytes/macrophages (Mo/Mφ), and classically activated Ly6C<sup>+</sup> Mo/Mφ in blood (4 h) and tumor (12 h). In blood, cRGD-NPs were taken up predominantly by neutrophils and Ly6C<sup>-</sup> monocytes, and in tumors, neutrophils were the main contributor to cRGD-NP uptake (also see Figure S8). (c) Heatmap of liposomes and nanoemulsion uptake (median fluorescence intensity) in blood ( $n = 3-6$  per formulation per time point). The highest uptake was observed in neutrophils and monocytes, which also showed significant preference for cRGD-NPs. Also see Figure S9a. (d) Heatmap of NP uptake ( $n = 3-6$  per formulation per time point) by tumor-associated phagocytes. Four h post-injection, more neutrophils and Ly6C<sup>-</sup> macrophages contained cRGD- than cRAD-NPs. At later time-points, this was still the case for neutrophils, whereas more Ly6C<sup>-</sup> macrophages contained cRAD- than cRGD-NPs (also see Figure S9b). *P*-values: \* <0.05, \*\* <0.01, \*\*\* <0.001, \*\*\*\* <0.0001.

negligible cRAD-NP leukocyte uptake (Movies S4 and S5). This pattern persisted up to 24 h post-injection, with extensive cRGD-NP cellular uptake (Movies S6 and S7) and some, but significantly fewer, cRAD-NP positive circulating leukocytes (Movies S8 and S9). We corroborated these observations by studying NP coinjections with IVM (Figure 3d and Movies S10 and S11) as well as *ex vivo* microscopy on isolated leukocytes (Figure S5).

Interestingly, we observed cRGD-NP-positive cells (appearing as red clusters) to bind suddenly in the tumor vasculature (Figure 3e). By designing region of interests (ROIs) around the spots where these “red clusters” appeared and plotting the NP fluorescence as a function of time, it became evident that the binding events were a result of sudden “targeting” of an entire cluster, rather than gradual accumulation/binding of individual cRGD-NPs (Figure 3e). The use of CD45 antibodies allowed us to verify the NPs’ presence in blood vessel-associated CD45<sup>+</sup> immune cells (Figure 3c,f). Unexpectedly, but in accordance with the idea that immune cells play a significant role in cRGD-NP tumor targeting, only a fraction of accumulated cRGD-NPs were associated with (GFP-expressing) tumor blood vessel endothelial cells (Figure 3g and Figure S6a). Even more surprisingly, a small portion of accumulated cRAD-NPs directly associated with tumor vascular endothelium as well (Figure 3h and Figure S6b,c).

cRGD-NP-positive cells not only bound to the luminal endothelium but also interacted closely with (GFP-expressing) endothelial cells (Figure 3i). Finally, these cRGD-NP-loaded cells were observed to extravasate extensively into and migrate through the tumor tissue (Figure 3j, Movies S12–S16). This phenomenon, which was not observed for cRAD-NPs (Movie S17), demonstrates immune cells’ critical role in cRGD-NPs’ distribution through the tumor interstitium.

### Cellular Interactions Studied with Flow Cytometry.

The dynamic IVM in mice with endothelial GFP expression was pivotal for concluding that (CD45<sup>+</sup>) leukocytes contribute considerably to cRGD-NP tumor targeting. Moreover, the IVM observation that nanoparticle-loaded cells migrate through the tumor interstitium may also explain the cRGD-NP redistribution from the tumor periphery to the core as observed with PET imaging. To determine which immune cells were involved, we performed flow cytometry on blood and tumor single cell suspensions at 4, 12, and 24 h after NP administration to immunocompetent BALB/c mice bearing orthotopic 66cl4 tumors. Facilitated by a comprehensive antibody panel, we identified the main phagocyte (myeloid-derived immune cell) populations: neutrophils (CD45<sup>+</sup>CD11b<sup>+</sup>Ly6G<sup>+</sup>) and monocytes/macrophages (CD45<sup>+</sup>CD11b<sup>+</sup>Ly6G<sup>-</sup>Ly6C<sup>-</sup>F4/80<sup>+</sup> or CD45<sup>+</sup>CD11b<sup>+</sup>Ly6G<sup>-</sup>Ly6C<sup>+</sup>F4/80<sup>intermediate</sup>).

Monocytes constitute 5–10% of circulating immune cells. Upon entering tissues, these cells differentiate into macrophages, a key function of which is to phagocytose pathogens and extracellular debris.<sup>33</sup> Depending on the cytokines in their surroundings, these phagocytes are subject to classical or alternative activation. Classically activated monocytes/macrophages (CD45<sup>+</sup>CD11b<sup>+</sup>Ly6G<sup>-</sup>Ly6C<sup>+</sup>F4/80<sup>intermediate</sup>) usually fulfill pro-inflammatory functions, as during infection to fight pathogens. Tumor-associated macrophages (TAMs) (CD45<sup>+</sup>CD11b<sup>+</sup>Ly6G<sup>-</sup>Ly6C<sup>-</sup>F4/80<sup>+</sup>) are typically alternatively activated, similar to macrophages in wound healing, and contribute to the local immune suppression in the tumor microenvironment.<sup>33</sup>

Neutrophils are by far the most abundant leukocyte in blood (50–70% of circulating immune cells). These phagocytes are produced at very high rates (more than 10<sup>11</sup> per day in humans) and are constantly replenished at recruitment sites like inflamed tissues and solid tumors.<sup>34</sup> Largely due to their short lifespan, neutrophils have long been considered merely a first line of defense against infections, but are now recognized as important actors in cancer development and progression. Not unlike TAMs, neutrophils appear to play pivotal roles in tumor initiation, immune suppression, angiogenesis, and metastasis (as reviewed in refs 34 and 35).

Figure 4a illustrates the key gating strategy used to detect the three main phagocytic myeloid cell populations in blood and tumor single-cell suspensions. Additionally, we studied NP interactions with dendritic cells (CD11c<sup>+</sup>CD103<sup>+</sup> or CD11c<sup>+</sup>CD103<sup>-</sup>), which are innate immune cells that shape adaptive immune responses by presenting antigens. Finally, we also defined the main lymphoid-derived cell populations comprising the adaptive immune system: natural killer (CD11b<sup>-</sup>FSC<sup>low</sup>SSC<sup>low</sup>NK-1.1<sup>+</sup>), T-, and B-cells (CD11b<sup>-</sup>FSC<sup>low</sup>SSC<sup>low</sup>NK-1.1<sup>-</sup>). Figure S7 describes the full gating strategy.

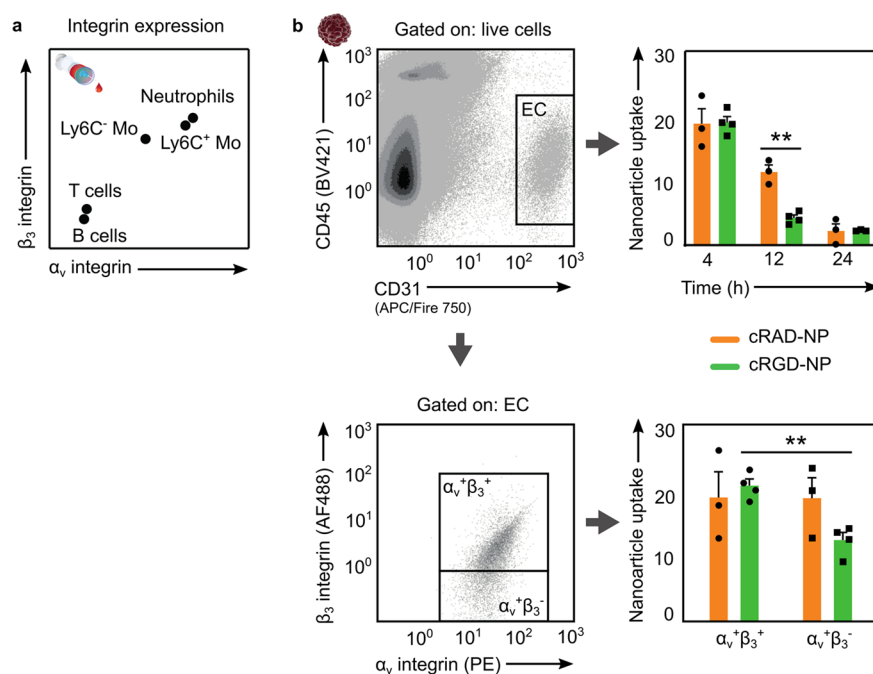
In blood, we observed extensive interactions between CD11b<sup>+</sup> myeloid cells and cRGD-NPs, which were predominantly found in Ly6G<sup>+</sup> neutrophils and Ly6G<sup>-</sup>/Ly6C<sup>-</sup> alternatively activated monocytes (Figure 4b,c and Figures S8a,b and S9a). cRAD-NPs were found in the same cell populations (Figure 4b,c), but at significantly lower levels. We observed few interactions between either of the NPs and lymphocytes (Figure 4c and Figures S8a,b and S9a). This confirmed, as expected, that phagocytic myeloid cells extensively ingested intravenously administered NPs.

CD11b<sup>+</sup> myeloid cells made up 75% of the immune cells in 66cl4 breast tumors (Figure S10). We observed striking differences between cRGD- and cRAD-NP levels in the different myeloid cell populations. Although circulating monocytes/macrophages showed a significant preference for cRGD-NPs, TAMs tended to contain higher or similar amounts of nonspecific cRAD-NPs as compared to cRGD-NPs (Figure 4b,d and Figure S8c,d and S9b). In the context of cRAD-NPs’ longer blood half-life and higher tumor uptake, this apparent TAM preference can be best explained by cRAD-NPs’ extravasation (passive accumulation) into tumor tissue and their subsequent uptake by TAMs.

Interestingly,  $\alpha_v\beta_3$ -integrin-specific cRGD-NPs displayed a notable proclivity toward neutrophils in both blood and tumor (Figure 4c,d and Figure S8 and S9). This indicates that these cRGD-NPs actively associate with neutrophils in the blood before migrating to the tumor. These clear differences in NP presence in tumor-associated neutrophils and macrophages strongly suggest that neutrophils are the main phagocytic contributor to cRGD-NPs hitchhiking to the tumor. The short life span of neutrophils, which are highly motile cells,<sup>36</sup> requires their tumor infiltration to occur rapidly and continuously. These cellular features and our observations support the idea that neutrophils are indeed capable of NP delivery to tumors and intratumoral nanomaterial distribution.

**Associating cRGD-Nanoparticle Uptake with  $\alpha_v\beta_3$ -Integrin Expression.** The interesting observations obtained by IVM and flow cytometry revealed a clear preference of myeloid cells and especially neutrophils for cRGD-NPs. The staining of all main immune cell populations with  $\alpha_v$  and  $\beta_3$  integrin subunits revealed a high co-expression of these





**Figure 5.** Association of nanoparticle uptake with  $\alpha_v$  and  $\beta_3$  integrin. (a) Integrin co-expression for the key cell populations in blood, showing high  $\alpha_v$  and  $\beta_3$  integrin co-expression by myeloid cells (see also Figure S11). (b) In tumor, endothelial cells (EC) associated with cRAD-nanoemulsions (liposomes in Figure S12) similarly as (4 and 24 h) or even more than (12 h post-injection) with cRGD-NPs. However, significant differences in cRGD-NP uptake (median fluorescence intensity) by activated ( $\alpha_v^+\beta_3^+$ ) and non-activated ( $\alpha_v^+\beta_3^-$ ) cells were detected at 4 h post-injection. Error bars: Standard error of the mean. *P*-values: \*\* < 0.01.

integrins by myeloid cells. Hence, the increased cRGD–NP interaction with phagocytes may be an effect of  $\alpha_v\beta_3$ -integrin expression (Figure 5a and Figure S11) on these cells.

Finally, in line with IVM observations, flow cytometry showed that cRGD- and cRAD-NPs were taken up by CD31<sup>+</sup> endothelial cells to similar extents. However, as expected, cRGD-NPs showed a significant preference for activated endothelial cells, characterized by dual expression of both  $\alpha_v$  and  $\beta_3$  integrin subunits (Figure 5b and Figure S12).

**General Discussion.** Taken together, the distinctly detailed results from our complementary approach have here untangled phagocytes' contribution to active cRGD-decorated lipid NP tumor targeting. cRGD-NP accumulation kinetics in the tumor core, as observed by PET imaging, could not be explained by established vascular extravasation and direct endothelium targeting mechanisms. IVM showed that cRGD-NPs hitchhike with immune cells to the tumor vasculature. Moreover, these immune cells were found to extravasate and distribute the internalized nanomaterial into the tumor. Carefully designed flow cytometry experiments corroborated these findings and revealed neutrophils to be the main phagocytic contributor to this intricate process.

Interestingly, we also observed extensive uptake of cRGD-NPs in circulating immune cells in healthy Balb/c mice in a previous study.<sup>37</sup> Furthermore, in healthy tissue, these cRGD-NP loaded circulating immune cells behaved as can be expected; freely circulating as well as some of them rolling along the vessel wall.<sup>38,39</sup> However, no extravasating immune cells were observed (Figure S13 and Movies 18 and 19). Taken together this strongly suggests that the cRGD-NP uptake by myeloid cells occurs independent of cancer and that hitchhiking depends on (patho)physiological immune cell infiltration.

The difference between our study and older studies attempting to explain or exploit the cRGD-mediated targeting is that our highly complementary approach allowed for gaining information from macroscopic to submicron level. High spatial resolution IVM revealed cRGD-NP clusters bound to tumor vasculature (e.g., Figure 3g,h), which we<sup>40</sup> and others<sup>13,41–44</sup> have observed before and interpreted as cRGD–NP interaction with endothelium. The high IVM temporal resolution (0.2–0.3 s) we utilized allowed for recording very rapid events and detecting fluorescent clusters in the circulation that were not NP aggregates, but immune cells that had endocytosed cRGD-NP. Subsequent flow cytometric analysis enabled us not only to discover the extensive association between immune cells and cRGD-NPs but also to establish “immune cell hitchhiking” as an important mechanism through which tumor targeting of lipidic cRGD-NPs can be mediated.

Particularly for cRGD-NP, the contribution of myeloid cells and specifically neutrophils to tumor targeting has not been investigated before. However, interestingly, a few other studies also report myeloid cell proclivity for cRGD-decorated-liposomes as well as for 1200 nm iron oxide loaded vesicles and show this phenomenon may be used as a tool to permeate the blood brain barrier.<sup>45,46</sup> This establishes the important role of phagocytes in the *in vivo* behavior of a wide variety of cRGD-decorated lipidic systems. Furthermore, it is well established that decoration of polymeric/metal/lipidic NPs with cRGD or other ligands does not necessarily lead to an improved tumor targeting.<sup>47–49</sup> Although in these studies the role of the immune system was not specifically studied, we believe that a deep investigation into this aspect of the *in vivo* behavior of ligand-decorated nanomedicine might “revive” their *in vivo* utilization and allow for alternative *in vivo* applications.



Although we gained these insights into the cRGD-NP tumor targeting, additional methodologies must be developed to quantify the relative contribution of NP hitchhiking. Emerging disciplines like nanoinformatics<sup>50</sup> combined with computational advances may facilitate this.<sup>51</sup> For example, machine learning-based algorithms may be able to distill and attribute relevance to the different kinetic targeting processes from multidimensional data sets.<sup>52</sup>

cRGD-decorated NPs are extensively applied for therapeutic purposes aimed to deliver drugs to tumors and angiogenic vasculature.<sup>8,21,53,54</sup> First of all, we here demonstrate that in terms of total NP levels in the tumor, cRGD NPs could only “match” (at early time points post-injection only) but never “exceed” those of cRAD-NPs. Moreover, our results suggest that NP cRGD decoration will also result in elevated drug levels in both circulating and angiogenesis-associated phagocytes. This stresses the importance of a careful consideration of the pathological conditions in which cRGD-conjugated drug delivery systems may be useful as well as what kind of drugs to incorporate in such formulations. In the case of conventional cytostatic nanomedicines, cRGD decoration could actually reduce the amount of therapeutic agent ending up in the tumor and intensify adverse side effects such as neutropenia. Instead, the selection of an immunomodulatory payload that will aim to alter myeloid immune cells toward an antitumor phenotype<sup>55</sup> or to specifically inhibit an immune subpopulation<sup>56</sup> seems more appropriate.

Another widely studied application for cRGD-NPs is molecular imaging of tumor angiogenesis.<sup>29,57–59</sup> We here showed that the accepted mode of angiogenesis targeting with cRGD-NPs on which this application is based, namely direct binding to integrin positive endothelium, only holds partially. Nevertheless, several studies report good correlation between *in vivo* imaging readouts and microvessel density<sup>29</sup> or other *ex vivo* analysis.<sup>58</sup> It is well established that neutrophils and monocytes interact with activated endothelium and accumulate at inflamed and angiogenic sites. Hence, on typically used macroscopic imaging modalities like MRI, or PET, cRGD-NP accumulation with these phagocytes in angiogenic/inflamed tissues could still report quite accurately on angiogenic activity, explaining the relative success of preclinical cRGD-NP application for angiogenesis imaging.

A number of recent studies support the insight that phagocytes can affect NP *in vivo* performance considerably. TAMs have been shown to internalize nanodrugs, migrate to hypoxic tumor regions,<sup>60</sup> and release the drug payload to neighboring tumor cells.<sup>6</sup> Furthermore, neutrophil-facilitated migration of NPs from circulation has been observed in inflamed<sup>61</sup> and cancerous tissues.<sup>5</sup> Importantly, recent studies in cancer patients demonstrate that significant interactions between nanomedicines and monocytes affect nanodrug pharmacokinetics and pharmacodynamics.<sup>62,63</sup> Hence, as we here showed for cRGD-decorated lipid NPs, phagocytic immune cells can play a notable yet poorly understood role in nanodrug *in vivo* applications. To facilitate clinical translation of ligand-conjugated formulations, which despite enormous efforts has not yet occurred, and to increase nanomedicines' clinical utility in general, we consider thoroughly assessing NP *in vivo* behavior, as we have presented here, to be critically important. Possibly, screening of NP libraries composed of different types of NPs (particle size, composition, shape, surface chemistry, charge, and ligand decoration) may also contribute to stratify NPs for different

applications including immune cell, vascular, and tumor delivery.<sup>64</sup>

## CONCLUSIONS

To conclude, using an integrative approach combining highly complementary experimental tools, we studied the complex *in vivo* fate of cRGD-decorated NPs. We established “phagocyte hitchhiking” as an important mechanism for cRGD-NP active tumor targeting. This demonstrates that despite RGD-NPs' widespread use and enrollment in clinical trials, we have a lot left to learn about their *in vivo* behavior. From a broader perspective, we believe that thorough investigations into NP *in vivo* behavior, like we presented here, are important for the future of nanomedicine. The resulting understanding and insights into their interaction with and effects on the immune system will be pivotal in the development of improved *in vivo* applications. Finally, as immunotherapy continues its recent rise, harnessing immune cell–NP interactions can become a viable strategy for developing immunomodulating nanomedicines.

## EXPERIMENTAL SECTION

**Experimental Design.** The objective of the current study was to obtain a detailed understanding of the role for phagocytes in ligand-mediated tumor targeting. To do so, we synthesized two different lipidic NPs conjugated with cRGD, a ligand widely used to target  $\alpha_v\beta_3$ -integrin expressed on activated tumor vasculature. To address cRGD-NPs' biodistribution and 66cl4 (murine mammary carcinoma) tumor targeting, we utilized *in vivo* PET imaging and *ex vivo* gamma counting. cRGD-NPs' interactions with phagocytes and other cells were assessed using real-time IVM and *ex vivo* flow cytometry on blood and tumor single cell suspensions. Figure 1 and Figure S14 provide a study outline and an overview over the number of animals used throughout the study, respectively.

**Nanoparticle Materials.** 1,2-Distearoyl-*sn*-glycero-3-phosphocholine (DSPC), cholesterol, and 1,2-distearoyl-*sn*-glycero-3-phosphoethanolamine-*N*-[methoxy(polyethylene glycol)-2000] (PEG2000-DSPE) were purchased from Avanti Polar Lipids, Inc. cRGD-PEG2000-DSPE and cRAD-PEG2000-DSPE were purchased from SyMO-Chem (Eindhoven, The Netherlands). Fluorescent lipids, 1,2-dioleoyl-*sn*-glycero-3-phosphoethanolamine labeled with Atto 633 (Atto633-DOPE) and 1,2-dipalmitoyl-*sn*-glycero-3-phosphoethanolamine-*N*-(lissamine rhodamine B sulfonyl) (Rho-PE) were purchased from Sigma-Aldrich and Avanti Polar Lipids, respectively. Soybean oil and solvents were purchased from Sigma-Aldrich.

**Nanoparticle Synthesis.** NPs, liposomes, and nanoemulsions were synthesized according to established methods<sup>37,40</sup> (Table S1). Liposomes consisted of DSPC:cholesterol:PEG2000-DSPE:X-PEG2000-DSPE (where X is cRGD or cRAD) at 62:33:4:1 molar ratios. For radiolabeling and fluorescent labeling, we, respectively, added 0.5 mol % DFO-DSPE and 0.2 mol % fluorescent lipid (Atto633-DOPE or Rhodamine-PE). To obtain nanoemulsions, identical lipid mixtures were used, and 2.5 mg of soybean oil per  $\mu\text{mol}$  of lipids was added. All lipids and oil (in case of nanoemulsions) were dissolved in 4:1 chloroform: methanol. The lipid mixture was dripped slowly into 2 mL of preheated phosphate buffered saline (PBS) of pH 7.4, at 70 °C, under vigorous stirring (700 rpm) upon which solvents evaporated. The obtained crude NPs were downsized using tip sonication (30 W and 20 kHz; 15 min/25% duty cycle for 20  $\mu\text{mol}$  lipids for liposomes, and 25 min/50% duty cycle for 20  $\mu\text{mol}$  lipids for nanoemulsions). NP size (diameter, nm), dispersity ( $\mathcal{D}$ ), and  $\zeta$  potential (mV) were determined using dynamic light scattering (DLS, Malvern's Zetasizer Nano). Liposomes and nanoemulsions were 100 and 150 nm, respectively, with  $\mathcal{D} < 0.15$ , and  $\zeta$  potential around  $-20$  and  $-30$  mV, respectively (Figure S1a). NP suspension had a final lipid concentration of 10

mM, and in all experiments, mice received an intravenous lipid dose of 80  $\mu\text{mol}$  per kg, resulting in injection volumes of approximately 200  $\mu\text{L}$ .

**Nanoparticle Radiolabeling.** For the NP radiolabeling and the preparation of the phospholipid chelator DSPE-DFO, similar methods were used as previously described.<sup>19,65,66</sup>  $^{89}\text{Zr}$ -oxalate was produced at Memorial Sloan Kettering Cancer Center (New York) on an EBCO TR19/9 variable-beam energy cyclotron (EBCO Industries Inc., BC, Canada) via the  $^{89}\text{Y}(p,n)^{89}\text{Zr}$  reaction and purified in accordance with previously reported methods to yield  $^{89}\text{Zr}$  with a specific activity of 195–497 MBq/ $\mu\text{g}$ .<sup>66</sup> Activity measurements were made using a Capintec CRC-15R Dose Calibrator (Capintec, Ramsey, NJ).  $^{89}\text{Zr}$ -oxalate was neutralized with aqueous  $\text{Na}_2\text{CO}_3$  (1 M), added to a solution of 0.5 mol % DFO-bearing liposomes or nanoemulsions in PBS, and stirred at 37 °C for 2 h. The labeled NPs were separated from free  $^{89}\text{Zr}$  using a PD-10 desalting column (GE Healthcare), with PBS as the eluent and collecting 0.5 mL fractions. The radiochemical yield was typically >90%. The fractions containing most activity (~1.5–2.0 mL) were used for the experiments. These had a radiochemical purity of >98%, as determined by SEC radio-HPLC analyses performed on a Shimadzu system equipped with a Superdex 10/300 SEC column and a Lablogix Scan-RAM radio-TLC/HPLC detector. A flow rate of 1 mL/min was used with demineralized water as the eluent.

**Animals.** For PET/CT, gamma counting, histology, and flow cytometry experiments, about 200 female BALB/c mice aged 8–12 weeks were purchased from Janvier Laboratories or The Jackson Laboratory. For IVM, we crossbred STOCK Tg(TIE2GFP)287Sato/J mice (expressing green fluorescent protein in vascular endothelial cells, strain 003658, The Jackson Laboratory) and BALB/c nude mice (Foxn1nu/nu, Envigo). Obtained breeder pairs consisted of immunodeficient (Foxn1nu/nu) male TIE2GFP mice and immunocompetent TIE2GFP females (Foxn1nu/+, immunodeficient females do not reproduce). Mice were kept under pathogen-free conditions at 20 °C, 50–60% humidity, and 65 air changes per hour and were allowed food and water ad libitum. All procedures were approved by the Norwegian Animal Research Authorities (IVM and flow cytometry) or by the Institutional Animal Care and Use Committees of the Icahn School of Medicine at Mount Sinai (PET/CT and gamma counting).

**Tumor Model.** The 66cl4 cell line (derived from 4T1 murine mammary carcinoma) was obtained from Barbara Ann Karmanos Cancer Institute. The cells were cultured in DMEM (Lonza, BioWhittaker, BE-604F) supplemented with 10% fetal calf serum (FCS, Thermo Fischer Scientific, #10270-106), 2 mM L-glutamine (Lonza Group, DE-17-605E), and 50 U/mL penicillin-streptomycin (Thermo Fischer Scientific, Gibco, #15070-063). The cells were incubated at 37 °C with 5%  $\text{CO}_2$ .

All mice were orthotopically inoculated into the fourth mammary fat pad with  $5 \times 10^6$  66cl4 cells in 50  $\mu\text{L}$  of sterile NaCl. The tumor growth and weight were regularly monitored. Animals were used for experiments at 14–21 days post-inoculation when tumors had an approximate volume of 250  $\text{mm}^3$ .

Window chambers were implanted on immunodeficient TIE2GFP female mice with similar procedures as previously reported.<sup>67</sup> The day after chamber implantation,  $1.5 \times 10^6$  66cl4 cells were implanted in the chamber. 12–16 days post-chamber implantation, when tumors filled 30–90% of the window area, the animals were subjected to IVM experiments. For IVM on orthotopically grown tumors, we implanted 66cl4 cells in 50  $\mu\text{L}$  1:1 sterile NaCl/matrigel (734-1101, Corning) in immunocompetent TIE2GFP mice. The 66cl4 tumor model immunoprofiling, based on data collected from tumor single cells suspensions ( $n = 57$ ), is given in Figure S10.

**Gamma Counting Experiments.** To determine pharmacokinetic profiles (Figure S1b), 20 female BALB/c mice bearing 66cl4 orthotopic tumors were administered NPs (80  $\mu\text{mol}$  lipid per kg; 108.2  $\pm$  2.7  $\mu\text{Ci}$  for liposomes, 45.5  $\pm$  6.2  $\mu\text{Ci}$  for nanoemulsions;  $n = 5$  per NP) via the lateral tail vein. At 1, 5, 30 min and 1, 2, 4, 6, 12/24 h, about 5  $\mu\text{L}$  of blood was collected from a small incision in the lateral tail vein. The blood samples were placed in preweighed

polystyrene gamma-counting tubes and gamma counted using a Wizard2 1-Detector Gamma Counter, PerkinElmer. The obtained gamma counts were plotted as %ID/g versus time, and the circulation half-life values were determined via monoexponential fitting with Matlab.

To determine NP biodistribution profiles, 60 female BALB/c mice bearing 66cl4 orthotopic tumors received NPs (80  $\mu\text{mol}$  lipid per kg; 104.2  $\pm$  7.8  $\mu\text{Ci}$  for liposomes, 53.8  $\pm$  7.9  $\mu\text{Ci}$  for nanoemulsions;  $n = 15$  per NP) via the lateral tail vein. At 1, 4, 12 h for nanoemulsions or 1, 4, 24 h for liposomes ( $n = 4$ –6 per NP per time point), mice were sacrificed, and 0.5–1 mL of blood was collected via heart puncturing. Then, all mice were perfused with PBS (10–15 mL) until all blood was visibly perfused out of liver and kidneys. Tumor, liver, spleen, kidneys, brain, bone, muscle, celiac lymph nodes, lungs, and heart were collected from each animal and placed in preweighed polystyrene gamma-counting tubes. The activity from each organ was determined by using a Wizard2 1-Detector Gamma Counter, PerkinElmer. Analysis of gamma counts revealed the biodistribution profile for all four NPs at three time points (Figure 2c and Figure S2).

To determine the absolute number of NPs that accumulated in the tumor, the %ID/g plots were transformed to number of NPs/g of tissue by considering the surface covered by an individual lipid (0.64, 0.25, and 0.72  $\text{nm}^2$  for DSPC, cholesterol, and PEG<sub>2000</sub>-DSPE, respectively), their molar ratio in the final recipe (62, 33, 5% for DSPC, cholesterol, and PEG<sub>2000</sub>-DSPE respectively), and the radius of the different NPs (50 and 75 nm for liposomes and nanoemulsions). Based on these, the total number of lipids per NP (121 and 137 K lipids per liposome and nanoemulsion respectively), the total injected amount (80  $\mu\text{mol}$  lipid per kg), and the absolute number of NPs reaching the tumor at the various time points were calculated (Figure S3).

**PET/CT Acquisition and Reconstruction.** All 80 mice used for pharmacokinetic and biodistribution profile determinations underwent PET/CT at 1, 2, 4, and 12/24 h ( $n = 5$  per NP per time point, totaling 80 scans). Before the scan, mice were anesthetized with isoflurane (Baxter Healthcare, Deerfield, IL)/oxygen gas mixture (2% for induction, 1% for maintenance), and subsequently imaged on a Mediso nanoScan PET/CT scanner (Mediso, Budapest, Hungary). The PET acquisition time for all scans was 20 min. Prior to the PET acquisition, a 3 min whole-body CT scan was performed (energy 50 kVp, current 180  $\mu\text{As}$ , isotropic voxel size at 0.25  $\times$  0.25 mm). The coincidences were filtered with an energy window between 400 and 600 keV. The voxel size was isotropic with 0.6 mm width, and the reconstruction was applied for two full iterations, six subsets per iteration. PET data were reconstructed using CT-based attenuation correction. Reconstruction was performed using the TeraTomo 3D reconstruction algorithm from the Mediso Nucline software. Immediately after the PET/CT scan, animals were euthanized.

**PET/CT Analysis.** For NP biodistribution comparisons (Figure 2a,b and Figure S4a) for each scanned animal 11–12 ROIs for liver and tumor, 6–7 ROIs for spleen and heart covering the whole organ on a 2–3 slices step were averaged. For NP spatial distribution in the tumor (Figure 2d–f and Figure S4d,e), whole tumors and tumor cores (defined as the volumes at half the tumor diameter, resulting in approximately 20% of total tumor volume) were defined to obtain the standardized uptake value (SUV) in corresponding volumes. For ROI designing and activity quantification OsiriX MD software was used.

For PET/CT image presentation, 3D Slicer software and InVesalius software were used. For the 3D reconstructions of animals injected with cRGD-NPs (Figure 1), the scanning bed was masked out from the CT images, after which 3D renderings of the skeletal bones were produced. These were overlaid with 3D renderings of the tumor and biodistribution of the NPs as visualized by PET. For the biodistribution images (Figure 2a and Figure S4a), the CT images were window-leveled from –1000 to 2000 HU and the PET images were window-leveled from 0 to 1.25 for nanoemulsions (Figure 2a) or 0 to 3 MBq/mL for liposomes (Figure S4a). Finally, the CT images were overlain with their

corresponding PET images at 50% opacity. For presentation of NP distribution through the tumors (Figure 2d and Figure S4d), the tumor was manually delineated in a single coronal slice approximately through the middle of the tumor. The intensity of the PET images was window-leveled from 0 to 100 for nano-emulsions and from 0 to 250 kBq/mL for liposomes.

**Intravital Microscopy.** In all IVM experiments, mice were anesthetized (subcutaneous injection of a mixture of fentanyl (0.05 mg/kg), medetomidine (0.5 mg/kg), midazolam (0.5 mg/kg), and water (2:1:2:5) at a dose of 0.1 mL per 10 g of body weight) and cannulated in the tail vein. Window chamber mice were fixed on a custom heated microscope stage. Directly after intravenous administration of a NP bolus (80  $\mu$ mol lipid per kg,  $n = 3-5$  per NP), laser power and detector gain were adjusted, and from 1 to 2 min post-injection, animals were imaged dynamically (temporal resolution of 30 s) for at least 30 min. To study uptake in circulating cells, we employed high-speed imaging with temporal resolution of 0.3–2 s at various time points post-injection. Static images and z-stacks were obtained at various time points up to 24 h post-injection as well. To study differences between cRGD- and cRAD-formulations *in vivo*, we co-injected mice with cRGD- and cRAD-NPs labeled with ATTO633-PE and rhodamine-PE, respectively (40  $\mu$ mol lipid per kg for each formulation,  $n = 2$  for both NE and liposomes). For IVM on orthotopic tumors, an incision around 70% of the tumor was made. The skin and tumor were carefully detached from the mouse such that a ‘skinflap’ with the tumor on it was obtained. Using an operating microscope, the fat tissue surrounding the tumor was carefully removed, without damaging blood vessels. Subsequently, the tumor was submerged in a small ‘bath’ of PBS (contained in a rubber ring glued to a coverglass) with the skin side facing upward. The tumor was fixed using a second rubber ring and tape. For IVM on healthy ear tissue, the mice ( $n = 2$  per NP combination) were placed in a supine position, and the ear was placed on a 0.17 mm microscope glass slide and embedded in glycerine. A cover glass was carefully taped on top of each ear to level the skin on the glass slide. Care was taken to allow blood flow through the ear after fixation. The cover glass was stabilized with tape. Finally, the mice were placed on a heated microscope stage, injected with NPs ( $n = 2$  per NP), and imaged up to 3 h post-injection. IVM was performed on a confocal laser scanning microscope (Leica SP8) using a 20 $\times$ /0.5 air objective in case of window chambers and a 25 $\times$ /1.95 water immersion objective in case of orthotopic tumor imaging. This system is equipped with a tunable white light laser as well as a tunable bandpass detection system. Brilliant Violet 421-CD45-antibody (4  $\mu$ g, intravenously injected to stain circulating immune cells) was excited with 405 nm and detected at 420–470 nm. GFP and Alexa Fluor 488-Dextran (2 MDa, intravenously injected to delineate circulating cells that do not take up dextran) were both excited at 488 nm, but detected at 500–515 nm and 500–550 nm, respectively. ATTO633 was excited at 633 nm and detected at 645–700 nm, and rhodamine was excited at 560 nm and detected at 570–600 nm. Obtained images and image stacks were analyzed and prepared for publication using ImageJ, and graphs were plotted with Matlab.

**Ex Vivo Microscopy.** To study NP uptake in circulating cells with *ex vivo* confocal laser scanning microscopy (CLSM), we co-injected cRGD- and cRAD-NPs labeled with ATTO633-PE and rhodamine-PE, respectively (40  $\mu$ mol lipid per kg for each formulation). To ensure fluorescent labeling did not affect our read-out, we also co-injected oppositely labeled formulations. We collected 0.5–1 mL of blood *via* a heart puncture at 5 or 10 min or 2 h post-injection in 1 mL of 1 IU/mL heparin in 0.9% NaCl. White cells were isolated using Lymphoprep (Nycomed Pharma) according to the manufacturer’s instructions, concentrated using centrifugation, transferred to ibidi wells, and imaged directly with the Leica SP8 using the same laser and detector configurations as during IVM. Images and z-stacks were acquired using a 40 $\times$ /1.1 water objective. Figure S5 reports the results for *ex vivo* CLSM on isolated white cells.

**Flow Cytometry.** For flow cytometrical analysis, mice bearing 66cl4 orthotopic tumors were injected with NPs (80  $\mu$ mol lipid per

kg,  $n = 3-6$  per NP per time point). At 4, 12, and 24 h post-injection mice were sacrificed, and approximately 1 mL of blood was collected from the heart and placed in 15 mL tube containing 1 mL of heparin 1 IU in 0.9% NaCl. Then, mice were perfused with PBS (10–15 mL) until all blood was visibly perfused out of liver and kidneys. Tumors were resected, weighted, and placed in serum-free RPMI on ice until ready to mince and digest. To prepare blood samples for flow cytometry, red cell lysis buffer (1 part 0.17N Tris (pH = 7.65) mixed with 50 parts 0.16N NH<sub>4</sub>Cl, final pH = 7.2) was made fresh, and red cells were lysed (1:5 blood volume/lysis buffer volume) at room temperature for 5 min. The tube with blood and lysis buffer was filled with serum-free RPMI, and the suspension was centrifuged at 340 g for 8 min at 4 °C. The supernatant was aspirated out, and the lysis process was repeated once. The obtained immune cell pellet was resuspended in 500  $\mu$ L of FACS buffer (PBS, supplemented with 2% fetal calf serum and 2 mM EDTA) and incubated with 5.0  $\mu$ g of TruStain fcX (antimouse CD16/32) antibody for 10 min on ice.

To obtain tumor single cell suspensions, tumors were minced to 0.5 mm pieces in an uncoated Petri dish containing 2 mL of serum-free RPMI using a scalpel. The minced tumor was placed in a tube containing 4 mL of enzyme solution (120  $\mu$ L Liberase DL (0.835 U/mL), 240  $\mu$ L Liberase TL (0.835 U/mL), and 40  $\mu$ L DNase I (13 U/mL) and incubated for 30 min at 37 °C on a shaking incubator. Then, the sample was filtered through a 70  $\mu$ m filter. Tumor pieces not able to go through the filter were further ground using a syringe plunger. The tumor single-cell suspension was washed with serum-free RPMI, and one red cell lysis step was performed. The cells were suspended in 500  $\mu$ L of FACS buffer and incubated with 5.0  $\mu$ g of TruStain fcX (antimouse CD16/32) antibody for 10 min on ice. 100  $\mu$ L of blood and tumor samples were incubated for 30 min on ice with fluorescent antibodies, as described in Table S2.

Right before the flow cytometry run, each sample was supplemented with one test of 7-AAD live/dead marker (A1310, ThermoFisher Scientific). For compensating fluorophore detection, single color samples were run under the same laser voltage conditions. The flow cytometry data were analyzed using Kaluza software. Cellular fragments and debris were gated out of the analysis by utilizing forward and side angle light scatter signal. As a negative control for determination of NP uptake, blood and tumor single cell suspensions from non-injected mice were used. Panel 1 was used to determine immune cell  $\alpha_v$  and  $\beta_3$  integrin subunit expression (Figure 5a and Figure S11). Panels 2–4 (for which the gating strategy is presented in Figure S7) were utilized for studying the cRGD- and cRAD-NP uptake by circulating immune cells (Figures S8a,b and S9a), tumor associated immune cells (Figures S8c,d and S9b), and tumor endothelial cells (Figure S12). The antibodies were purchased from Biolegend (Brilliant Violet 421-CD11b, clone M1/70, 101235; Brilliant Violet 421-CD45, clone 30-F11, 103133; Brilliant Violet 510-Ly6G, clone 1A8, 127633; Alexa Fluor 488-CD61, clone 2C9.G2, 104311; Alexa Fluor 488-CD103, clone 2E7, 121408; PE-CD51, clone RMV-7, 104105; PE-F4/80, clone BM8, 123110; PE/Cy7-CD19, clone 6D5, 115520; PE/Cy7-NK-1.1, clone PK136, 108714; PE/Cy7-CD31, clone MEC13.3, 102524; Alexa Fluor 700-CD11c, clone N418, 117320; Alexa Fluor 700-CD45, clone 30-F11, 103128; APC-Fire750-Ly6C, clone HK1.4, 128046; APC-Fire750-CD31, clone MEC13.3, 102528) or ThermoFisher (APC-eFluor780-CD3e, clone HK1.4, 128046) (Table S3). To evaluate NP uptake and determine percent positive cells, we set the border between negative and positive cells such that the negative control cells were <2% positive. For the reported  $X_{\text{median}}$  values, the following calculation was conducted:

$$X_{\text{median reported}} = X_{\text{median of cells from mice injected with nanoparticles}} - X_{\text{median of cells from non-injected mice}}$$

**Histology.** For histological analysis of tumors, 18 BALB/c mice were inoculated orthotopically with  $5 \times 10^6$  66cl4 cells. 14–21 days post-inoculation, mice were sacrificed and perfused with 10–15 mL PBS. Tumors were collected and embedded with OCT compound



and snap frozen in 2-methylbutane (M32631, Sigma-Aldrich Inc.) 2-Methylbutane was cooled down to  $-20\text{ }^{\circ}\text{C}$  in dry ice for approximately 1 min. The samples were placed on  $-80\text{ }^{\circ}\text{C}$  overnight. Next, the tissue was cut in a cryostat at  $8\text{ }\mu\text{m}$  in consecutive sections.

For histological staining, sections were fixed in 95% ethyl alcohol for 15 s. Samples were then dipped 10 times into formalin 10%, (phosphate buffered, SF100-4 10% buffered formalin phosphate, Fisher Scientific) and rinsed in water (10 dips). Slides were stained with hematoxylin stain (Part 1201A hematoxylin stain, Harris Modified, Newcomer Supply) for 30 s and washed in two changes of distilled water for 10 dips, followed by 95% ethyl alcohol for 10 dips. The counterstaining was done in eosin Y working solution (Part 1072A eosin Y working solution, Newcomer Supply) for 15 s, and the sections were subsequently dehydrated in two changes of 95% ethyl alcohol and two changes of 100% ethyl alcohol (UN1170 ethanol 200 proof Anhydrous, Decon Laboratories) (10 dips each). Slides were then cleared in two changes of xylene (10 dips each). The coverslip was done in cytooseal XYL (8312-4 Cytooseal XYL, Richard-Allan Scientific). For imaging hematoxylin and eosin (H&E) stained slices, a Zeiss Axio Imager Z2M with a  $10\times/0.3$  air and a  $20\times/0.8$  air objective were used (Figure S 4d).

For immunofluorescence staining, tumor slices neighboring the H&E stained sections were fixed in  $-20\text{ }^{\circ}\text{C}$  acetone for 10 min and washed twice with PBS for 5 min. The slices were blocked with TBS buffer (37581 SuperBlock Blocking Buffer, Life Technologies) for 20 min at room temperature in a humid chamber. After the samples were washed with PBS, slides were incubated with  $5\text{ }\mu\text{g}/\text{mL}$  Alexa Fluor 594 antimouse CD31 antibody (102432, Biolegend) in TBS buffer overnight at  $4\text{ }^{\circ}\text{C}$ . This was followed by three wash steps of 5 min each with PBS. Excess of liquid was removed and slides were mounted with VECTASHIELD Vibrance Antifade Mounting Medium with DAPI (H-1800-10, Vector laboratories). Fluorescence-labeled tumor slices were imaged using a Leica SP5 DMI confocal microscope equipped with  $20\times/0.7$  air,  $40\times/1.25$  oil, and  $63\times/1.4$  oil objectives. DAPI and Alexa Fluor 594 were excited with a UV diode at 405 nm and a DPSS 561 nm laser and detected at 430–490 nm and 590–630 nm, respectively (Figure S4c). Both H&E and CD31 staining revealed an homogeneous physiology in the 66c14 tumors, the absence of a necrotic core, and homogeneous tumor vascularization (Figure S4b,c).

**Statistical Analysis.** All data are presented as mean  $\pm$  standard error of the mean (SEM), unless otherwise indicated. Flow cytometry data were tested for significance using a two-tailed unpaired *t* test for comparing two groups in GraphPad Prism version 8.0.2 software. *P*-values  $< 0.05$  were considered significant, with levels of significance being indicated as follows: \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ ; \*\*\*\*  $p < 0.0001$ ; *ns*, not significant. Number of animals per experiment is indicated in the figure legends. An overview of the total number of animals used in each experimental procedure is provided in Figure S14.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acsnano.9b08693>.

Figure S1: Nanoparticle characterization and pharmacokinetic profile. Figure S2: Nanoparticle biodistribution profile. Figure S3: Nanoparticle distribution in the tumor tissue over time. Figure S4: PET/CT of liposome injected animals and histological analysis of 66c14 tumors. Figure S5: *Ex vivo* CLSM on white blood cells isolated at 5 or 10 min or 2 h post-injection. Figure S6: IVM images obtained in tumors growing in window chamber mice. Figure S7: Flow cytometry gating strategy on blood, tumor, and tumor endothelium. Figure S8: Flow cytometry histograms reveal nanoparticle uptake by circulating and tumor associated

immune cells. Figure S9: Flow cytometry-based bar chart shows nanoparticle uptake by circulating and tumor associated immune cells. Figure S10: 66c14 tumor cell profiling. Figure S11: Association of circulating immune cells with  $\alpha_v$  and  $\beta_3$  integrin subunits. Figure S12: Nanoparticle uptake by tumor endothelial cells as a function of  $\alpha_v$  and  $\beta_3$  integrin expression levels. Figure S13: cRGD-NP positive immune cells do not extravasate in healthy ear tissue. Figure S14: Number of mice used per experimental procedure. Table S1: Nanoparticle composition. Table S2: Flow cytometry panels. Table S3: Used antibodies (PDF)

Movies S1–S13: Movies obtained in tumors grown in dorsal window chambers showing extensive cRGD-nanoparticle uptake in circulating immune cells and hitchhiking of cRGD-liposomes and cRGD-nanoemulsions (and not of cRAD-nanoparticles) with cells from the vasculature into the tumor tissue. Movie S1: Immune cell uptake in the circulation for cRGD-nanoemulsions during the first hour post-injection (AVI)

Movie S2: Immune cell uptake in the circulation for cRGD-liposomes during the first hour post-injection (AVI)

Movie S3: Intravital CD45-staining confirmed the presence of bound and circulating cRGD-nanoparticle positive immune cells (AVI)

Movie S4: cRAD-nanoemulsions were hardly taken up by immune cells during the first hour post-injection (AVI)

Movie S5: cRAD-liposomes were hardly taken up by immune cells during the first hour post-injection (AVI)

Movie S6: Significant immune cell uptake of circulating cRGD-nanoemulsions at later time points post-injection (AVI)

Movie S7: Significant immune cell uptake of circulating cRGD-liposomes at later time points post-injection (AVI)

Movie S8: cRAD-nanoemulsions were taken-up by circulating cells, but to much lower extent (AVI)

Movie S9: cRAD-liposomes were taken-up by circulating cells, but to much lower extent (AVI)

Movie S10: cRAD- and cRGD-nanoemulsions co-injections corroborated that cRGD-nanoemulsions were taken up to higher extents than cRAD-nanoemulsions by circulating cells (AVI)

Movie S11: cRAD- and cRGD-liposomes co-injections corroborated that cRGD-liposomes were taken up to higher extents than cRAD-liposomes by circulating cells (AVI)

Movie S12: In tumors grown in window chamber mice we observed cRGD-nanoemulsions hitchhiking with immune cells from the circulation into the tumor tissue (AVI)

Movie S13: In tumors grown in window chamber mice we observed cRGD-liposomes hitchhiking with immune cells from the circulation into the tumor tissue (AVI)

Movies 14–17: Movies obtained in tumors grown orthotopically showing that cRGD-nanomaterial hitchhikes with cells into the tumor tissue where these cells display high motility. Hardly any accumulation of cRGD-nanoparticles in TAMs was observed. Movie



S14: cRGD-nanoparticle hitchhiking (nanoemulsion) (AVI)

Movie S15: cRGD-nanoparticles hardly accumulated in dextran positive TAMs up to 3 hours post-injection, but rather presented in mobile cells in the tumor tissue (liposomes) (AVI)

Movie S16: cRGD-nanoparticles hardly accumulated in dextran positive TAMs up to 3 hours post-injection, but rather presented in mobile cells in the tumor adjacent fat tissue (liposomes) (AVI)

Movie S17: Hardly any cRAD-nanoparticle hitchhiking was observed (nanoemulsion) (AVI)

Movies 18–19: Movies obtained in healthy ear tissue showing the same nanoparticle uptake patterns in healthy mice as in the tumor bearing animals and the absence of nanoparticle hitchhiking with cells into healthy tissue. Movie S18: cRGD-liposomes (24 h post-injection) injected in non-tumor bearing mice showed an extensive uptake by circulating immune cells in comparison to limited uptake of the control cRAD-liposomes by these cells (AVI)

Movie S19: cRGD-nanoemulsions (2 h post-injection) injected in non-tumor bearing mice showed an extensive uptake by circulating immune cells in comparison to limited uptake of the control cRAD-nanoemulsions by these cells (AVI)

## AUTHOR INFORMATION

### Corresponding Authors

**Alexandros Marios Sofias** – Department of Circulation and Medical Imaging, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU), 7030 Trondheim, Norway; BioMedical Engineering and Imaging Institute, Icahn School of Medicine at Mount Sinai, New York, New York 10029, United States; Department of Nanomedicine and Theranostics, Institute for Experimental Molecular Imaging, Faculty of Medicine, RWTH Aachen University, S2074 Aachen, Germany; [orcid.org/0000-0002-7475-2526](https://orcid.org/0000-0002-7475-2526); Email: [alexandros Sofias@outlook.com](mailto:alexandros Sofias@outlook.com), [asofias@ukaachen.de](mailto:asofias@ukaachen.de)

**Sjoerd Hak** – Department of Circulation and Medical Imaging, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU), 7030 Trondheim, Norway; Department of Biotechnology and Nanomedicine, SINTEF Industry, 7034 Trondheim, Norway; Email: [sjoerd.hak@ntnu.no](mailto:sjoerd.hak@ntnu.no), [sjoerd.hak@sintef.no](mailto:sjoerd.hak@sintef.no)

### Authors

**Yohana C. Toner** – BioMedical Engineering and Imaging Institute, Icahn School of Medicine at Mount Sinai, New York, New York 10029, United States

**Anu E. Meerwaldt** – BioMedical Engineering and Imaging Institute, Icahn School of Medicine at Mount Sinai, New York, New York 10029, United States; Biomedical MR Imaging and Spectroscopy Group, Center for Image Sciences, University Medical Center Utrecht and Utrecht University, 3584 CX Utrecht, The Netherlands

**Mandy M. T. van Leent** – BioMedical Engineering and Imaging Institute, Icahn School of Medicine at Mount Sinai, New York, New York 10029, United States; Department of Medical Biochemistry, Amsterdam University Medical Centers, 1105 AZ Amsterdam, The Netherlands

**Georgios Soultanidis** – BioMedical Engineering and Imaging Institute, Icahn School of Medicine at Mount Sinai, New York, New York 10029, United States

**Mattijs Elschot** – Department of Circulation and Medical Imaging, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU), 7030 Trondheim, Norway; Department of Radiology and Nuclear Medicine, St. Olavs Hospital, Trondheim University Hospital, 7030 Trondheim, Norway

**Haruki Gonai** – BioMedical Engineering and Imaging Institute, Icahn School of Medicine at Mount Sinai, New York, New York 10029, United States

**Kristin Grendstad** – Department of Physics, Faculty of Natural Sciences, Norwegian University of Science and Technology (NTNU), 7034 Trondheim, Norway

**Åsmund Flobak** – The Cancer Clinic, St. Olav's University Hospital, 7030 Trondheim, Norway; Department of Clinical and Molecular Medicine, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU), 7030 Trondheim, Norway

**Ulrike Neckmann** – Department of Biomedical Laboratory Science, Faculty of Natural Sciences and Centre of Molecular Inflammation Research (CEMIR), Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU), 7030 Trondheim, Norway

**Camilla Wolowczyk** – Department of Biomedical Laboratory Science, Faculty of Natural Sciences and Centre of Molecular Inflammation Research (CEMIR), Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU), 7030 Trondheim, Norway

**Elizabeth L. Fisher** – BioMedical Engineering and Imaging Institute, Icahn School of Medicine at Mount Sinai, New York, New York 10029, United States

**Thomas Reiner** – Department of Radiology, Memorial Sloan Kettering Cancer Center, New York, New York 10065, United States; Department of Radiology, Weill Cornell Medical College, New York, New York 10065, United States;

[orcid.org/0000-0002-7819-5480](https://orcid.org/0000-0002-7819-5480)

**Catharina de Lange Davies** – Department of Physics, Faculty of Natural Sciences, Norwegian University of Science and Technology (NTNU), 7034 Trondheim, Norway

**Geir Bjørkøy** – Department of Clinical and Molecular Medicine, Faculty of Medicine and Health Sciences, Department of Biomedical Laboratory Science, Faculty of Natural Sciences, and Centre of Molecular Inflammation Research (CEMIR), Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU), 7030 Trondheim, Norway

**Abraham J. P. Teunissen** – BioMedical Engineering and Imaging Institute, Icahn School of Medicine at Mount Sinai, New York, New York 10029, United States; [orcid.org/0000-0002-0401-8262](https://orcid.org/0000-0002-0401-8262)

**Jordi Ochando** – Department of Oncological Sciences, Icahn School of Medicine at Mount Sinai, New York, New York 10029, United States; Transplant Immunology Unit, National Center of Microbiology, Instituto de Salud Carlos III, 28220 Madrid, Spain

**Carlos Pérez-Medina** – BioMedical Engineering and Imaging Institute, Icahn School of Medicine at Mount Sinai, New York, New York 10029, United States; Centro Nacional de Investigaciones Cardiovasculares Carlos III (CNIC), 28029 Madrid, Spain

**Willem J. M. Mulder** – BioMedical Engineering and Imaging Institute, Icahn School of Medicine at Mount Sinai, New York, New York 10029, United States; Department of Medical Biochemistry, Amsterdam University Medical Centers, 1105 AZ Amsterdam, The Netherlands; Laboratory of Chemical Biology, Department of Biochemical Engineering, Eindhoven University of Technology, 5612 AP Eindhoven, The Netherlands;  
✉ [orcid.org/0000-0001-8665-3878](https://orcid.org/0000-0001-8665-3878)

Complete contact information is available at:  
<https://pubs.acs.org/10.1021/acsnano.9b08693>

### Author Contributions

S.H. and W.J.M.M. conceptualized the study. A.M.S., C.P.M., W.J.M.M., and S.H. designed the study. A.M.S. and S.H. planned the experiments and A.M.S. coordinated their execution. Nanoparticles were synthesized and characterized by A.M.S. and S.H. and radiolabeled by A.J.P.T., T.R., and C.P.M. Gamma counting experiments were conducted by A.M.S., Y.C.T., A.E.M., and M.M.T.v.L. PET/CT imaging and analysis was conducted by A.M.S., G.S., M.E., H.G., and E.L.F. IVM and *ex vivo* microscopy was conducted by A.M.S. and S.H. Dorsal window chambers were implanted by K.G. Flow cytometry experiments and data analysis were conducted by A.M.S. Flow cytometry protocols were developed by A.M.S., U.N., C.W., and S.H. Immunological insights were provided by G.B. and J.O. Histology experiments and analysis were conducted by A.M.S. and Y.C.T. S.H. and A.M.S. wrote the manuscript and produced the figures. All authors contributed to data interpretation and reviewed, edited, and approved the manuscript before submission.

### Notes

The authors declare the following competing financial interest(s): J.O. and W.J.M.M. are scientific founders of Trained Therapeutix Discovery.

### ACKNOWLEDGMENTS

This work was supported by the Central Norway Regional Health Authority “Helse Midt-Norge” [A.M.S.: Ph.D. stipend (90062100) and travel grant (90284100); S.H.: researcher grant (90262100)], the National Institutes of Health (W.J.M.M.: R01 CA220234, T.R.: P30 CA00574), the American Heart Association (C.P.M.: 16SDG31390007), the Norwegian Research Council (S.H.: 230788/F20), and the Tromsø Research Foundation and Trond Mohn Foundation (S.H.: 180°N project).

### REFERENCES

- (1) Barenholz, Y. Doxil® — The First FDA-Approved Nano-Drug: Lessons Learned. *J. Controlled Release* **2012**, *160*, 117–134.
- (2) van der Meel, R.; Vehmeijer, L. J. C.; Kok, R. J.; Storm, G.; van Gaal, E. V. B. Ligand-Targeted Particulate Nanomedicines Undergoing Clinical Evaluation: Current Status. *Adv. Drug Delivery Rev.* **2013**, *65*, 1284–1298.
- (3) Witzigmann, D.; Hak, S.; van der Meel, R. Translating Nanomedicines: Thinking beyond Materials? A Young Investigator's Reply to ‘The Novelty Bubble’. *J. Controlled Release* **2018**, *290*, 138–140.
- (4) Hua, S.; de Matos, M. B. C.; Metselaar, J. M.; Storm, G. Current Trends and Challenges in the Clinical Translation of Nanoparticulate Nanomedicines: Pathways for Translational Development and Commercialization. *Front. Pharmacol.* **2018**, *9*, 790.
- (5) Chu, D.; Zhao, Q.; Yu, J.; Zhang, F.; Zhang, H.; Wang, Z. Nanoparticle Targeting of Neutrophils for Improved Cancer Immunotherapy. *Adv. Healthcare Mater.* **2016**, *5*, 1088–1093.

- (6) Miller, M. A.; Zheng, Y.-R.; Gadde, S.; Pfirschke, C.; Zope, H.; Engblom, C.; Kohler, R. H.; Iwamoto, Y.; Yang, K. S.; Askevold, B.; Kolishetti, N.; Pittet, M.; Lippard, S. J.; Farokhzad, O. C.; Weissleder, R. Tumour-Associated Macrophages Act as a Slow-Release Reservoir of Nano-Therapeutic Pt(IV) Pro-Drug. *Nat. Commun.* **2015**, *6*, 8692.
- (7) Tang, J.; Baxter, S.; Menon, A.; Alaarg, A.; Sanchez-Gaytan, B. L.; Fay, F.; Zhao, Y.; Ouimet, M.; Braza, M. S.; Longo, V. A.; Abdel-Atti, D.; Duivenvoorden, R.; Calcagno, C.; Storm, G.; Tsimikas, S.; Moore, K. J.; Swirski, F. K.; Nahrendorf, M.; Fisher, E. A.; Pérez-Medina, C.; et al. Immune Cell Screening of a Nanoparticle Library Improves Atherosclerosis Therapy. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, E6731–E6740.
- (8) Duro-Castano, A.; Gallon, E.; Decker, C.; Vicent, M. J. Modulating Angiogenesis with Integrin-Targeted Nanomedicines. *Adv. Drug Delivery Rev.* **2017**, *119*, 101–119.
- (9) Wu, P.-H.; Onodera, Y.; Ichikawa, Y.; Rankin, E.; Giaccia, A.; Watanabe, Y.; Qian, W.; Hashimoto, T.; Shirato, H.; Nam, J.-M. Targeting Integrins with RGD-Conjugated Gold Nanoparticles in Radiotherapy Decreases the Invasive Activity of Breast Cancer Cells. *Int. J. Nanomed.* **2017**, *12*, 5069–5085.
- (10) Ng, Q. K. T.; Olariu, C. L.; Yaffee, M.; Taelman, V. F.; Marincek, N.; Krause, T.; Meier, L.; Walter, M. A. Indium-111 Labeled Gold Nanoparticles for *In Vivo* Molecular Targeting. *Biomaterials* **2014**, *35*, 7050–7057.
- (11) Zhan, C.; Gu, B.; Xie, C.; Li, J.; Liu, Y.; Lu, W. Cyclic RGD Conjugated Poly(Ethylene Glycol)-Co-Poly(Lactic Acid) Micelles Enhance Paclitaxel Anti-Glioblastoma Effect. *J. Controlled Release* **2010**, *143*, 136–142.
- (12) Huang, Y.; Liu, W.; Gao, F.; Fang, X.; Chen, Y. C(RGDyK)-Decorated Pluronic Micelles for Enhanced Doxorubicin and Paclitaxel Delivery to Brain Glioma. *Int. J. Nanomed.* **2016**, *11*, 1629.
- (13) Schifferers, R. M.; Koning, G. A.; ten Hagen, T. L.; Fens, M. H. A.; Schraa, A. J.; Janssen, A. P. C.; Kok, R. J.; Molema, G.; Storm, G. Anti-Tumor Efficacy of Tumor Vasculature-Targeted Liposomal Doxorubicin. *J. Controlled Release* **2003**, *91*, 115–122.
- (14) Miura, Y.; Takenaka, T.; Toh, K.; Wu, S.; Nishihara, H.; Kano, M. R.; Ino, Y.; Nomoto, T.; Matsumoto, Y.; Koyama, H.; Cabral, H.; Nishiyama, N.; Kataoka, K. Cyclic RGD-Linked Polymeric Micelles for Targeted Delivery of Platinum Anticancer Drugs to Glioblastoma through the Blood-Brain Tumor Barrier. *ACS Nano* **2013**, *7*, 8583–8592.
- (15) Anselmo, A. C.; Mitragotri, S. Nanoparticles in the Clinic: An Update. *Bioeng. Transl. Med.* **2019**, *4*, e10143.
- (16) Sofias, A. M.; Dunne, M.; Storm, G.; Allen, C. The Battle of “Nano” Paclitaxel. *Adv. Drug Delivery Rev.* **2017**, *122*, 20–30.
- (17) Patti, B. S.; Chupin, V. V.; Torchilin, V. P. New Developments in Liposomal Drug Delivery. *Chem. Rev.* **2015**, *115*, 10938–10966.
- (18) Hörmann, K.; Zimmer, A. Drug Delivery and Drug Targeting with Parenteral Lipid Nanoemulsions — A Review. *J. Controlled Release* **2016**, *223*, 85–98.
- (19) Pérez-Medina, C.; Abdel-Atti, D.; Tang, J.; Zhao, Y.; Fayad, Z. A.; Lewis, J. S.; Mulder, W. J. M.; Reiner, T. Nanoreporter PET Predicts the Efficacy of Anti-Cancer Nanotherapy. *Nat. Commun.* **2016**, *7*, 11838.
- (20) Danhier, F.; Le Breton, A.; Pr at, V. RGD-Based Strategies To Target Alpha(v) Beta(3) Integrin in Cancer Therapy and Diagnosis. *Mol. Pharmacol.* **2012**, *9*, 2961–2973.
- (21) Schifferers, R. M.; Ansari, A.; Xu, J.; Zhou, Q.; Tang, Q.; Storm, G.; Molema, G.; Lu, P. Y.; Scaria, P. V.; Woodle, M. C. Cancer siRNA Therapy by Tumor Selective Delivery with Ligand-Targeted Sterically Stabilized Nanoparticle. *Nucleic Acids Res.* **2004**, *32*, No. e149.
- (22) Fan, Z.; Chang, Y.; Cui, C.; Sun, L.; Wang, D. H.; Pan, Z.; Zhang, M. Near Infrared Fluorescent Peptide Nanoparticles for Enhancing Esophageal Cancer Therapeutic Efficacy. *Nat. Commun.* **2018**, *9*, 2605.
- (23) Yang, K.; Liu, Y.; Liu, Y.; Zhang, Q.; Kong, C.; Yi, C.; Zhou, Z.; Wang, Z.; Zhang, G.; Zhang, Y.; Khashab, N. M.; Chen, X.; Nie,



Z. Cooperative Assembly of Magneto-Nanovesicles with Tunable Wall Thickness and Permeability for MRI-Guided Drug Delivery. *J. Am. Chem. Soc.* **2018**, *140*, 4666–4677.

(24) Yu, G.; Yang, Z.; Fu, X.; Yung, B. C.; Yang, J.; Mao, Z.; Shao, L.; Hua, B.; Liu, Y.; Zhang, F.; Fan, Q.; Wang, S.; Jacobson, O.; Jin, A.; Gao, C.; Tang, X.; Huang, F.; Chen, X. Polyrotaxane-Based Supramolecular Theranostics. *Nat. Commun.* **2018**, *9*, 766.

(25) Kimball, K. J.; Preuss, M. A.; Barnes, M. N.; Wang, M.; Siegal, G. P.; Wan, W.; Kuo, H.; Saddekni, S.; Stockard, C. R.; Grizzle, W. E.; Harris, R. D.; Aurigemma, R.; Curiel, D. T.; Alvarez, R. D. A Phase I Study of a Tropism-Modified Conditionally Replicative Adenovirus for Recurrent Malignant Gynecologic Diseases. *Clin. Cancer Res.* **2010**, *16*, 5277–5287.

(26) Pesonen, S.; Diaconu, I.; Cerullo, V.; Escutenaire, S.; Raki, M.; Kangasniemi, L.; Nokisalmi, P.; Dotti, G.; Guse, K.; Laasonen, L.; Partanen, K.; Karli, E.; Haavisto, E.; Oksanen, M.; Karioja-Kallio, A.; Hannuksela, P.; Holm, S.-L.; Kauppinen, S.; Joensuu, T.; Kanerva, A.; et al. Integrin Targeted Oncolytic Adenoviruses Ad5-D24-RGD and Ad5-RGD-D24-GMCSF for Treatment of Patients with Advanced Chemotherapy Refractory Solid Tumors. *Int. J. Cancer* **2012**, *130*, 1937–1947.

(27) Phillips, E.; Penate-Medina, O.; Zanzonico, P. B.; Carvajal, R. D.; Mohan, P.; Ye, Y.; Humm, J.; Gönen, M.; Kalaigian, H.; Schöder, H.; Strauss, H. W.; Larson, S. M.; Wiesner, U.; Bradbury, M. S. Clinical Translation of an Ultrasmall Inorganic Optical-PET Imaging Nanoparticle Probe. *Sci. Transl. Med.* **2014**, *6*, 260ra149.

(28) Dubey, P. K.; Mishra, V.; Jain, S.; Mahor, S.; Vyas, S. P. Liposomes Modified with Cyclic RGD Peptide for Tumor Targeting. *J. Drug Target.* **2004**, *12*, 257–264.

(29) Mulder, W. J. M.; van der Schaft, D. W. J.; Hautvast, P. A. I.; Strijkers, G. J.; Koning, G. A.; Storm, G.; Mayo, K. H.; Griffioen, A. W.; Nicolay, K. Early *In Vivo* Assessment of Angiostatic Therapy Efficacy by Molecular MRI. *FASEB J.* **2007**, *21*, 378–383.

(30) Hak, S.; Cebulla, J.; Huuse, E. M.; Davies, C. d. L.; Mulder, W. J. M.; Larsson, H. B. W.; Haraldseth, O. Periodicity in Tumor Vasculature Targeting Kinetics of Ligand-Functionalized Nanoparticles Studied by Dynamic Contrast Enhanced Magnetic Resonance Imaging and Intravital Microscopy. *Angiogenesis* **2014**, *17*, 93–107.

(31) Mulder, W. J. M.; Castermans, K.; van Beijnum, J. R.; oude Egbrink, M. G. A.; Chin, P. T. K.; Fayad, Z. A.; Löwik, C. W. G. M.; Kaijzel, E. L.; Que, I.; Storm, G.; Strijkers, G. J.; Griffioen, A. W.; Nicolay, K. Molecular Imaging of Tumor Angiogenesis Using Av $\beta$ 3-Integrin Targeted Multimodal Quantum Dots. *Angiogenesis* **2009**, *12*, 17–24.

(32) Sofias, A. M.; Åslund, A. K. O.; Hagen, N.; Grendstad, K.; Hak, S. Simple and Robust Intravital Microscopy Procedures in Hybrid TIE2GFP-BALB/c Transgenic Mice. *Mol. Imaging Biol.* **2019**, *1*–8.

(33) DeNardo, D. G.; Ruffell, B. Macrophages as Regulators of Tumour Immunity and Immunotherapy. *Nat. Rev. Immunol.* **2019**, *19*, 369–382.

(34) Coffelt, S. B.; Wellenstein, M. D.; de Visser, K. E. Neutrophils in Cancer: Neutral No More. *Nat. Rev. Cancer* **2016**, *16*, 431–446.

(35) Moses, K.; Brandau, S. Human Neutrophils: Their Role in Cancer and Relation to Myeloid-Derived Suppressor Cells. *Semin. Immunol.* **2016**, *28*, 187–196.

(36) Barros-Becker, F.; Lam, P.-Y.; Fisher, R.; Huttenlocher, A. Live Imaging Reveals Distinct Modes of Neutrophil and Macrophage Migration within Interstitial Tissues. *J. Cell Sci.* **2017**, *130*, 3801–3808.

(37) Sofias, A. M.; Andreassen, T.; Hak, S. Nanoparticle Ligand-Decoration Procedures Affect *In Vivo* Interactions with Immune Cells. *Mol. Pharmaceutics* **2018**, *15*, 5754–5761.

(38) Summers, C.; Rankin, S. M.; Condliffe, A. M.; Singh, N.; Peters, A. M.; Chilvers, E. R. Neutrophil Kinetics in Health and Disease. *Trends Immunol.* **2010**, *31*, 318–324.

(39) Colditz, I. G. Margination and Emigration of Leucocytes. *Pathol. Immunopathol. Res.* **1985**, *4*, 44–68.

(40) Hak, S.; Helgesen, E.; Hektoen, H. H.; Huuse, E. M.; Jarzyna, P. A.; Mulder, W. J. M.; Haraldseth, O.; Davies, C. de L. The Effect of Nanoparticle Polyethylene Glycol Surface Density on Ligand-Directed Tumor Targeting Studied *In Vivo* by Dual Modality Imaging. *ACS Nano* **2012**, *6*, 5648–5658.

(41) Smith, B. R.; Cheng, Z.; De, A.; Koh, A. L.; Sinclair, R.; Gambhir, S. S. Real-Time Intravital Imaging of RGD-Quantum Dot Binding to Luminal Endothelium in Mouse Tumor Neovasculature. *Nano Lett.* **2008**, *8*, 2599–2606.

(42) Smith, B. R.; Cheng, Z.; De, A.; Rosenberg, J.; Gambhir, S. S. Dynamic Visualization of RGD-Quantum Dot Binding to Tumor Neovasculature and Extravasation in Multiple Living Mouse Models Using Intravital Microscopy. *Small* **2010**, *6*, 2222–2229.

(43) Danhier, F.; Vroman, B.; Lecouturier, N.; Crockart, N.; Pourcelle, V.; Freichels, H.; Jérôme, C.; Marchand-Brynaert, J.; Feron, O.; Préat, V. Targeting of Tumor Endothelium by RGD-Grafted PLGA-Nanoparticles Loaded with Paclitaxel. *J. Controlled Release* **2009**, *140*, 166–173.

(44) Amin, M.; Mansourian, M.; Koning, G. A.; Badiie, A.; Jaafari, M. R.; ten Hagen, T. L. M. Development of a Novel Cyclic RGD Peptide for Multiple Targeting Approaches of Liposomes to Tumor Region. *J. Controlled Release* **2015**, *220*, 308–315.

(45) Hou, J.; Yang, X.; Li, S.; Cheng, Z.; Wang, Y.; Zhao, J.; Zhang, C.; Li, Y.; Luo, M.; Ren, H.; Liang, J.; Wang, J.; Wang, J.; Qin, J. Accessing Neuroinflammation Sites: Monocyte/Neutrophil-Mediated Drug Delivery for Cerebral Ischemia. *Sci. Adv.* **2019**, *5*, No. eaau8301.

(46) Jain, S.; Mishra, V.; Singh, P.; Dubey, P.; Saraf, D.; Vyas, S. RGD-Anchored Magnetic Liposomes for Monocytes/Neutrophils-Mediated Brain Targeting. *Int. J. Pharm.* **2003**, *261*, 43–55.

(47) Kunjachan, S.; Pola, R.; Gremse, F.; Theek, B.; Ehling, J.; Moeckel, D.; Hermanns-Sachweh, B.; Pechar, M.; Ulbrich, K.; Hennink, W. E.; Storm, G.; Lederle, W.; Kiessling, F.; Lammers, T. Passive versus Active Tumor Targeting Using RGD- and NGR-Modified Polymeric Nanomedicines. *Nano Lett.* **2014**, *14*, 972–981.

(48) Kirpotin, D. B.; Drummond, D. C.; Shao, Y.; Shalaby, M. R.; Hong, K.; Nielsen, U. B.; Marks, J. D.; Benz, C. C.; Park, J. W. Antibody Targeting of Long-Circulating Lipidic Nanoparticles Does Not Increase Tumor Localization but Does Increase Internalization in Animal Models. *Cancer Res.* **2006**, *66*, 6732–6740.

(49) Dai, Q.; Wilhelm, S.; Ding, D.; Syed, A. M.; Sindhvani, S.; Zhang, Y.; Chen, Y. Y.; MacMillan, P.; Chan, W. C. W. Quantifying the Ligand-Coated Nanoparticle Delivery to Cancer Cells in Solid Tumors. *ACS Nano* **2018**, *12*, 8423–8435.

(50) Panneerselvam, S.; Choi, S. Nanoinformatics: Emerging Databases and Available Tools. *Int. J. Mol. Sci.* **2014**, *15*, 7158–7182.

(51) Liarski, V. M.; Sibley, A.; van Panhuys, N.; Ai, J.; Chang, A.; Kennedy, D.; Merolle, M.; Germain, R. N.; Giger, M. L.; Clark, M. R. Quantifying *In Situ* Adaptive Immune Cell Cognate Interactions in Humans. *Nat. Immunol.* **2019**, *20*, 503–513.

(52) Giger, M. L. Machine Learning in Medical Imaging. *J. Am. Coll. Radiol.* **2018**, *15*, 512–520.

(53) Kim, H. J.; Lee, S.-M.; Park, K.-H.; Mun, C. H.; Park, Y.-B.; Yoo, K.-H. Drug-Loaded Gold/Iron/Gold Plasmonic Nanoparticles for Magnetic Targeted Chemo-Photothermal Treatment of Rheumatoid Arthritis. *Biomaterials* **2015**, *61*, 95–102.

(54) Koning, G. A.; Schiffelers, R. M.; Wauben, M. H. M.; Kok, R. J.; Mastrobattista, E.; Molema, G.; ten Hagen, T. L. M.; Storm, G. Targeting of Angiogenic Endothelial Cells at Sites of Inflammation by Dexamethasone Phosphate-Containing RGD Peptide Liposomes Inhibits Experimental Arthritis. *Arthritis Rheum.* **2006**, *54*, 1198–1208.

(55) Rodell, C. B.; Arlauckas, S. P.; Cuccarese, M. F.; Garris, C. S.; Li, R.; Ahmed, M. S.; Kohler, R. H.; Pittet, M. J.; Weissleder, R. TLR7/8-Agonist-Loaded Nanoparticles Promote the Polarization of Tumour-Associated Macrophages to Enhance Cancer Immunotherapy. *Nat. Biomed. Eng.* **2018**, *2*, 578–588.

(56) Dong, R.; Gong, Y.; Meng, W.; Yuan, M.; Zhu, H.; Ying, M.; He, Q.; Cao, J.; Yang, B. The Involvement of M2 Macrophage

Polarization Inhibition in Fenretinide-Mediated Chemopreventive Effects on Colon Cancer. *Cancer Lett.* **2017**, *388*, 43–53.

(57) Sipkins, D. A.; Cheresch, D. A.; Kazemi, M. R.; Nevin, L. M.; Bednarski, M. D.; Li, K. C. P. Detection of Tumor Angiogenesis *In Vivo* by Av $\beta$ 3-Targeted Magnetic Resonance Imaging. *Nat. Med.* **1998**, *4*, 623–626.

(58) Melemenidis, S.; Jefferson, A.; Ruparelia, N.; Akhtar, A. M.; Xie, J.; Allen, D.; Hamilton, A.; Larkin, J. R.; Perez-Balderas, F.; Smart, S. C.; Muschel, R. J.; Chen, X.; Sibson, N. R.; Choudhury, R. P. Molecular Magnetic Resonance Imaging of Angiogenesis *In Vivo* Using Polyvalent Cyclic RGD-Iron Oxide Microparticle Conjugates. *Theranostics* **2015**, *5*, 515–529.

(59) Kitagawa, T.; Kosuge, H.; Uchida, M.; Iida, Y.; Dalman, R. L.; Douglas, T.; McConnell, M. V. RGD Targeting of Human Ferritin Iron Oxide Nanoparticles Enhances *In Vivo* MRI of Vascular Inflammation and Angiogenesis in Experimental Carotid Disease and Abdominal Aortic Aneurysm. *J. Magn. Reson. Imaging* **2017**, *45*, 1144–1153.

(60) Huang, W.-C.; Chen, S.-H.; Chiang, W.-H.; Huang, C.-W.; Lo, C.-L.; Chern, C.-S.; Chiu, H.-C. Tumor Microenvironment-Responsive Nanoparticle Delivery of Chemotherapy for Enhanced Selective Cellular Uptake and Transportation within Tumor. *Biomacromolecules* **2016**, *17*, 3883–3892.

(61) Chu, D.; Gao, J.; Wang, Z. Neutrophil-Mediated Delivery of Therapeutic Nanoparticles across Blood Vessel Barrier for Treatment of Inflammation and Infection. *ACS Nano* **2015**, *9*, 11800–11811.

(62) Giovinazzo, H.; Kumar, P.; Sheikh, A.; Brooks, K. M.; Ivanovic, M.; Walsh, M.; Caron, W. P.; Kowalsky, R. J.; Song, G.; Whitlow, A.; Clarke-Pearson, D. L.; Brewster, W. R.; Van Le, L.; Zamboni, B. A.; Bae-Jump, V.; Gehrig, P. A.; Zamboni, W. C. Technetium Tc 99m Sulfur Colloid Phenotypic Probe for the Pharmacokinetics and Pharmacodynamics of PEGylated Liposomal Doxorubicin in Women with Ovarian Cancer. *Cancer Chemother. Pharmacol.* **2016**, *77*, 565–573.

(63) Song, G.; Petschauer, J.; Madden, A.; Zamboni, W. Nanoparticles and the Mononuclear Phagocyte System: Pharmacokinetics and Applications for Inflammatory Diseases. *Curr. Rheumatol. Rev.* **2014**, *10*, 22–34.

(64) van der Meel, R.; Sulheim, E.; Shi, Y.; Kiessling, F.; Mulder, W. J. M.; Lammers, T. Smart Cancer Nanomedicine. *Nat. Nanotechnol.* **2019**, *14*, 1007–1017.

(65) Pérez-Medina, C.; Abdel-Atti, D.; Zhang, Y.; Longo, V. A.; Irwin, C. P.; Binderup, T.; Ruiz-Cabello, J.; Fayad, Z. A.; Lewis, J. S.; Mulder, W. J. M.; Reiner, T. A Modular Labeling Strategy for *In Vivo* PET and Near-Infrared Fluorescence Imaging of Nanoparticle Tumor Targeting. *J. Nucl. Med.* **2014**, *55*, 1706–1711.

(66) Holland, J. P.; Sheh, Y.; Lewis, J. S. Standardized Methods for the Production of High Specific-Activity Zirconium-89. *Nucl. Med. Biol.* **2009**, *36*, 729–739.

(67) Seynhaeve, A. L. B.; ten Hagen, T. L. M. Intravital Microscopy of Tumor-Associated Vasculature Using Advanced Dorsal Skinfold Window Chambers on Transgenic Fluorescent Mice. *J. Visualized Exp.* **2018**, *131*, No. e55115.





# Defining the boundaries: challenges and advances in identifying cells in microscopy images

Nodar Gogoberidze and Beth A Cimini



Segmentation, or the outlining of objects within images, is a critical step in the measurement and analysis of cells within microscopy images. While improvements continue to be made in tools that rely on classical methods for segmentation, deep learning-based tools increasingly dominate advances in the technology. Specialist models such as Cellpose continue to improve in accuracy and user-friendliness, and segmentation challenges such as the Multi-Modality Cell Segmentation Challenge continue to push innovation in accuracy across widely varying test data as well as efficiency and usability. Increased attention on documentation, sharing, and evaluation standards is leading to increased user-friendliness and acceleration toward the goal of a truly universal method.

## Address

Imaging Platform, Broad Institute, Cambridge, MA 02142, USA

Corresponding author: Cimini, Beth A ([bcimini@broadinstitute.org](mailto:bcimini@broadinstitute.org))

**Current Opinion in Biotechnology** 2024, **85**:103055

This review comes from a themed issue on **Analytical Biotechnology**

Edited by **Andreas E. Vasdekis** and **Scott E. Baker**

For complete overview of the section, please refer to the article collection, "**Analytical Biotechnology (2024)**"

Available online 23 December 2023

<https://doi.org/10.1016/j.copbio.2023.103055>

0958-1669/© 2023 Elsevier Ltd. All rights reserved.

## Introduction

Segmentation plays a pivotal role in microscopy analysis and refers to the automatic delineation of individual objects (often cells or cellular components) within complex scientific images. It is an important step before measuring properties of those biological entities. Approaches for cell segmentation have benefitted from advancements in more general segmentation problems in traditional computer vision (CV), machine learning (ML), and in recent years deep learning (DL) [1,2]. Accurate segmentation allows the quantification and analysis of cellular features, such as morphology, staining intensity, and spatial relationships, which capture valuable cellular phenotypes. While computational methods now achieve better-than-human accuracy on a number of specific tasks, in general, given the wide range of cell

types, imaging modalities, and experimental conditions, the problem remains an ongoing challenge.

As state-of-the-art (SOTA) methodologies for segmentation have progressed, the community has also tried to provide access to these methods to less-computational users in the form of user-friendly software interfaces and intuitive tools that improve reproducibility. Widespread adoption will require methods with few-or-no tunable parameters, models that are efficient in terms of computational runtime and memory requirements, and an ecosystem of tools for their use. The past two years reviewed here have seen a proliferation of new local and cloud-oriented software and workflows, the adoption of several user-oriented models, and the development of next-generation model architectures. We herein review progress in approaches utilizing classical CV techniques and specialist DL networks, as well as progress toward and current needs related to making high-quality-accessible generalist networks that will reduce the 'time to science' for the broader community.

## Progress in classical approaches

While advancements in segmentation accuracy are largely driven via DL approaches, they are not always a suitable solution, as some require large annotated datasets and interpretability (though often unnecessary for segmentation tasks) can be a challenge [3,4]. We therefore begin with advancements in non-deep ML and classic image processing. Kartezio [5] is a recent exemplar of non-deep ML, using Cartesian Genetic Programming to combine classic CV algorithms into a fully interpretable image pipeline for segmentation. It performs comparably to DL approaches such as Cellpose [6], StarDist [7], and Mask R-CNN [8], and importantly requires relatively few images to train a computationally efficient and explainable workflow.

Existing CV tools such as Fiji/ImageJ [9], CellProfiler [10], and Napari [11] are receiving continuous extensions via the growing ecosystem of plugins and integrations [1,12-15], allowing these tools to adapt to a broader range of tasks. One example, General Image Analysis of Nuclei-based Images [16], is a Fiji plugin for segmentation of cells in 3D microscopy images. Similarly, LABKIT [17] is a Fiji plugin specifically oriented toward efficient segmentation in large, multiterabyte, images. Other tools such as Tonga [18] prioritize ease of

installation and customization to a specific task to appeal to nontechnical and nonexpert users.

### Lowering time-to-science in deep learning

As new DL models rapidly emerge, users who want to employ them must surmount computational hurdles. First, installing the models and their corresponding software dependencies can often be technically complex and time-consuming, often involving usage of the Command Line Interface (CLI), compiling source code, navigating environment-related conflicts, and other tasks that may confound nontechnical users. Second, pretrained models are most effective on the type of data they have been trained on, which may differ from a user's data in a number of ways such as the phenotype of interest, how the biological samples were prepared, the imaging modality (fluorescence, histological stains, phase contrast, etc.), and the experimental conditions. The model's *generality* is how well it is able to perform across these differences, without loss in segmentation accuracy. If the model's generality is not sufficient to perform well on the user's data, it must be fine-tuned by feeding in new training data (Figure 1). The complexity of this task can vary even more dramatically than installation. Depending on the tools made available by the model's developers, it may involve writing custom code, which can be a time-intensive task even for expert programmers, or may simply require the use of a purpose-built tool.

The less time spent installing, tuning, and configuring models and software, the more bandwidth is available to concentrate on addressing scientific questions. Unfortunately, usability varies wildly across tools and documentation for trained DL models, from only the model parameters (weights) without documentation or source code, to models that come with extensive documentation and entire libraries for utilizing the model, including data loading and processing, fine-tuning, and configuration, to models that come with several interfaces, including CLI or Graphical User Interfaces (GUI), easy-to-use installers, and guides or tutorials on how to fine-tune and configure the model. It is no accident that some of the most commonly used networks, discussed below, are those that in addition to high performance have emphasized usability.

### Current progress toward useful specialist networks

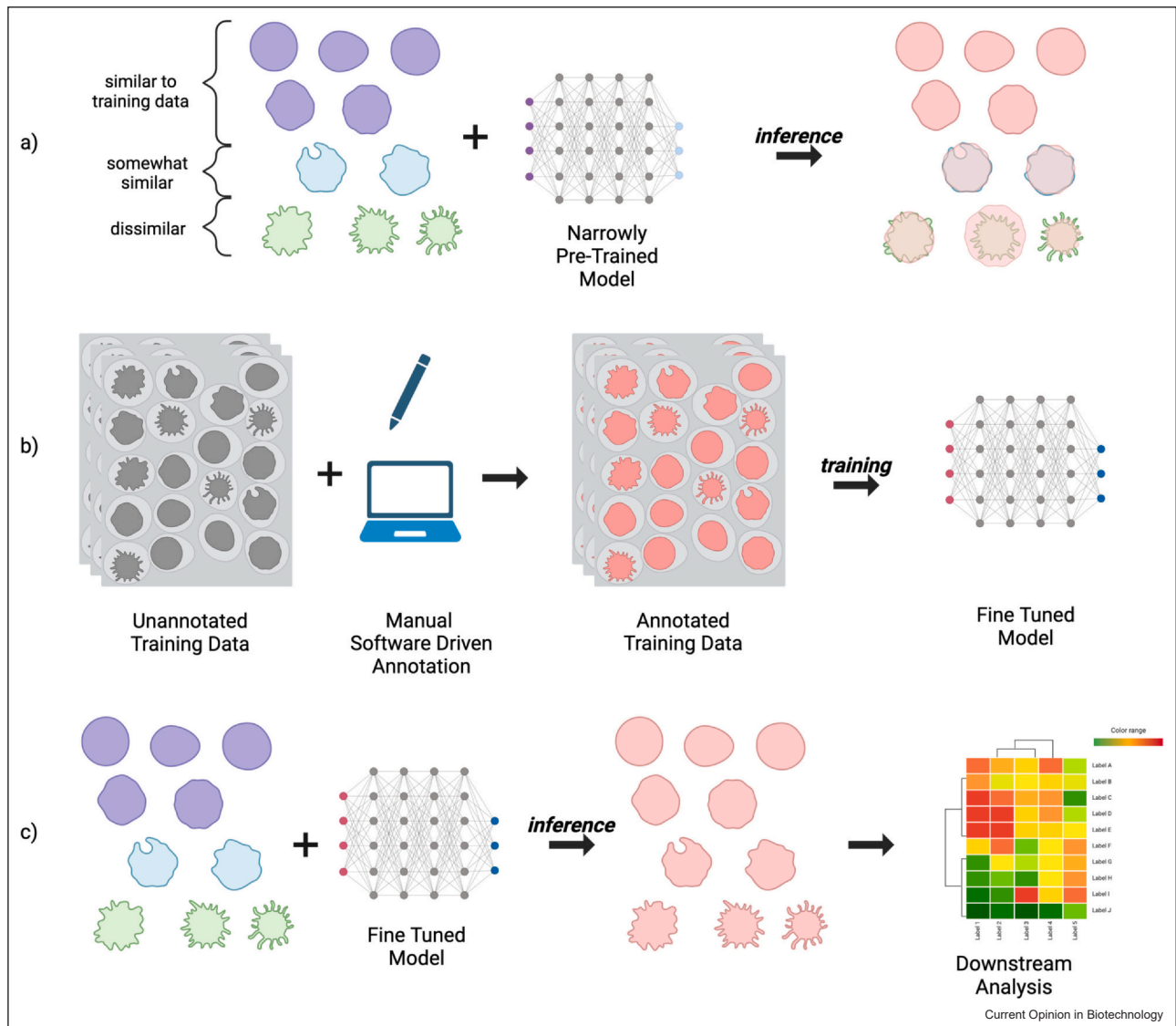
As shown in Figure 1, models typically underperform on new, unseen data. Before fine-tuning, these data are considered out of distribution (OOD). Intuitively, OOD data can be thought of as being drawn from a different distribution than that of the original training set (such as the purple versus green cells in Figure 1). A shift between the in-distribution training set and OOD data can be referred to as a difference in *style* between the two sets of data, for instance, differences in acquisition

parameters, staining methods, or imaging modalities [19,20]. A process known as *style transfer* can be utilized to address changes in the data distribution by training a model that is able to pixel-wise map an image of one style to that of another, ideally with minimal loss in semantic content [21]. For instance, a style transfer model can be trained to transform an image of one modality, such as brightfield, to that of another, such as fluorescence; or a model can be trained to transform an image's annotation mask into the image itself. The nucleAIzer [22] model utilizes the latter approach as a means of achieving greater generalization capabilities, allowing the model to more easily be adapted to OOD data. Although developed over three years ago, it is still one of the top-performing models at cell segmentation [23]. In order to improve usability, a plugin was developed for CellProfiler 3 [24] allowing users of the tool to perform inference through the use of a GUI. While the plugin allows ease of use via the CellProfiler interface, manual installation of nucleAIzer and its dependencies are still necessary, which is a challenge for noncomputational users. A simple web interface is also available, but requires upload of the data to a central server, limiting use for large batches of images.

StarDist predates nucleAIzer, however, iterations of it are still being introduced [25,26], and general usage remains quite high due to a great deal of time being spent on documenting its usage, as well as making it available across environments and in a variety of graphical tools [12] such as Fiji/ImageJ, Napari, QuPath [27,28], Icy [29], CellProfiler, and KNIME [30]. While a variety of pretrained variants are available, it is still limited to specific modalities (such as fluorescence or histology stains) and even with fine-tuning is oriented to segmenting objects that are *star-convex* — shapes where line segments can be drawn from any point along the border to some single interior point — that makes it a poor choice for very irregular cell shapes or neurons, since shapes with very large bends or curves may not be star-convex.

A primary goal of Cellpose was to develop a generalist model by training on a large dataset of manually segmented images from a variety of modalities. Its preprocessing method, focusing on transforming input data to spatial gradients, allows it to generalize to a larger variety of shapes. While the architecture was developed for generalist purposes, fine-tuning is still often necessary. Cellpose 2.0 [31] was introduced as a package that included several pretrained models, a human-in-the-loop pipeline for fine-tuning custom models with small datasets, and an improved set of graphical software tools to aid in its usage. Omnipose [32] extends Cellpose to work better on elongated cells common in bacteria by adding distance field prediction similar to StarDist. It is similarly well-documented, packaged, and provides a number of interfaces in the form of a library, CLI, and

Figure 1



Segmentation and fine-tuning process. **(a)** The inference dataset contains a variety of samples: some are similar to the original training data the model was pretrained on, while others are slightly or very different, leading to low segmentation accuracy. **(b)** New training data matching the characteristics of the full distribution are annotated either manually, through software such as CellProfiler, or a human-in-the-loop model such as Cellpose, and used to fine-tune the model. **(c)** The fine-tuned model produces more accurate segmentations on the inference dataset, which can then be used for downstream tasks. Figures created with [BioRender.com](https://BioRender.com).

GUI. While it allows for training, it is not available in Cellpose 2.0, and thus is not focused on human-in-the-loop fine-tuning. Cellpose and Cellpose 2.0 include a custom GUI, and extensive documentation; broad usage is further supported by their availability via plugins from many of the same tools as StarDist.

Mesmer is a DL pipeline trained on the largest public tissue dataset of annotated nuclei and whole cells, TissueNet [33]. It provides access to a remotely hosted instance model through several interfaces, including a

web portal and plugins for Fiji/ImageJ and QuPath. It also provides a Docker container (described below, in Making Models Findable, Accessible, Interoperable, and Reusable (FAIR)) to run the model in a self-hosted manner, with access provided through a Jupyter Notebook or CLI.

### A vision for the state-of-the-art

Many of those working on DL for image segmentation are aiming to create a truly generalist model, often referred to as a foundation model, capable of greater-than-human

accuracy, across the wide variety of imaging modalities, in a parameter-free manner [34–36]. For biologists, this would mean a model that has no need for manual annotation or fine-tuning, preferably in a form factor that is easily accessible, configurable, and invocable. Although there is still much progress to be made on this front, a number of architectural developments can be highlighted as mile markers along the path.

#### Dataset availability

The effectiveness of DL models relies on sufficient similarity between the training data and the user's own data. The specific criteria for what constitutes 'sufficient similarity' will naturally differ based on the methods and architectural choices employed by the model, but it therefore follows that creating foundation models (capable of segmenting a wide variety of biological images) will require a diverse and comprehensive corpus of training data to ensure that the segmentation model can generalize effectively. The dataset should encompass a broad spectrum of microscope modalities, across a variety of imaging conditions, and include many distinct cell types.

The volume of publicly available datasets is ever-increasing, particularly driven from the development of specialist models and challenges wherein teams compete on segmentation-oriented tasks. TissueNet is the largest collection of annotated tissue images, while LIVECell [37] is the largest collection of high-quality, manually annotated, and expert-validated phase-contrast images. The Cell Tracking Challenge (CTC) [38] is an ongoing benchmark and reference in cell segmentation and tracking algorithms, which in recent years has extended the available benchmarks with the Cell Segmentation Benchmark (CSB). The Multi-Modality Cell Segmentation Challenge (MMCSC) [36] consolidated a modestly sized labeled dataset with a particular emphasis on diversity in modalities.

Challenges stand as an excellent pointer for future progress in a given area; trends among the top-ranking team's architectures and techniques often form the basis of future implementations available to the wider field. For instance, data augmentation — where existing training data are perturbed and transformed in various ways, such as rotations, scaling, intensity adjustments, or the infusion of random noise — was highlighted in the MMCSC as a particularly important feature in pre-training top-performing models, aiding them in their generalizability. Entirely synthetic datasets are also often useful, as demonstrated by their inclusion in a subset of the CTC datasets, and in the development of frameworks for their generation [39].

#### Next-generation models

Many architectures are being researched and explored in the quest for evermore general and robust models, as

demonstrated in the MMCSC. The KIT-GE model [40] is among the top-3 performing models of CTC CSB, and was therefore used as one of the baselines in the MMCSC alongside Cellpose, Cellpose 2.0, and OmniPose [6,31,32]. Analysis of the top-ranking solutions in the MMCSC shows that choices in backbone networks are particularly important for next-generation models. While U-Net inspired architectures form the basis of the widely used contemporary models such as StarDist, Cellpose, and KIT-GE, the winning solutions in MMCSC employ backbones such as SegFormer [41], ConvNeXt [42], and ResNeXt [43]. CTC reports that segmentation performance increases with techniques such as self-configured neural networks (e.g. nnU-Net) [44], neural architecture search, and multibranch prediction.

The generalist capabilities provided by Cellpose 2.0 rely on fine-tuning, which may cause the model to suffer from severe loss of performance on tasks outside of the fine-tuning dataset [36]. The top-performing model [45] of the MMCSC was able to outperform the pretrained generalist Cellpose and Omnipose models, as well as a Cellpose 2.0 model fine-tuned on the challenge's training data. The testing set included images that were distinct from the training data, and sourced from new biological experiments, meaning successful models needed to show a strong ability to generalize across data without additional fine-tuning.

While the winning solution of MMCSC and its associated code is available on GitHub, it remains to be seen whether it or any of the top-ranking models will be available with documentation and interface tools in a way comparable to StarDist, Cellpose 2.0, or Omnipose. If so, we may be one step closer to a truly general, easy-to-set up, easy-to-use, one-click segmentation model, with no additional tuning. Short of that, alternative interfaces for model configuration may come to prominence in the form of dialog-driven large language models (LLMs) [46,47]. It is also as yet unclear whether future enhancements will be driven primarily by transformer architectures [34,48], whether advancements in convolutional networks [42] will keep pace, as demonstrated by the second- and third-place solutions in the MMCSC, or whether hybrid approaches will dominate [49,50].

It is also of great research and commercial interest to develop foundation models. Meta AI Research recently released a family of foundation models for segmentation, referred to as the Segment Anything Model [51] (SAM), the largest of which was trained on 1.1 billion high-quality segmentation labels, across 11 million high-resolution images. While the images included in the dataset were mostly photographs of natural scenes, it did include a small number of microscopy images taken from the 2018 Data Science Bowl [52]. In short order, Segment Anything for Microscopy [53] was developed, in which SAM



was extended to generalize across many imaging modalities by fine-tuning the original model using a variety of datasets. Important to SAM's architecture is its interactive segmentation capabilities, where a subset of the user's data is first annotated with a small amount of either point annotations or rectangular bounding boxes. Annotations of this type are significantly less time-consuming than pixel-level mask annotations, and provide SAM with enough guidance to output full segmentation masks on the user's dataset. Segment Anything for Microscopy therefore adopts this capability and includes a Napari plugin for interactive and automatic segmentation. There is a mechanism for automatic segmentation, however, in order to get generalist accuracy above that of Cellpose, some manual annotations must be made.

### **Making models Findable, Accessible, Interoperable, and Reusable**

Alongside progress in model development, there has been a greater push toward the dissemination of models such that they are FAIR [54–56]. Not only should models be available on publicly accessible platforms, but the associated code for asset loading, data preprocessing, data postprocessing, and model training and model inference, should also be made available in a well-packaged and documented form. Container platforms such as Docker [57] can alleviate many installation and setup complexities, providing an isolated and controlled environment in which software is installed, and a pre-configured installation process for the target software. This 'containerization' of software and its dependencies dramatically decreases the barriers for reuse. In addition, making models readily accessible, configurable, and tunable in a low- or no-code manner via interactive code notebooks or GUI encourages broader adoption of SOTA models.

The CTC proposed guidelines for algorithm developers to make their workflows both available and reproducible, while currently optional, they will be mandatory in the future. At minimum, the source code should be available on a public repository, and contain clear instructions for installing dependencies, initializing the model, loading weights, and training with new data. They also pushed for source code to be available via notebooks such as Jupyter and Google Colab. In the same vein, the MMCSC required all participants to place their solutions in Docker containers; the winning teams have made these available on public image registries and also made their algorithms publicly available on GitHub alongside processing source code. In addition, the top-three solutions were encouraged to develop Napari [11] plugins.

A missing component in full adherence to FAIR principles is the interoperability of models. While research and development in model architectures are healthy and

vital, there is no agreed-upon specification on the inputs and outputs of models. The difference in the model outputs between, for example, StarDist and Cellpose, is stark, and the postprocessing that is needed is correspondingly distinct. While the outputs are necessary by-products of the model architectures, the lack of a standard makes interoperability with existing tools difficult as custom code needs to be written to mirror the post-processing steps.

### **Making models efficient**

Models vary widely in terms of algorithmic efficiency, which will affect their adoption, especially in low-resource settings. While some challenges such as the CTC emphasize segmentation accuracy alone, others such as MMCSC evaluate efficiency as an explicit criterion, and the top-performing models had good trade-offs between accuracy and efficiency in runtime and memory usage. There are additional efforts in bringing model optimization tools to the bioimage community [58,59], as well as reducing runtime of bioimage analysis pipelines in general [59,60]. Efficient models, and model optimization tools will become increasingly important for training and inference tasks in local desktop and web-based tools [2,61], especially in contexts where moving data to the cloud is not viable or allowed.

### **Improving tool access and availability**

Though local-first software and algorithms remain important to grow and maintain, cloud-oriented tools and resources for bioimage analysis are becoming more prevalent and easier to use as demand for large-scale cloud-based workflows increases. In addition to providing large storage capacity and high-performance hardware, cloud-based tools increase the availability and accessibility of models, and often move the technical complexity away from the end user.

Notebooks allow code, explanatory text, and interactive elements to live together in a single package, providing alternatives or complements to libraries, documentation, and GUI interfaces. ZeroCostDL4Mic [62] provides rigorously documented and annotated code notebooks with prewritten code that can be customized for specific workflows through the exposed settings. A major benefit to these notebooks is that they can be deployed either locally or on the Google Colab platform, which eases hardware requirements and allows running moderately sized workloads for free. Behind the scenes, a container is initialized in the cloud and the installation occurs via preconfigured installation scripts contained within the notebooks.

Beyond notebooks, other tools provide a larger degree of customization and control, albeit at the expense of additional complexity. The BioContainers project is an

open-source and community-driven framework that provides cloud resources for defining, building, and distributing containers for biological tools [63]. The BioContainers Registry was developed with FAIR principles in mind, and provides both web and RESTful API interfaces to search for bioinformatics tools [64]. BIAFLOWS [65] is a community-driven, open-source web platform that allows deployment of and access to a wide variety of reproducible image analysis workflows. The platform provides a framework to import data, encapsulate workflows in container images, batch-process data, visualize data, and assess performance using widely accepted benchmark metrics on a large collection of public datasets. BioImageIT [66] is a more recent, plugin-oriented, workflow tool for data management and analysis. It has a unique emphasis on reconciling existing data management and data analysis tools, and although run locally has the ability to tap into remote data stores and job runners.

The BioImage Model Zoo [67] provides a community-driven repository for pretrained DL models and promotes a standard model description format for describing metadata. Community partners can work with the BioImage Model Zoo to support execution of the models and include many common bioimage tools. In addition, model execution can be performed via the BioEngine application framework, on top of the ImJoy plugin framework [68], allowing inference both on the BioImage Model Zoo web application and other web applications using the client ImJoy software. Behind the scenes, multiple containers are being run and managed with a container orchestration tool called Kubernetes.

For moderately more technical users, who are comfortable with using tools for deploying their own container orchestration workloads, there are some additional options. DeepCell Kiosk [69] is a cloud-native tool for dynamic scaling of image analysis workflows, utilizing Kubernetes orchestration similarly to the BioEngine inference engine. The tool is managed from several interfaces, including a web portal and Fiji plugin. Distributed-Something [70] takes a script-based approach to scale and distribute arbitrary containerized jobs on AWS, automatically configuring the AWS infrastructure for container orchestration, monitoring, and data handling. It runs the work in a cost-effective manner, and cleans up the infrastructure when the work has been completed.

## Conclusion

The landscape of segmentation algorithms, enabling tools, workflow management systems, repositories, benchmarks, and challenges, is constantly shifting. This very active landscape makes it all the more important to create community standards for reporting on methods and robust segmentation quality metrics, on which there

has been recent guidance [35,71–73]. While there is still much work to do, the past two years have seen essential strides made in democratizing the use of advanced segmentation methods through user-friendly interfaces and improved documentation. Integrating tools and scaling up reproducible workflows fosters a more collaborative and robust ecosystem; these continuing trends will empower researchers from diverse backgrounds to collectively explore the intricate universe of single-cell biology, ultimately accelerating the pace of discovery and innovation in this vital field of study.

## Funding

This publication has been made possible in part by Chan Zuckerberg Initiative Grants 2020-225720 (DOI: 10.37921/977328pjbca) and 2021-238657 (DOI: 10.37921/365498zdfyyk) from the Chan Zuckerberg Initiative DAF, an advised fund of Silicon Valley Community Foundation (funder DOI 10.13039/100014989). This work was also supported by the Center for Open Bioimage Analysis (COBA) funded by the National Institute of General Medical Sciences P41 GM135019. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the paper.

## CRedit authorship contribution statement

**Nodar Gogoberidze:** Conceptualization, Visualization, Writing – original draft, Writing – review & editing. **Beth A Cimini:** Conceptualization, Visualization, Writing – review & editing, Funding acquisition.

## Data Availability

No data were used for the research described in the article.

## Declaration of Competing Interest

The authors declare that there are no competing interests associated with the paper.

## Acknowledgements

The authors thank the members of the Imaging Platform for helpful discussions and contributions.

Figures created with BioRender.com

## References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
  - of outstanding interest
1. Hollandi R, Moshkov N, Paavolainen L, Tasnadi E, Piccinini F, Horvath P: **Nucleus segmentation: towards automated solutions.** *Trends Cell Biol* 2022, **32**:295-310, <https://doi.org/10.1016/j.tcb.2021.12.004>.

A comprehensive review, up to 2021, on the state of nucleus segmentation.

2. Lucas AM, Ryder PV, Li B, Cimini BA, Eliceiri KW, Carpenter AE: **Open-source deep-learning software for bioimage segmentation.** *Mol Biol Cell* 2021, **32**:823-829, <https://doi.org/10.1091/mbc.E20-10-0660>
  3. Rudin C: **Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead.** *Nat Mach Intell* 2019, **1**:206-215, <https://doi.org/10.1038/s42256-019-0048-x>
  4. Karim MR, Islam T, Shajalal Md, Beyan O, Lange C, Cochez M, Rebholz-Schuhmann D, Decker S: **Explainable AI for bioinformatics: methods, tools and applications.** *Brief Bioinform* 2023, **24**:bbad236, <https://doi.org/10.1093/bib/bbad236>
  5. Cortacero K, McKenzie B, Müller S, Khazen R, Lafouresse F, Corsaut G, Van Acker N, Frenois F-X, Lamant L, Meyer N, Vergier B, Wilson DG, Luga H, Staufer O, Dustin ML, Valitutti S, Cussat-Blanc S, **Kartezio: evolutionary design of explainable pipelines for biomedical image analysis.** *arXiv [cs.CV]*; 2023. <http://arxiv.org/abs/2302.14762>.
- Introduction of a tool that uses machine learning techniques to develop pipelines composed of combinations of classical, interpretable image processing methods for segmentations.
6. Stringer C, Wang T, Michaelos M, Pachitariu M: **Cellpose: a generalist algorithm for cellular segmentation.** *Nat Methods* 2021, **18**:100-106, <https://doi.org/10.1038/s41592-020-01018-x>
  7. Schmidt U, Weigert M, Broaddus C, Myers G: **Cell detection with Star-Convex polygons.** Springer International Publishing; 2018:265-273, [https://doi.org/10.1007/978-3-030-00934-2\\_30](https://doi.org/10.1007/978-3-030-00934-2_30)
  8. He K, Gkioxari G, Dollár P, Girshick R: **Mask R-CNN.** *IEEE Trans Pattern Anal Mach Intell* 2020, **42**:386-397, <https://doi.org/10.1109/TPAMI.2018.2844175>
  9. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez J-Y, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A: **Fiji: an open-source platform for biological-image analysis.** *Nat Methods* 2012, **9**:676-682, <https://doi.org/10.1038/nmeth.2019>
  10. Stirling DR, Swain-Bowden MJ, Lucas AM, Carpenter AE, Cimini BA, Goodman A: **CellProfiler 4: improvements in speed, utility and usability.** *BMC Bioinform* 2021, **22**:433, <https://doi.org/10.1186/s12859-021-04344-9>
  11. Ahlers J, Althviz Moré D, Amsalem O, Anderson A, Bokota G, Boone P, Bragantini J, Buckley G, Burt A, Bussonnier M, Can Solak A, Caporal C, Doncila Pop D, Evans K, Freeman J, Gaifas L, Gohlke C, Gunalan K, Har-Gil H, Harfouche M, Harrington KIS, Hilsenstein V, Hutchings K, Lambert T, Lauer J, Lichtner G, Liu Z, Liu L, Lowe A, Marconato L, Martin S, McGovern A, Migas L, Miller N, Muñoz H, Müller J-H, Nuroth-Kreß C, Nunez-Iglesias J, Pape C, Pevey K, Peña-Castellanos G, Pierré A, Rodríguez-Guerra J, Ross D, Royer L, Russell CT, Selzer G, Smith P, Sobolewski P, Sofiiuk K, Sofroniew N, Stansby D, Sweet A, Vierdag W-M, Wadhwa P, Weber Mendonça M, Windhager J, Winston P, Yamauchi K: **napari: a multi-dimensional image viewer for Python.** *Zenodo* 2023,1-2, <https://doi.org/10.5281/ZENODO.3555620>
  12. Haase R, Fazeli E, Legland D, Doube M, Culley S, Belevich I, Jokitalo E, Schorb M, Klemm A, Tischer C: **A Hitchhiker's guide through the bio-image analysis software universe.** *FEBS Lett* 2022, **596**:2472-2485, <https://doi.org/10.1002/1873-3468.14451>.
- A review of the state of bio-image analysis tools, for segmentation and beyond, up to 2022.
13. Selzer GJ, Rueden CT, Hiner MC, Evans EL 3rd, Harrington KIS, Eliceiri KW: **napari-imagej: ImageJ ecosystem access from napari.** *Nat Methods* 2023, **20**:1443-1444, <https://doi.org/10.1038/s41592-023-01990-0>
  14. Weisbart E, Tromans-Coia C, Diaz-Rohrer B, Stirling DR, Garcia-Fossa F, Senft RA, Hiner MC, de Jesus MB, Eliceiri KW, Cimini BA: **CellProfiler plugins – an easy image analysis platform integration for containers and Python tools.** *J Microsc* 2023, **1**:1-8, <https://doi.org/10.1111/jmi.13223>
  15. Rueden CT, Hiner MC, Evans EL 3rd, Pinkert MA, Lucas AM, Carpenter AE, Cimini BA, Eliceiri KW: **PylmageJ: a library for integrating ImageJ and Python.** *Nat Methods* 2022, **19**:1326-1327, <https://doi.org/10.1038/s41592-022-01655-4>
  16. Barry DJ, Gerri C, Bell DM, D'Antuono R, Niakan KK: **GIANI – open-source software for automated analysis of 3D microscopy images.** *J Cell Sci* 2022, **135**:jcs259511 (<https://journals.biologists.com/jcs/article-abstract/135/10/jcs259511/275435>).
  17. Arzt M, Deschamps J, Schmiech C, Pietzsch T, Schmidt D, Tomancak P, Haase R, Jug F: **LABKIT: labeling and segmentation Toolkit for Big Image Data.** *Front Comput Sci* 2022, **4**:1-12, <https://doi.org/10.3389/fcomp.2022>
  18. Ritchie A, Laitinen S, Katajisto P, Englund JI: **“Tonga”: a novel toolbox for straightforward bioimage analysis.** *Front Comput Sci* 2022, **4**:1-9, <https://doi.org/10.3389/fcomp.2022>
  19. Uhlmann V, Donati L, Sage D: **A practical guide to supervised deep learning for bioimage analysis: challenges and good practices.** *IEEE Signal Process Mag* 2022, **39**:73-86, <https://doi.org/10.1109/msp.2021.3123589>
  20. Tian J, Hsu Y-C, Shen Y, Jin H, Kira Z: **Exploring covariate and concept shift for detection and calibration of out-of-distribution data.** *arXiv [cs.LG]*; 2021. (<http://arxiv.org/abs/2110.15231>).
  21. Isola P, Zhu J-Y, Zhou T, Efros AA: **Image-to-image translation with conditional adversarial networks.** *arXiv [cs.CV]*; 2016:1125-1134. ([http://openaccess.thecvf.com/content\\_cvpr\\_2017/html/Isola\\_Image-To-Image\\_Translation\\_With\\_CVPR\\_2017\\_paper.html](http://openaccess.thecvf.com/content_cvpr_2017/html/Isola_Image-To-Image_Translation_With_CVPR_2017_paper.html)) [Accessed 27 November 2023].
  22. Hollandi R, Szkalitsy A, Toth T, Tasnadi E, Molnar C, Mathe B, Grexa I, Molnar J, Balind A, Gorbe M, Kovacs M, Migh E, Goodman A, Balassa T, Koos K, Wang W, Caicedo JC, Bara N, Kovacs F, Paaolainen L, Danka T, Kriston A, Carpenter AE, Smith K, Horvath P: **nucleAIzer: a parameter-free deep learning framework for nucleus segmentation using image style transfer.** *Cell Syst* 2020, **10**:453-458.e6, <https://doi.org/10.1016/j.cels.2020.04.003>
  23. Lee MY, Bedia JS, Bhate SS, Barlow GL, Phillips D, Fantl WJ, Nolan GP, Schürch CM: **CellSeg: a robust, pre-trained nucleus segmentation and pixel quantification software for highly multiplexed fluorescence images.** *BMC Bioinform* 2022, **23**:46, <https://doi.org/10.1186/s12859-022-04570-9>
  24. McQuin C, Goodman A, Chernyshev V, Kamentsky L, Cimini BA, Karhohs KW, Doan M, Ding L, Rafelski SM, Thirstrup D, Wiegreae W, Singh S, Becker T, Caicedo JC, Carpenter AE: **CellProfiler 3.0: next-generation image processing for biology.** *PLoS Biol* 2018, **16**:e2005970, <https://doi.org/10.1371/journal.pbio.2005970>
  25. Mandal S, Uhlmann V: **SplineDist: automated cell segmentation with spline curves.** *bioRxiv*; 2021. 2020.10.27.357640. (<https://doi.org/10.1101/2020.10.27.357640>).
  26. Walter FC, Damrich S, Hamprecht FA: **MultiStar: instance segmentation of overlapping objects with star-convex polygons.** *arXiv [cs.CV]*; 2020. <http://arxiv.org/abs/2011.13228>.
  27. Bankhead P, Loughrey MB, Fernández JA, Dombrowski Y, McArt DG, Dunne PD, McQuaid S, Gray RT, Murray LJ, Coleman HG, James JA, Salto-Tellez M, Hamilton PW: **QuPath: open source software for digital pathology image analysis.** *Sci Rep* 2017, **7**:16878, <https://doi.org/10.1038/s41598-017-17204-5>
  28. Humphries MP, Maxwell P, Salto-Tellez M: **QuPath: the global impact of an open source digital pathology system.** *Comput Struct Biotechnol J* 2021, **19**:852-859, <https://doi.org/10.1016/j.csbj.2021.01.022>
  29. de Chaumont F, Dallongeville S, Chenouard N, Hervé N, Pop S, Provoost T, Meas-Yedid V, Pankajakshan P, Lecomte T, Le Montagner Y, Lagache T, Dufour A, Olivo-Marin J-C: **Icy: an open bioimage informatics platform for extended reproducible research.** *Nat Methods* 2012, **9**:690-696, <https://doi.org/10.1038/nmeth.2075>
  30. Fillbrunn A, Dietz C, Pfeuffer J, Rahn R, Landrum GA, Berthold MR: **KNIME for reproducible cross-domain analysis of life science data.** *J Biotechnol* 2017, **261**:149-156, <https://doi.org/10.1016/j.jbiotec.2017.07.028>



31. Pachitariu M, Stringer C: **Cellpose 2.0: how to train your own model**. *Nat Methods* 2022, **19**:1634-1641, <https://doi.org/10.1038/s41592-022-01663-4>.  
User-friendly software for human-in-the-loop fine-tuning of DL models applied to cell segmentation.
32. Cutler KJ, Stringer C, Lo TW, Rappaport L, Stroustrup N, Brook Peterson S, Wiggins PA, Mougous JD: **Omnipose: a high-precision morphology-independent solution for bacterial cell segmentation**. *Nat Methods* 2022, **19**:1438-1448, <https://doi.org/10.1038/s41592-022-01639-4>
33. Greenwald NF, Miller G, Moen E, Kong A, Kagel A, Dougherty T, Fullaway CC, McIntosh BJ, Leow KX, Schwartz MS, Pavelchek C, Cui S, Camplisson I, Bar-Tal O, Singh J, Fong M, Chaudhry G, Abraham Z, Moseley J, Warshawsky S, Soon E, Greenbaum S, Risom T, Hollmann T, Bendall SC, Keren L, Graf W, Angelo M, Van Valen D: **Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning**. *Nat Biotechnol* 2022, **40**:555-565, <https://doi.org/10.1038/s41587-021-01094-0>
34. Ma J, Wang B: **Towards foundation models of biological image segmentation**. *Nat Methods* 2023, **20**:953-955, <https://doi.org/10.1038/s41592-023-01885-0>
35. Laine RF, Arganda-Carreras I, Henriques R, Jacquemet G:  
• **Avoiding a replication crisis in deep-learning-based bioimage analysis**. *Nat Methods* 2021, **18**:1136-1144, <https://doi.org/10.1038/s41592-021-01284-3>.  
Opinion on best practices for implementing and reporting on the use of development of DL image analysis tools.
36. Ma J, Xie R, Ayyadhury S, Ge C, Gupta A, Gupta R, Gu S, Zhang Y, Lee G, Kim J, Lou W, Li H, Upschulte E, Dickscheid T, de Almeida JG, Wang Y, Han L, Yang X, Labagnara M, Rahi SJ, Kempster C, Pollitt A, Espinosa L, Mignot T, Middeke JM, Eckardt J-N, Li W, Li Z, Cai X, Bai B, Greenwald NF, Van Valen D, Weisbart E, Cimini BA, Li Z, Zuo C, Brück O, Bader GD, Wang B: **The multi-modality cell segmentation challenge: towards universal solutions**. *arXiv [eess.IV]*; 2023. (<http://arxiv.org/abs/2308.05864>).  
Report on the results of the Multi-modality Cell Segmentation Challenge, wherein teams competed to develop state-of-the-art generalist models for cell segmentation, utilizing the latest advancements in developing DL architectures, preprocessing techniques, and data augmentation.
37. Edlund C, Jackson TR, Khalid N, Bevan N, Dale T, Dengel A, Ahmed S, Trygg J, Sjögren R: **LIVECell — a large-scale dataset for label-free live cell segmentation**. *Nat Methods* 2021, **18**:1038-1045, <https://doi.org/10.1038/s41592-021-01249-6>
38. Maška M, Ulman V, Delgado-Rodriguez P, Gómez-de-Mariscal E, Nečasová T, Guerrero Peña FA, Ren TI, Meyerowitz EM, Scherr T, Löffler K, Mikut R, Guo T, Wang Y, Allebach JP, Bao R, Al-Shakarji NM, Rahmon G, Toubal IE, Palaniappan K, Lux F, Matula P, Sugawara K, Magnusson KEG, Aho L, Cohen AR, Arbelle A, Ben-Haim T, Raviv TR, Isensee F, Jäger PF, Maier-Hein KH, Zhu Y, Ederra C, Urbioa A, Meijering E, Cunha A, Muñoz-Barrutia A, Kozubek M, Ortiz-de-Solórzano C: **The cell tracking challenge: 10 years of objective benchmarking**. *Nat Methods* 2023, **20**:1010-1020, <https://doi.org/10.1038/s41592-023-01879-y>
39. Dey N, Mazdak Abulnaga S, Billot B, Turk EA, Ellen Grant P, Dalca AV, Golland P: **AnyStar: Domain randomized universal star-convex 3D instance segmentation**. *arXiv [cs.CV]*; 2023. (<http://arxiv.org/abs/2307.07044>).
40. Scherr T, Löffler K, Böhlend M, Mikut R: **Cell segmentation and tracking using CNN-based distance predictions and a graph-based matching strategy**. *PLoS One* 2020, **15**:e0243219, <https://doi.org/10.1371/journal.pone.0243219>
41. Xie E, Wang W, Yu Z, Anandkumar A, Alvarez JM, Luo P: **SegFormer: simple and efficient design for semantic segmentation with transformers**. *Adv Neural Inf Process Syst* 2021, **34**:12077-12090 (<https://proceedings.neurips.cc/paper/2021/hash/64f1f27bf1b4ec22924fd0acb550c235-Abstract.html>).
42. Liu Z, Mao H, Wu C-Y, Feichtenhofer C, Darrell T, Xie S: **A ConvNet for the 2020s**. *arXiv [cs.CV]*; 2022:11976-11986. ([http://openaccess.thecvf.com/content/CVPR2022/html/Liu\\_A\\_ConvNet\\_for\\_the\\_2020s\\_CVPR\\_2022\\_paper.html](http://openaccess.thecvf.com/content/CVPR2022/html/Liu_A_ConvNet_for_the_2020s_CVPR_2022_paper.html)) [Accessed 28 September 2023].
43. S. Xie, R. Girshick, P. Dollár, Z. Tu, K. He: **Aggregated residual transformations for deep neural networks**. *arXiv [cs.CV]*; 2016:1492-1500. ([http://openaccess.thecvf.com/content\\_cvpr\\_2017/html/Xie\\_Aggregated\\_Residual\\_Transformations\\_CVPR\\_2017\\_paper.html](http://openaccess.thecvf.com/content_cvpr_2017/html/Xie_Aggregated_Residual_Transformations_CVPR_2017_paper.html)) [Accessed September 28, 2023].
44. Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH: **nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation**. *Nat Methods* 2021, **18**:203-211, <https://doi.org/10.1038/s41592-020-01008-z>
45. G. Lee, S. Kim, J. Kim, S.-Y. Yun: **MEDIAR: harmony of data-centric and model-centric for multi-modality microscopy**. *arXiv [cs.CV]*; 2022. (<http://arxiv.org/abs/2212.03465>).
46. Royer LA: **The future of bioimage analysis: a dialog between mind and machine**. *Nat Methods* 2023, **20**:951-952, <https://doi.org/10.1038/s41592-023-01930-y>
47. C. Wu, S. Yin, W. Qi, X. Wang, Z. Tang, N. Duan: **Visual ChatGPT: talking, drawing and editing with visual foundation models**. *arXiv [cs.CV]*; 2023. (<http://arxiv.org/abs/2303.04671>).
48. Li X, Zhang Y, Wu J, Dai Q: **Challenges and opportunities in bioimage analysis**. *Nat Methods* 2023, **20**:958-961, <https://doi.org/10.1038/s41592-023-01900-4>
49. Gao Y, Zhou M, Metaxas DN: **UTNet: a hybrid transformer architecture for medical image segmentation**. Springer International Publishing; 2021:61-71, , [https://doi.org/10.1007/978-3-030-87199-4\\_6](https://doi.org/10.1007/978-3-030-87199-4_6)
50. Wang R, Lei T, Cui R, Zhang B, Meng H, Nandi AK: **Medical image segmentation using deep learning: a survey**. *IET Image Proc* 2022, **16**:1243-1267, <https://doi.org/10.1049/ipr2.12419>
51. A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A.C. Berg, W.-Y. Lo, P. Dollár, R. Girshick: **Segment anything**. *arXiv [cs.CV]*; 2023. (<http://arxiv.org/abs/2304.02643>).
52. Caicedo JC, Goodman A, Karhohs KW, Cimini BA, Ackerman J, Haghighi M, Heng C, Becker T, Doan M, McQuinn C, Rohban M, Singh S, Carpenter AE: **Nucleus segmentation across imaging experiments: the 2018 Data Science Bowl**. *Nat Methods* 2019, **16**:1247-1253, <https://doi.org/10.1038/s41592-019-0612-7>
53. A. Archit, S. Nair, N. Khalid, P. Hilt, V. Rajashekar, M. Freitag, S. Gupta, A. Dengel, S. Ahmed, C. Pape: **Segment anything for microscopy**. *bioRxiv*; 2023:2023.08.21.554208. (<https://doi.org/10.1101/2023.08.21.554208>).
54. Paul-Gilloteaux P: **Bioimage informatics: Investing in software usability is essential**. *PLoS Biol* 2023, **21**:e3002213, <https://doi.org/10.1371/journal.pbio.3002213>
55. Kemmer I, Keppler A, Serrano-Solano B, Rybina A, Özdemir B, Bischof J, El Ghadraoui A, Eriksson JE, Mathur A: **Building a FAIR image data ecosystem for microscopy communities**. *Histochem Cell Biol* 2023, **160**:199-209, <https://doi.org/10.1007/s00418-023-02203-7>
56. Wilkinson MD, Dumontier M, Aalbersberg IJJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJG, Groth P, Goble C, Grethe JS, Heringa J, 't Hoen PAC, Hooft R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone S-A, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B: **The FAIR Guiding Principles for scientific data management and stewardship**. *Sci Data* 2016, **3**:160018, <https://doi.org/10.1038/sdata.2016.18>
57. Merkel D: **Docker: lightweight Linux containers for consistent development and deployment**. *Linux J* 2014;2, <https://doi.org/10.5555/2600239.2600241>
58. Y. Zhou, J. Sonneck, S. Banerjee, S. Dörr, A. Grüneboom, K. Lorenz, J. Chen: **EfficientBioAI: making bioimaging AI models efficient in energy, latency and representation**. *arXiv [cs.LG]*; 2023. (<http://arxiv.org/abs/2306.06152>).
59. B.M. Saraiva, I.M. Cunha, A.D. Brito, G. Follain, R. Portela, R. Haase, P.M. Pereira, G. Jacquemet, R. Henriques: **NanoPyc:**



- super-fast bioimage analysis powered by adaptive machine learning.** *bioRxiv*; 2023: 2023.08.13.553080. (<https://doi.org/10.1101/2023.08.13.553080>).
60. R. Haase, A. Jain, S. Rigaud, D. Vorkel, P. Rajasekhar, T. Suckert, T.J. Lambert, J. Nunez-Iglesias, D.P. Poole, P. Tomancak, E.W. Myers: **Interactive design of GPU-accelerated Image Data Flow Graphs and cross-platform deployment using multi-lingual code generation.** *bioRxiv*; 2020: 2020.11.19.386565. (<https://doi.org/10.1101/2020.11.19.386565>).
  61. Ouyang W, Eliceiri KW, Cimini BA: **Moving beyond the desktop: prospects for practical bioimage analysis via the web.** *Front Bioinform* 2023, **3**:1233748, <https://doi.org/10.3389/fbinf.2023.1233748>
  62. von Chamier L, Laine RF, Jukkala J, Spahn C, Krentzel D, Nehme E, Lerche M, Hernández-Pérez S, Mattila PK, Karinou E, Holden S, Solak AC, Krull A, Buchholz T-O, Jones ML, Royer LA, Leterrier C, Shechtman Y, Jug F, Heilemann M, Jacquemet G, Henriques R: **Democratising deep learning for microscopy with ZeroCostDL4Mic.** *Nat Commun* 2021, **12**:2276, <https://doi.org/10.1038/s41467-021-22518-0>
  63. da Veiga Leprevost F, Grüning BA, Alves Aflitos S, Röst HL, Uszkoreit J, Barsnes H, Vaudel M, Moreno P, Gatto L, Weber J, Bai M, Jimenez RC, Sachsenberg T, Pfeuffer J, Vera Alvarez R, Griss J, Nesvizhskii AI, Perez-Riverol Y: **BioContainers: an open-source and community-driven framework for software standardization.** *Bioinformatics* 2017, **33**:2580-2582, <https://doi.org/10.1093/bioinformatics/btx192>
  64. Bai J, Bandla C, Guo J, Vera Alvarez R, Bai M, Vizcaino JA, Moreno P, Grüning B, Sallou O, Perez-Riverol Y: **BioContainers registry: searching bioinformatics and proteomics tools, packages, and containers.** *J Proteome Res* 2021, **20**:2056-2061, <https://doi.org/10.1021/acs.jproteome.0c00904>
  65. Rubens U, Mormont R, Paavolainen L, Bäcker V, Pavie B, Scholz LA, Michiels G, Maška M, Ünay D, Ball G, Hoyoux R, Vandaele R, Golani O, Stanciu SG, Sladoje N, Paul-Gilloteaux P, Marée R, Tosi S: **BIAFLOWS: a collaborative framework to reproducibly deploy and benchmark bioimage analysis workflows.** *Patterns* 2020, **1**:100040, <https://doi.org/10.1016/j.patter.2020.100040>
  66. Prigent S, Valades-Cruz CA, Leconte L, Maury L, Salamero J, Kervrann C: **BiolmageIT: open-source framework for integration of image data management with analysis.** *Nat Methods* 2022, **19**:1328-1330, <https://doi.org/10.1038/s41592-022-01642-9>
  67. W. Ouyang, F. Beuttenmueller, E. Gómez-de-Mariscal, C. Pape, T. Burke, C. Garcia-López-de-Haro, C. Russell, L. Moya-Sans, C. de-la-Torre-Gutiérrez, D. Schmidt, D. Kutra, M. Novikov, M. Weigert, U. Schmidt, P. Bankhead, G. Jacquemet, D. Sage, R. Henriques, A. Muñoz-Barrutia, E. Lundberg, F. Jug, A. Kreshuk: **Biolmage Model Zoo: a community-driven resource for accessible deep learning in bioimage analysis.** *bioRxiv*; 2022: 2022.06.07.495102. (<https://doi.org/10.1101/2022.06.07.495102>).
- Paper introducing a repository of bioimage-specific DL models, a platform for their execution, and a framework for interoperability with bioimage analysis tools.
68. Ouyang W, Mueller F, Hjelmare M, Lundberg E, Zimmer C: **ImJoy: an open-source computational platform for the deep learning era.** *Nat Methods* 2019, **16**:1199-1200, <https://doi.org/10.1038/s41592-019-0627-0>
  69. Bannon D, Moen E, Schwartz M, Borba E, Kudo T, Greenwald N, Vijayakumar V, Chang B, Pao E, Osterman E, Graf W, Van Valen D: **DeepCell Kiosk: scaling deep learning-enabled cellular image analysis with Kubernetes.** *Nat Methods* 2021, **18**:43-45, <https://doi.org/10.1038/s41592-020-01023-0>
  70. Weisbart E, Cimini BA: **Distributed-Something: scripts to leverage AWS storage and computing for distributed workflows at scale.** *Nat Methods* 2023, **20**:1120-1121, <https://doi.org/10.1038/s41592-023-01918-8>
  71. Schmiech C, Nelson MS, Avilov S, Bakker G-J, Bertocchi C, Bischof J, Boehm U, Brocher J, Carvalho MT, Chiritescu C, Christopher J, Cimini BA, Conde-Sousa E, Ebner M, Ecker R, Eliceiri K, Fernandez-Rodriguez J, Gaudreault N, Gelman L, Grunwald D, Gu T, Halidi N, Hammer M, Hartley M, Held M, Jug F, Kapoor V, Koksoy AA, Lacoste J, Le Dévédec S, Le Guyader S, Liu P, Martins GG, Mathur A, Miura K, Montero Llopis P, Nitschke R, North A, Parslow AC, Payne-Dwyer A, Plantard L, Ali R, Schroth-Diez B, Schütz L, Scott RT, Seitz A, Selchow O, Sharma VP, Spitaler M, Srinivasan S, Strambio-De-Castilla C, Taaftjes D, Tischer C, Jambor HK: **Community-developed checklists for publishing images and image analyses.** *Nat Methods* 2023, **1-9**, <https://doi.org/10.1038/s41592-023-01987-9>
  72. Hirling D, Tasnadi E, Caicedo J, Caroprese MV, Sjögren R, Aubreville M, Koos K, Horvath P: **Segmentation metric misinterpretations in bioimage analysis.** *Nat Methods* 2023, <https://doi.org/10.1038/s41592-023-01942-8>
  73. L. Maier-Hein, A. Reinke, P. Godau, M.D. Tizabi, F. Buettner, E. Christodoulou, B. Glocker, F. Isensee, J. Kleesiek, M. Kozubek, M. Reyes, M.A. Riegler, M. Wiesenfarth, A. Emre Kavur, C.H. Sudre, M. Baumgartner, M. Eisenmann, D. Heckmann-Nötzel, A. Tim Rädtsch, L. Acion, M. Antonelli, T. Arbel, S. Bakas, A. Benis, M. Blaschko, M. Jorge Cardoso, V. Cheplygina, B.A. Cimini, G.S. Collins, K. Farahani, L. Ferrer, A. Galdran, B. van Ginneken, R. Haase, D.A. Hashimoto, M.M. Hoffman, M. Huisman, P. Jannin, C.E. Kahn, D. Kainmueller, B. Kainz, A. Karargyris, A. Karthikesalingam, H. Kenngott, F. Kofler, A. Kopp-Schneider, A. Kreshuk, T. Kurc, B.A. Landman, G. Litjens, A. Madani, K. Maier-Hein, A.L. Martel, P. Mattson, E. Meijering, B. Menze, K.G.M. Moons, H. Müller, B. Nichyporuk, F. Nickel, J. Petersen, N. Rajpoot, N. Rieke, J. Saez-Rodríguez, C.I. Sánchez, S. Shetty, M. van Smeden, R.M. Summers, A.A. Taha, A. Tlupin, S.A. Tsaftaris, B. Van Calster, G. Varoquaux, P.F. Jäger: **Metrics reloaded: recommendations for image analysis validation.** *arXiv [cs.CV]*; 2022. (<http://arxiv.org/abs/2206.01653>).
- A wide-sweeping and comprehensive guide on the use of metrics for biomedical analysis tasks in Machine Learning contexts.

## Review

## Mechanobiology of organelles: illuminating their roles in mechanosensing and mechanotransduction

Santosh Phuyal,<sup>1</sup> Patrizia Romani,<sup>2</sup> Sirio Dupont,<sup>2,\*</sup> and Hesso Farhan <sup>1,3,\*</sup>

**Mechanobiology studies the mechanisms by which cells sense and respond to physical forces, and the role of these forces in shaping cells and tissues themselves. Mechanosensing can occur at the plasma membrane, which is directly exposed to external forces, but also in the cell's interior, for example, through deformation of the nucleus. Less is known on how the function and morphology of organelles are influenced by alterations in their own mechanical properties, or by external forces. Here, we discuss recent advances on the mechanosensing and mechanotransduction of organelles, including the endoplasmic reticulum (ER), the Golgi apparatus, the endo-lysosomal system, and the mitochondria. We highlight open questions that need to be addressed to gain a broader understanding of the role of organelle mechanobiology.**

**Organelles: the overlooked players in mechanobiology**

Cells are exposed to a wide variety of mechanical stimuli, resulting in the deformation of both cells and the extracellular matrix. This occurs under physiological conditions such as the deformations of the lung's epithelial cells during breathing [1], or forces exerted on the bone and cartilage during body movement [2,3]. Mechanical forces are also relevant under pathological conditions such as fibrosis or cancer, where the stiffness of tissues can increase by up to tenfold [4,5]. Therefore, understanding cellular responses to mechanical stimuli is highly relevant for both physiology and pathology. Mechanical forces may elicit rapid adaptive changes, such as an increase in the strength of cell–matrix and cell–cell adhesions, or alignment with the direction of fluid flow [6]. Alternatively, cellular responses extend to alterations in gene expression, which ultimately affect cell fate [7–11]. This implies that cells dispose of structures by which they 'sense' physical forces, in addition to the mechanotransduction mechanisms by which forces are translated into biochemical information [12,13]. Initially, the search for cellular mechanosensors has focused on the plasma membrane because it is directly impacted by external forces, as well as by internal forces such as those driving the formation of protrusion, blebbing, or cell contraction [14]. Another reason for the dominance of the plasma membrane in the mechanobiology field is the scarcity of tools to impose physical forces and measure their effect at intracellular locations. However, increasing evidence has indicated that intracellular membranous organelles can receive, modulate, and even initiate mechanotransduction signaling. Moreover, the development of novel tools to measure mechanical forces at organelle membranes is expected to boost discoveries in this research area.

The link between organelles and mechanobiology can occur at multiple levels. First, organelles are delimited by membranes with a defined composition, shape, and biophysical properties, which are fundamental for their biological functions, and they are often connected to the cytoskeleton for structural support. Thus, organelles have their own mechanical properties, which can

**Highlights**

A growing body of evidence indicates that intracellular membrane-bound organelles can receive, modulate, and may even initiate mechanotransduction signaling.

Organelles may sense mechanical forces in various ways such as through membrane deformation, changes in lipid packing, alterations in the membrane contact sites, through the cytoskeleton, or by altering the intraluminal fluid flow within organelles.

The endoplasmic reticulum is equipped with mechanosensitive channels that can directly sense mechanical forces but was also shown to be the target of mechanotransduction signaling pathways that affect its functional organization.

Mitochondrial fusion and fission were shown to be altered in response to extracellular and intracellular mechanical forces, thereby affecting the metabolic function of these organelles.

The Golgi apparatus and the endosomes are involved in the mechanics of cell migration, as well as in mechanosensing and cellular mechanoprotection.

<sup>1</sup>Institute of Basic Medical Sciences, University of Oslo, Oslo, Norway

<sup>2</sup>Department of Molecular Medicine, University of Padua, Padua, Italy

<sup>3</sup>Institute of Pathophysiology, Biocenter, Medical University of Innsbruck, Innsbruck, Austria

\*Correspondence: [sirio.dupont@unipd.it](mailto:sirio.dupont@unipd.it) (S. Dupont) and [hesso.farhan@i-med.ac.at](mailto:hesso.farhan@i-med.ac.at) (H. Farhan).



change depending on the varying intra- and extracellular needs. Second, mechanical stimuli are known to affect biological processes that are intimately connected to the organelles of the endomembrane system. For instance, the function and the positioning of the Golgi apparatus regulates the migration and invasion of cancer cells [15]. The activation of immune cells often requires adaptation of the secretory capacity of the cell [16]. Fibrosis is caused by the secretion of collagen, which is a bulky cargo that requires adaptation of the export machinery of the ER [17]. Third, because many organelles can be deformed in response to external forces, they can play a primary role in mechanosensing and in the processes of mechanotransduction. In the current review, we will highlight and discuss recent evidence showing the role of organelles in mechanobiology. For the sake of space, we have focused on selected membrane-bound organelles. We only mention the role of the nucleus in relation to the ER and leave a more in-depth analysis of this topic to other excellent reviews [18–20].

### Principles of mechanosensing and mechanotransduction at the endomembranes

The most common mechanical forces acting on cells include stretching (tensile forces), compression, shear stress, hydrostatic pressure, and the stiffness of the surrounding extracellular matrix [21]. The cellular response to such forces proceeds in three steps. The first step is that of the mechanical force via mechanosensors. Secondly, the force is converted into biochemical signals via mechanotransduction pathways. Of note, mechanotransduction is not limited to classical signaling pathways, but also includes metabolic changes as well as the regulation of proteins' structure, activity, and trafficking. The third step involves the cellular response, which might be purely adaptive (e.g., aiming to maintain the cellular state) or a change in the cell's state (e.g., the induction of proliferation).

Mechanical force from inside or outside the cell causes different degrees of cellular deformation, depending on the type and magnitude of the mechanical stimulus. The deformation of the cells' surface may then be transmitted to the intracellular membranes, either by direct physical deformation, indirectly via the cytoskeleton, or through the viscoelastic nature of the cytosol. However, mechanotransduction signals originating at the cell's surface may be transmitted to the organelles and affect their structure and function. Whether and how these complementary mechanotransduction pathways intersect each other remains poorly understood.

Although much is unknown about mechanosensing at the endomembranes, here, we present five main modes by which mechanical forces might (in principle) be sensed at/by the organelles:

- (i) Membrane deformation: as outlined above, the mechanical force might result in a deformation in the shape of the organelle, which will cause a change in the tension and lipid packing of its membrane. This will then result in an alteration in the dynamics and activity of the transmembrane proteins. Analogous to the plasma membrane, there might be mechanosensitive ion channels at the endomembranes. In fact, Piezo-1, a bona fide mechanosensory, was shown to be also present in the ER [22]. However, the same principle might well influence the activity of other proteins, including enzymes and signaling mediators. Finally, membrane deformation might affect the recruitment of curvature-sensing proteins to endomembranes.
- (ii) The membrane's contact sites have the potential to be mechanosensitive because mechanical forces might affect the formation and stability of these contact sites. Many proteins are known that stabilize the membranes of apposing organelles at a close distance and enable the exchange of ions, lipids, and proteins among the organelles. Alterations in number and size of contact sites will then trigger downstream effects that regulate cellular physiology [23,24]. However, little is known about the range mechanical force of these tethering events, or on how these adhesive forces modulate the membrane's contact sites.

- (iii) Morphological changes in an organelle's shape might alter the local intraluminal fluid flows. This could be particularly relevant for the ER, where the lumen acts as a reservoir for calcium, as it is also a site for protein folding and quality control. The active (energy-dependent) contraction of ER tubules has been reported to generate intraluminal fluid flow [25], indicating that mechanical factors that affect the shape of the ER could affect this flow. An alteration in the intraluminal flow in the ER was shown to affect the ability of chaperones to access their folding clients, which has consequences for cellular proteostasis [26].
- (iv) Organelles are known to be associated with the cytoskeleton, and the cytoskeleton has great potential for directly transmitting internal and external forces. Alterations in the interaction of organelles with the cytoskeleton might affect the functional organization of the organelle itself and thereby the biology of the cells. For instance, the structure of the Golgi ribbon is maintained not only by microtubules but also by actin [27,28] filaments. Altering the interaction of the Golgi ribbon with the microtubules will result in defective positioning and the fragmentation of this organelle, thereby causing defects in the cell's polarity and migration [15,29,30]. We also discuss how mechanotransduction affects the interaction of the ER with the actin cytoskeleton.

These five modes of mechanosensing at or by the organelles are largely theoretical in nature because the field of organelle-based mechanobiology is still in its infancy. However, the development of techniques to measure organelle rheology [31,32] and of novel tension sensors targeted to various organelles' will provide the community with tools to measure the impact of mechanical forces at the intracellular membranes. One example is the different responses of the ER's subdomains (sheets and tubules) and the mitochondria to hypertonic osmotic shock [33].

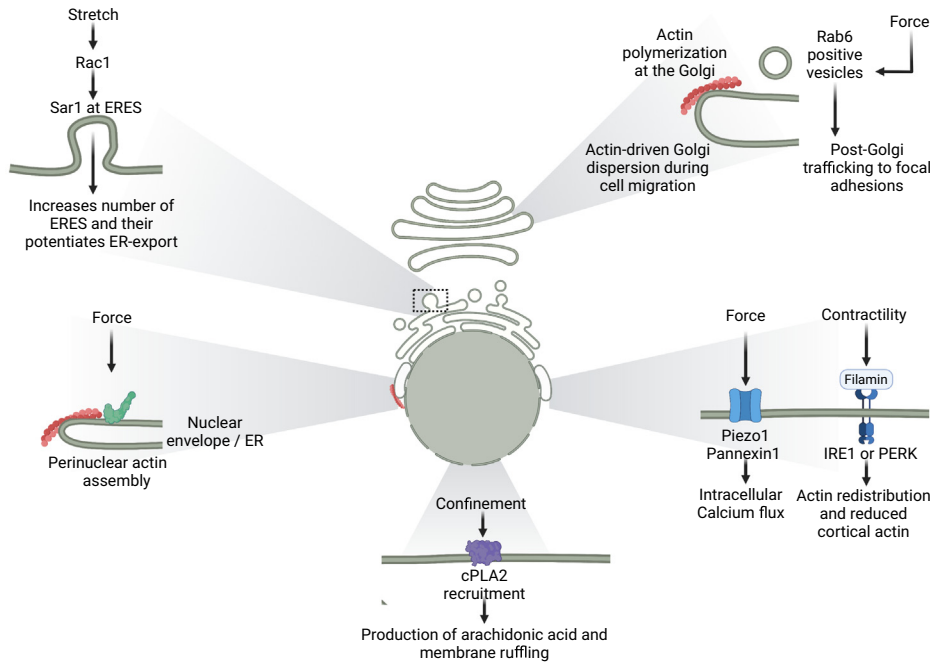
### Mechanobiology and the ER

The organization of the ER (Box 1) appears to be a favored site for mechanosensing and mechanotransduction for two main reasons. Firstly, it spans almost the entire cytosol and is therefore likely to easily undergo structural changes upon deformation of the cell by mechanical forces. Secondly, the ER has close contact with cytoskeletal factors such as actin and microtubules, which are likely to transmit mechanical forces to this organelle.

#### Box 1. The endoplasmic reticulum

In mammalian cells, the ER forms a network of interconnected dynamic tubules and sheets that undergo constant remodeling. The ER is the largest cellular organelle, making up 20–30% of the cellular volume. It is the site of the synthesis, the control of folding quality, and trafficking of a major fraction of the cellular proteome including all transmembrane proteins of the secretory and endocytic organelles and of the plasma membrane, and all the secreted proteins, with estimates assigning at least one-third of the eukaryotic proteome to this compartment [97]. As such, the ER is considered a hub for cellular proteostasis. Moreover, the ER houses a wide range of metabolic enzymes and exchanges metabolic intermediates with other organelles via vesicular transport and direct contact. It plays a key role in the homeostasis of cellular phospholipids and cholesterol, making it a key organelle involved in cellular metabolism. The rough ER is decorated with ribosomes and is the subdomain that is mainly responsible for protein synthesis. On the rough ER, ribosome-free regions can be found that are referred to as ER exit sites (ERES), which represent the sites for the assembly of the components of COPII that are responsible for exporting secretory proteins from the ER towards distal compartments [98–100]. ERES were shown to respond to a variety of signals in response to intracellular and extracellular needs [55,101,102]. Other factors that affect the ER (and the entire endomembrane system) include the shape and composition of lipids, which affect the function and the mechanical properties of membranes [103,104]. However, their regulation in response to mechanical stimuli has only very recently been investigated. The central role of the ER in cellular proteostasis is regulated by the unfolded protein response (UPR), a homeostatic ER-to-nucleus signaling route driven by three transmembrane sensors: IRE1 (inositol-requiring enzyme 1), ATF6 (activating transcription factor 6), and PERK (PKR-like ER kinase) [105]. The induction of ER stress through the accumulation of misfolded proteins activates the UPR, leading to the enhanced capacity of the ER to fold, degrade, and secrete proteins.





Trends in Cell Biology

Figure 1. Overview of some of the effects of mechanical forces on the secretory pathway. The following mechanisms are shown: stretching induced regulation of the functions of endoplasmic reticulum (ER) exit sites (ERES), the force-induced assembly of actin at the ER and the Golgi apparatus, confinement-induced recruitment of cPLA2 to the nuclear envelope, and finally how physical forces affect the Piezo channels or regulators of the unfolded protein response (UPR). The figure was created using BioRender.

Mechanosensitive ion channels have long been recognized as bona fide sensors and transducers of mechanical forces. Piezo-1 (also known as FAM38A) is among the best characterized mechanosensors that can increase ion conductance upon mechanical stimulation of the cell [34]. Piezo-1 was also found to be localized to the ER in its active form, where it recruits R-Ras to mediate the release of calcium and activation of the protease calpain, thus resulting in integrin activation [22]. This work implicated Piezo-1 in ER-based signaling, but it did not investigate whether this mechanism is sensitive to mechanical stress. In fact, a more recent study showed that stretching triggers nuclear deformation, leading to the Piezo-1-mediated release of calcium from the ER (Figure 1) to reduce lamina-associated heterochromatin formation, which in turn affects nuclear mechanics [35].

Another mechanosensitive channel with a link to the ER is Pannexin-1, a large-pore nonselective channel that is permeable to ions and other solutes. Pannexin-1 is well known to be mechanosensitive [36]. The work of two groups linked Pannexin-1 to mechanotransduction at the ER. Treatment of hippocampal neurons with thapsigargin, which blocks the ER-localized calcium pump, resulted in the activation of Pannexin-1 [37]. This effect was mediated via the ER-based calcium-sensor STIM1, which interacts with Pannexin-1. Whether or not Pannexin-1 is localized to the ER was not directly tested in previous work, which assumed that the STIM1–Pannexin interaction occurred via contacts between the ER and the plasma membrane. Another study reported the localization of Pannexin-1 at the ER and proposed a role for this channel in sensing the mechanical stress imposed by focused ultrasound [38]. The activation of ER-localized Pannexin-1 resulted in calcium efflux to the cytosol, unlike Pannexin-1 at the cell’s surface, which was found to primarily mediate the release of ATP.

Because the ER constitutes and is continuous with the outer layer of the nuclear envelope, it was hypothesized that nuclear deformation and positioning can affect the ER's dynamics and functions. Recent studies have indicated that cells measure mechanical confinement by a mechanism that depends on nuclear deformation [39–41]. This is mediated via the recruitment of cPLA2 to the nuclear envelope to promote a distinct type of cell migration based on membrane blebbing to escape confinement. The recruitment of cPLA2 requires the release of calcium from internal stores via channels activated by stretching, thus implying a role for the nuclear envelope and/or the perinuclear ER (Figure 1). The observation that cytoplasts (i.e., cells where the nucleus has been removed) do not respond to confinement was used to exclude a general role for the rest of the ER [39], even if the absence of the nucleus might eliminate the very structure that transmits deformation to the ER in the first instance. Therefore, proximity to the nucleus might endow this ER subdomain with the physical and biochemical properties to sense and respond to cellular confinement. More generally, this suggests that cells use the ER–nucleus unit to sense extreme deformations, a principle that, in future, might be extended to other organelles as well.

The ER can also play a more direct function in mediating cytoskeletal dynamics. During cell migration, the position of the nucleus is regulated, depending on the interaction of the actin cytoskeleton with the nuclear envelope, including the linker of nucleoskeleton and cytoskeleton (LINC) complexes that provide a direct tether between the two [42]. As cells polarized in a wound-healing assay, the ER accumulated on the ventral side (i.e., the side facing the glass slide) of the cell in a manner that depended on the ER protein CLIMP63, which regulates the formation of ER sheets and helps the ER wrap around actin fibers [28]. This was, however, not observed on the dorsal side of the cell (i.e., the side facing away from the glass slide). By shielding the actin fibers, the ER prevents the interaction of actin with the nuclear envelope on the ventral side. In turn, this asymmetry in the nucleus–cytoskeleton interaction results in the movement of the nucleus away from the leading edge, which is important for efficient cell migration. Earlier work investigating the cytoplasmic streaming during meiosis of *Caenorhabditis elegans* zygotes found that the ER is required for this collective flow by binding to microtubules and transmitting the force to adjacent regions [43], indicating the wide potential for the ER in regulating the ability of the cytoskeleton to produce and sense forces.

More recently, a mechanism of mechanocontrol of the ER exit sites (ERES) (Box 1 and Figure 1) was discovered. When the cells were exposed to mechanical strain, this led to activation of the Rac1 GTPase, which formed a complex with Sar1, another GTPase that initiated the assembly of the COPII coat and the formation of outward vesicles [44]. Consequently, mechanical strain increased the number of ERES and promoted secretory trafficking out of the ER. The potential physiological relevance of the mechanocontrol of ERES might be that the secretory trafficking gives the stability when exposed to mechanical strain. This is supported by the observation that cells with altered ER export exhibited more ruptures when exposed to mechanical stress. This might be potentially due to the fact that secretion supplies the plasma membrane with lipids, or it might be due to the trafficking of proteins that help the cells cope with mechanical stress.

The unfolded protein response (UPR) is a central regulator of proteostasis (Box 1). However, there is evidence suggesting roles of the UPR that are independent of proteostasis in the context of mechanobiology. Through the use of proximity-labeling proteomics, the actin cross-linking protein Filamin A was identified as a binding partner of PKR-like ER kinase (PERK) [45]. Notably, Filamin A has previously been shown to link the ER to the actin cytoskeleton, which is important for spreading the ER to the periphery of the cell [46]. The loss of PERK resulted in a skewed subcellular distribution of actin with an accumulation of actin at the cell cortex [45]. This was accompanied by an alteration in the formation of the ER–plasma membrane contact site. The trigger activating the PERK was cytosolic

calcium and was not related to its role in the UPR. This elegant work established an unprecedented proteostasis-independent role of PERK, linking it to the cellular remodeling of actin. This work did not directly test the involvement of mechanobiology in this novel role of PERK, but there is a strong rationale for assuming such a role, given the prominent function of actin in cellular mechanoregulation. The observation that PERK-deficient cells had more prominent cortical actin might indicate a change in cellular stiffness, which could be tested in future work. Later work showed that PERK is not the only UPR branch to interact with Filamin A, but IRE1 also interacts with this actin crosslinker [47]. The interaction of IRE1 with Filamin A is important for the regulation of the dynamics of cellular actin and cell migration. This interaction was dependent on the dimerization of IRE1, but was independent on its role in UPR, because enzymatically inactive mutants of IRE1 were still bound to Filamin A. This work provided another nice example of a link between ER-proteostasis regulators and the actin cytoskeleton (Figure 1). A more recent work capitalized on these earlier findings of the role of Filamin A in the ER and identified a key difference between mammary cells cultured in 2D versus 3D [48]. They found that cells cultured in 2D exhibited greater tension in the cortical actin and the stronger association of filamin with the ER stress sensor PERK, which increased ER stress and reduced cell survival. However, this effect might also be partly caused by differences in the substrate's stiffness, which was supported by findings showing that ER stress is differentially regulated by the stiffness of the extracellular matrix [49].

The ER may also respond to mechanical stimuli through its interaction with components of the cytoskeleton. Various elements of the cytoskeleton were previously shown to interact with the ER. For instance, in the past, the ER has been reported to be a site for the nucleation of actin in a manner that was dependent on INF2, which was shown to be localized in this compartment [50,51]. Mechanical force was applied to the cells' surface to generate a perinuclear actin ring, which was absent in INF2-depleted cells [52]. Whether this actin ring colocalized with the ER was not tested, but this is very likely because the perinuclear region is dominated by ER membranes. Another report showed that culturing normal epithelial cells on stiff substrate resulted in an accumulation of Filamin A in the perinuclear region [53]. Together with the findings above about the link of Filamin A to the ER, future work should analyze whether PERK and IRE1 might be involved in the cellular adaptation of stiffness via interactions with Filamin A (Figure 1).

The Pepperkok group used a systematic approach to investigate how signaling by the extracellular matrix affects the ER cytoskeleton crosstalk [54]. A library of 378 cytoskeletal proteins was depleted (by siRNA) and synergistic effects with the coat protein complex II (COPII) components Sec23A–B were explored. This approach identified proteins associated with focal adhesions (e.g., ROCK1 or PIK3CA), which led to a downregulation of Sec23A when silenced by siRNA. This indicates that a network of cytoskeleton-associated proteins links focal adhesions and extracellular matrix (ECM)-related signaling to the regulation of ER export [54]. In addition, this work indicated that prior systematic approaches used to study the export of the ER [55,56] need to be revisited and studied under conditions that consider the ECM, the stiffness of the substrate, and the dimensionality used for culturing the cells.

### Mechanobiology of the Golgi apparatus

Although the Golgi apparatus constantly exchanges content with the ER via vesicular-tubular carriers, the mechanobiology of the Golgi apparatus is likely to follow slightly different rules from the ER. Three factors should be considered. Firstly, the Golgi apparatus is substantially smaller than the ER; secondly, it has a much more confined subcellular localization; thirdly, unlike with the ER network, the cisternae that form the Golgi stack only have sparse intercisternal connections. This has consequences for the transmissibility of the forces applied to one cisterna on the other parts of the Golgi apparatus. By contrast, the Golgi apparatus is surrounded by actin filaments that play important structural roles [27,28], which might enable the transmission and control of forces to the Golgi apparatus as a whole.

On a tissue level, Szafranski *et al.* [57] used electron microscopy to explore the effects of compressing chondrocytes and their organelles. They found that the compression of cartilage tissues resulted in a reduction in the total cellular volume, together with a decrease in the volume of most organelles such as the ER, nuclei, and mitochondria. At the same time, the volume of the Golgi apparatus remained unaltered. Although these results need to be validated using more advanced methods and assays focused directly on cellular deformation, they nevertheless highlighted potential differences between the ER and the Golgi apparatus with respect to their responses to mechanical forces.

A classic example of a change in the architecture of the Golgi apparatus is in migrating cells. When a cell monolayer is wounded, the leading edge of the plasma membrane expands and the cells migrate directionally into the wound. Under such 2D conditions, the Golgi orients quickly in the direction of cell migration, anterior to the nucleus [15,58]. Reorientation also occurs in 3D, but in this case the Golgi apparatus is positioned posterior to the nucleus, rather than anterior to it. These changes in the positioning of the Golgi apparatus are mainly due to the reorientation of the microtubule network during cell migration. The role of actin remains poorly explored, but recent work reported the dispersion of the Golgi apparatus during collective cell migration that was driven by the association of the actin cytoskeleton with the Golgi apparatus [59].

The first direct demonstration of an effect of mechanical forces on the Golgi apparatus was provided by the Manneville group, who imposed mechanical stress directly onto the Golgi apparatus with an intracellular optical trap system [31]. They found that the Golgi apparatus is endowed with a distinct rheology, which was much higher than that observed by measuring the cytosol surrounding the Golgi apparatus (and hence, probably, the ER). This rheology depends on contractile actin filaments, indicating a mechano-active cytoskeleton surrounding the Golgi apparatus. In addition, the application of these mechanical constraints induced the production of Rab6-positive transport vesicles, indicating a potential link between this mechano-active cytoskeleton and the regulation of protein secretion (Figure 1). These vesicles have also been shown to be preferentially transported along the microtubules towards focal adhesions [60]. These studies suggested wide potential for the mechano-regulation of the secretory pathway departing from the Golgi apparatus through the forces of cell adhesion.

The rigidity of the Golgi apparatus was also shown to be coupled to the stiffness of the extracellular matrix [61]. This happens in the context of a metabolic response to matrix stiffness that originates from the Golgi apparatus: changes in the matrix stiffness or in the tension of the actomyosin regulated the activity of the Lipin1 phosphatidate phosphatase at the endomembranes, altering the ARF1 GTPase at the Golgi apparatus, and thereby affecting the shuttling of SCAP and SREBPs between the ER and the Golgi apparatus. This ultimately resulted in the activation of SREBPs at the Golgi apparatus, driving the expression of lipogenic enzymes and increasing the overall lipogenic tone of the cell in response to a soft matrix [61,62]. Applying force to the Golgi apparatus could increase the measured activity of Lipin 1, suggesting the physical and functional coupling of the Golgi apparatus and extracellular stiffnesses. Similar regulation of the activity of ARF1 in response to adhesion forces, associated with remodeling of the overall structure of the Golgi apparatus, was also observed by others [63], further reinforcing this possibility.

### Mechanobiology of the endolysosomal system

The pleomorphic membranous structures of the endolysosomal system modulate the recycling of biomolecules, thereby contributing to cellular proteostasis and the regulation of cell surface receptors. Moreover, endosomes and lysosomes act as dynamic signaling platforms by recruiting signaling molecules in response to environmental factors such as nutrients. The mechanoregulation of



the early events of the endocytic process at the plasma membrane has been reasonably well explored [64]. However, only limited knowledge exists regarding the contribution of the organelles of the endolysosomal system in cellular mechanotransduction. Because the spatial and temporal distribution of the organelles of the endolysosomal system dictate the rate of flow of signaling molecules, this might be crucial for cellular mechanoadaptation and mechanotransduction. In fact, cells lacking the secretory lysosome v-SNARE VAMP7 failed to properly adapt to changes in the substrate's rigidity [65]. That study also showed that the rigidity of the ECM determined the peripheral distribution of VAMP7-secreting lysosomes, and this peripheral pool was tightly controlled by LRRK1 and VARP in response to biomechanical constraints [65]. Another study also showed that the peripheral pools of secretory lysosomes were regulated in a manner that depended on the matrix's stiffness [66]. The endolysosomal system must also be able to rapidly adapt to changes in the endocytic flux, which means that endosomes and lysosomes must alter their surface-to-volume ratio due to fusion and fission events. Because of their small size, such changes in shape may have severe consequences with respect to the membrane tension of these organelles. This led the Grimm group to search for a mechanism that senses and adapts to these local mechanical forces. They identified the endolysosomal cation channel TRPML2 as a hypotonicity-sensitive protein that allows the endolysosomal system to adapt to mechanical stress [67].

Rather than being regulated by mechanical forces, endocytosis was also shown to regulate the tension of the plasma membrane. The CLIC/GEEC pathway (a dynamin-independent endocytic route) was found to regulate the tension of the plasma membrane, as its activation would increase tension. However, altering the plasma membrane's tension had no effect on the endocytosis of CLIC/GEEC [68]. Caveolae are characteristic plasma membrane invaginations, which are known to regulate endocytic trafficking. Caveolae make up a substantial fraction of the plasma membrane's surface and, as such, represent a reservoir that allows cells to respond to an increase in the cell's surface tension [69]. This mediates cellular mechanoprotection as well as the ability to sense forces. We refer the reader to the recent excellent reviews on the roles of the caveolae in mechanobiology [70].

Recently, the role of endosomes in regulating the mechanics of collective cell migration (a mode of locomotion that is important for wound healing, organogenesis, and metastasis) has also emerged. Within cell monolayers, cells are still able to migrate. However, the increasing cell-cell interactions associated with increased cell densities gradually constrain cell migration, first leading to long-range coordinated streams of motion and then to kinetical arrest, similar to 'jamming' or a transition of rigidity. Interestingly, the activation of endocytosis by Rab5 GTPase was sufficient to change the material properties of the monolayer, and to induce the emergence of large multicellular streams, similar to a 'flowing liquid' [71]. Collectively, these findings highlight the importance of endolysosomal system for cellular mechanoadaptation.

### Mechanobiology of mitochondria

Mitochondria are filamentous organelles that undergo rapid and dynamic cycles of fission and fusion. The mechanics of mitochondrial dynamics (Box 2) have only started to be analyzed. Fission entails the recruitment of DRP1 at the mitochondrial surface and the formation of linear polymers that encircle and constrict the mitochondria, leading to abscission of the two mitochondrial membranes [72–74]. It has been proposed that the mitochondrial membrane is under tension, which is instrumental to allowing fission [75]. Analogous to other membrane compartments, this indicates the importance of the lipid composition of the mitochondrial membrane as a key factor influencing the membrane's biophysical properties and mechanical behavior [76,77].

### Box 2. Mitochondria fusion and fission

The structure of the mitochondrial tubular network is very dynamic owing to the continuous separation of some elements (fission) and addition of others (fusion). This is mediated by a dedicated set of dynamic GTPases that mediate both fusion (Mitofusin1-2 and OPA1) and fission (DRP1) in the mitochondria [72]. The expression and activity of these factors is regulated by multiple inputs, so that the mitochondrial dynamics match the cells' metabolic demands [106–108]. Regulation of the mitochondrial dynamics also has a more structural role; for example, shortening of the mitochondria occurs at mitosis to allow undisturbed chromosome segregation and the homogeneous distribution of healthy mitochondria to the daughter cells. Shortening also occurs upon mitochondrial stress to facilitate the elimination of damaged mitochondria via mitophagy and, in its most extreme form, serves for the production of small mitochondrial vesicles, as part of the mechanisms of quality control [108–110]. An analysis of the topological properties of the mitochondrial network suggested that mitochondrial dynamics keep the network at criticality, near to a percolation phase transition characterized by maximal heterogeneity in the sizes of the network's subcomponents, which can undergo changes towards subcritical or supercritical conditions depending on the perturbations [111–113].

Mitochondrial fission is intimately linked with the actin cytoskeleton (Figure 2). Actin is polymerized in close vicinity to the sites of fission by ER- and mitochondria-anchored polymerizing factors (INF2 and Spire1c, respectively), and close to the activity of Arp2/3 complexes [78–80]. Recruitment of the actin-polymerizing protein INF2 to the ER drives the formation of perimitochondrial actin filaments when the ER's tubules encircle the mitochondria, initiating the process of fission [78]. The Spire1c splice isoform of the Spire1 actin nucleator was found to be associated with the outer mitochondrial membrane instead [80]. Little is known about the recruitment and activation of Arp2/3 actin nucleators at the mitochondria, apart from the observation that the MIEF1-2 transmembrane proteins that act as DRP1 receptors at the outer membrane of the mitochondria

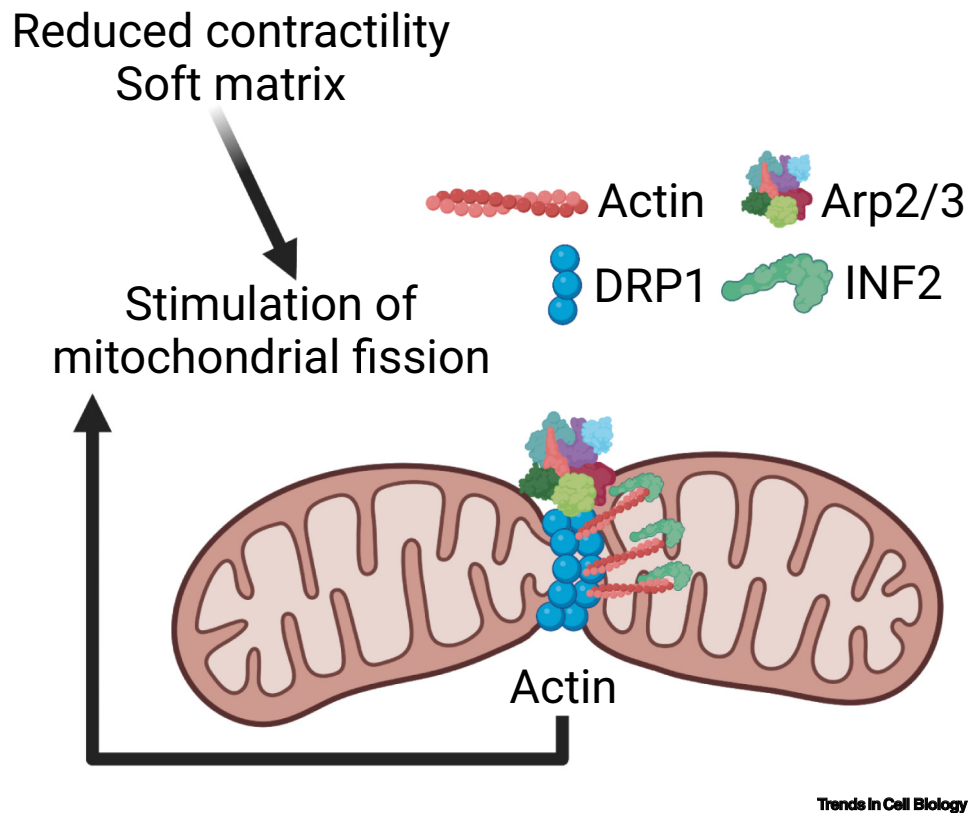


Figure 2. Overview of how mechanical forces affect mitochondrial fission via local polymerization of actin. The figure was created using BioRender.

are found in close proximity to Arp2/3-containing complexes [81]. These factors are widely required for mitochondrial fission, indicating the important role of perimitochondrial actin. It has been proposed that the function of F-actin in this context is a mechanical one that helps the deformation and constriction of the mitochondria. It is also possible that actin plays a role in stabilizing the formation of DRP1 filaments. In line with the first idea, inducing the polymerization of actin at the mitochondrial surface can be sufficient to drive mitochondrial fission, even in cells lacking the DRP1 protein [82,83]. The possibility also exists that the mitochondria have the inbuilt ability to undergo fission upon direct deformation. For example, it was shown that intracellular *Shigella* bacteria 'bumping' against mitochondria can induce mitochondrial fission [84]. Such fission can be recapitulated by simply applying local cell deformations, and this can happen even in the complete absence of filamentous actin. Further adding to this picture, the formation of mitochondrial actin 'comet tails' can play an active role in displacing and mixing mitochondria across the cytoplasm [85]. An emerging theme is that different fission events take place in the cell, which have different requirements in terms of these perimitochondrial actin polymerization factors.

Mitochondrial fission is not only a mechanical process, but also a mechanoresponsive event. Recent reports have indicated the wide potential for regulating of mitochondrial fission by actomyosin and by mechanical stimuli, even if the reports are sometimes divergent [86–93]. A coherent set of evidence indicates that the mitochondria shorten in response to conditions of reduced actomyosin tension across multiple cell lineages. In mammary epithelial and breast cancer cells, this sparks the production of mitochondrial reactive oxygen species (ROS) [92]. These ROS do not reach a level that induces toxicity, but activate a protective response mediated by the master transcription factor of antioxidants, NRF2, and involving the rewiring of the glutathione metabolism. This response makes the cells in a soft microenvironment able to better resist exogenous oxidative stress, including ROS-dependent drugs. This also represents a mechanism that is imposed on disseminating cancer cells by the soft environment of the metastatic target organs, accounting for the notorious resistance of these cells to chemotherapy [92]. In other cells, mitochondrial fission is associated with changes in the urea cycle and in the ATP-recycling system of creatine phosphagen and in the mitochondrial metabolism of proline [87,93]. Mitochondrial function can also be regulated by the regularity of mechanical forces, with monotonous stretching cycles leading to reduced ATP production and variable stretching cycles leading to enhanced ATP production [94]. Finally, the observation that shear stress regulates anti-inflammatory responses in the endothelia by regulating mitochondrial dynamics and the metabolism in the endothelia [91] suggests a broader than expected role for mechanically regulated mitochondrial dynamics in regulating multiple metabolic and signaling responses.

### Concluding remarks

Although it was realized early on that mechanical forces are of relevance for physiology and pathology, the field of mechanobiology is comparatively young and is thus a rapidly growing area of research. As such, considerable advances have been made with regard to our understanding of role of the plasma membrane in sensing mechanical forces and converting these into biochemical stimuli. More recently, the role of the nucleus as a mechanosensitive organelle has been explored. However, the role of mechanobiology in the regulation of other organelles such as the ER, Golgi apparatus, or mitochondria remains an interesting area for future research. As such, there is a growing interest in this research area and there are many questions that need to be resolved (see [Outstanding questions](#)). In the future, new approaches have to be used to study the mechanics of organelles, such as Brillouin microscopy [32], intracellular optic tweezers [95], or organelle-targeted Flipper-TR probes [33]. It is likely that these tools will be developed further, making them more accessible to a larger community. These and other novel tools will allow us to quantitatively and qualitatively measure forces at endomembranes in the future [96], thereby

### Outstanding questions

What are the forces that organelles experience? Better and more precise methods to measure forces at organelles have recently been introduced, and novel tools are likely to be developed to address this question.

To what extent are changes in the hydrostatic pressure of the cytosol transmitted to the lumen of organelles and what are the consequences thereof?

How do different mechanical forces affect the size, volume, morphology, and function of organelles? Do subcellular organelles harbor unidentified force-sensing molecules?

Are there autochthonous signaling events at organelles that are triggered by mechanical forces? To what extent do they intersect with mechanotransduction pathways originating from the plasma membrane?

Are organelles damaged by forces, how is such damage is dealt with, or how does it serve as an additional signaling mechanism in the cell?

How are these seemingly different organelle responses coordinated with the transcription of nuclear gene into a coherent biological response?

What is the broader relevance of organelles' mechanobiology in physiology and pathology?

allowing us to better understand the biophysical principles of how these forces are transmitted to and sensed by intracellular organelles. We will also need systematic explorations to identify bona fide mechanosensors and mechanotransducers that reside in different endomembranes. Finally, the mechanobiology of endomembranes needs to be investigated using adequate (patho)physiological model systems *in vitro* and *in vivo* to provide us with a better understanding of the broader relevance of organelles in the physiological and pathological responses to forces.

### Acknowledgments

H.F. acknowledges funding from the Norwegian Cancer Society (Grant 208015), the Norwegian Research Council (Grants 262652 and 302452), the Austrian Science Foundation (FWF Grants numbers P35832 and P36600), and the MSCA-H2020 Innovative Training Networks SAND and SECRET. S.D. acknowledges Worldwide Cancer Research (Grant 21-0156) and the AIRC Foundation (Investigator Grant 21392). P.R. acknowledges the Veronesi Foundation and the AIRC Foundation (MFAG 27453).

### Declaration of interests

The authors declare no competing interests.

### References

- Shiraishi, K. *et al.* (2023) Biophysical forces mediated by respiration maintain lung alveolar epithelial cell fate. *Cell* 186, 1478–1492
- Wang, J. *et al.* (2023) The role of mechanosensor Piezo1 in bone homeostasis and mechanobiology. *Dev. Biol.* 493, 80–88
- Ragan, P.M. *et al.* (1999) Down-regulation of chondrocyte aggrecan and type-II collagen gene expression correlates with increases in static compression magnitude and duration. *J. Orthop. Res.* 17, 836–842
- Levental, K.R. *et al.* (2009) Matrix crosslinking forces tumor progression by enhancing integrin signaling. *Cell* 139, 891–906
- Lu, P. *et al.* (2012) The extracellular matrix: a dynamic niche in cancer progression. *J. Cell Biol.* 196, 395–406
- Steward, R., Jr *et al.* (2015) Fluid shear, intercellular stress, and endothelial cell alignment. *Am. J. Physiol. Cell Physiol.* 308, C657–C664
- Vining, K.H. and Mooney, D.J. (2017) Mechanical forces direct stem cell behaviour in development and regeneration. *Nat. Rev. Mol. Cell Biol.* 18, 728–742
- Dupont, S. and Wickström, S.A. (2022) Mechanical regulation of chromatin and transcription. *Nat. Rev. Genet.* 23, 624–643
- Romani, P. *et al.* (2021) Crosstalk between mechanotransduction and metabolism. *Nat. Rev. Mol. Cell Biol.* 22, 22–38
- Discher, D.E. *et al.* (2017) Matrix mechanosensing: from scaling concepts in 'omics data to mechanisms in the nucleus, regeneration, and cancer. *Annu. Rev. Biophys.* 46, 295–315
- De Belly, H. *et al.* (2022) Interplay between mechanics and signalling in regulating cell fate. *Nat. Rev. Mol. Cell Biol.* 23, 465–480
- Iskratsch, T. *et al.* (2014) Appreciating force and shape—the rise of mechanotransduction in cell biology. *Nat. Rev. Mol. Cell Biol.* 15, 825–833
- Humphrey, J.D. *et al.* (2014) Mechanotransduction and extracellular matrix homeostasis. *Nat. Rev. Mol. Cell Biol.* 15, 802–812
- Le Roux, A.L. *et al.* (2019) The plasma membrane as a mechanochemical transducer. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 374, 20180221
- Millarte, V. and Farhan, H. (2012) The Golgi in cell migration: regulation by signal transduction and its implications for cancer cell metastasis. *ScientificWorldJournal* 2012, 498278
- Garcia-Manteiga, J.M. *et al.* (2011) Metabolomics of B to plasma cell differentiation. *J. Proteome Res.* 10, 4165–4176
- McCaughy, J. and Stephens, D.J. (2019) ER-to-Golgi transport: a sizeable problem. *Trends Cell Biol.* 29, 940–953
- Miroshnikova, Y.A. and Wickström, S.A. (2022) Mechanical forces in nuclear organization. *Cold Spring Harb. Perspect. Biol.* 14, a039685
- Niethammer, P. (2021) Components and mechanisms of nuclear mechanotransduction. *Annu. Rev. Cell Dev. Biol.* 37, 233–256
- Nader, G.P.F. *et al.* (2021) Nuclear deformations, from signaling to perturbation and damage. *Curr. Opin. Cell Biol.* 72, 137–145
- Janmey, P.A. *et al.* (2020) Stiffness sensing by cells. *Physiol. Rev.* 100, 695–724
- McHugh, B.J. *et al.* (2010) Integrin activation by Fam38A uses a novel mechanism of R-Ras targeting to the endoplasmic reticulum. *J. Cell Sci.* 123, 51–61
- Venditti, R. *et al.* (2021) Regulation and physiology of membrane contact sites. *Curr. Opin. Cell Biol.* 71, 148–157
- Prinz, W.A. *et al.* (2020) The functional universe of membrane contact sites. *Nat. Rev. Mol. Cell Biol.* 21, 7–24
- Holcman, D. *et al.* (2018) Single particle trajectories reveal active endoplasmic reticulum luminal flow. *Nat. Cell Biol.* 20, 1118–1125
- Chambers, J.E. *et al.* (2022) Z- $\alpha$ (1)-antitrypsin polymers impose molecular filtration in the endoplasmic reticulum after undergoing phase transition to a solid state. *Sci. Adv.* 8, eabm2094
- Ravichandran, Y. *et al.* (2020) The Golgi apparatus and cell polarity: roles of the cytoskeleton, the Golgi matrix, and Golgi membranes. *Curr. Opin. Cell Biol.* 62, 104–113
- Egea, G. *et al.* (2006) Actin dynamics at the Golgi complex in mammalian cells. *Curr. Opin. Cell Biol.* 18, 168–178
- Rivero, S. *et al.* (2009) Microtubule nucleation at the cis-side of the Golgi apparatus requires AKAP450 and GM130. *EMBO J.* 28, 1016–1028
- Tonucci, F.M. *et al.* (2015) Centrosomal AKAP350 and CIP4 act in concert to define the polarized localization of the centrosome and Golgi in migratory cells. *J. Cell Sci.* 128, 3277–3289
- Guet, D. *et al.* (2014) Mechanical role of actin dynamics in the rheology of the Golgi complex and in Golgi-associated trafficking events. *Curr. Biol.* 24, 1700–1711
- Antonacci, G. *et al.* (2020) Recent progress and current opinions in Brillouin microscopy for life science applications. *Biophys. Rev.* 12, 615–624
- Goujon, A. *et al.* (2019) Mechanosensitive fluorescent probes to image membrane tension in mitochondria, endoplasmic reticulum, and lysosomes. *J. Am. Chem. Soc.* 141, 3380–3384
- Coste, B. *et al.* (2010) Piezo1 and Piezo2 are essential components of distinct mechanically activated cation channels. *Science* 330, 55–60
- Nava, M.M. *et al.* (2020) Heterochromatin-driven nuclear softening protects the genome against mechanical stress-induced damage. *Cell* 181, 800–817.e822



36. Yang, K. *et al.* (2022) Mechanisms of Pannexin 1 (PANX1) channel mechanosensitivity and its pathological roles. *Int. J. Mol. Sci.* 23, 1523
37. Patil, C.S. *et al.* (2022) ER-resident STIM1/2 couples  $Ca^{2+}$  entry by NMDA receptors to pannexin-1 activation. *Proc. Natl. Acad. Sci. U. S. A.* 119, e2112870119
38. Lee, N.S. *et al.* (2020) Focused ultrasound stimulates ER localized mechanosensitive PANNEXIN-1 to mediate intracellular calcium release in invasive cancer cells. *Front. Cell Dev. Biol.* 8, 504
39. Lomakin, A.J. *et al.* (2020) The nucleus acts as a ruler tailoring cell responses to spatial constraints. *Science* 370, eaba2894
40. Venturini, V. *et al.* (2020) The nucleus measures shape changes for cellular proprioception to control dynamic cell behavior. *Science* 370, eaba2644
41. Enyedi, B. *et al.* (2016) The cell nucleus serves as a mechanotransducer of tissue damage-induced inflammation. *Cell* 165, 1160–1170
42. Janota, C.S. *et al.* (2022) Shielding of actin by the endoplasmic reticulum impacts nuclear positioning. *Nat. Commun.* 13, 2763
43. Kimura, K. *et al.* (2017) Endoplasmic-reticulum-mediated microtubule alignment governs cytoplasmic streaming. *Nat. Cell Biol.* 19, 399–406
44. Phuyal, S. *et al.* (2022) Mechanical strain stimulates COPII-dependent secretory trafficking via Rac1. *EMBO J.* 41, e110596
45. van Vliet, A.R. *et al.* (2017) The ER stress sensor PERK coordinates ER-plasma membrane contact site formation through interaction with Filamin-A and F-actin remodeling. *Mol. Cell* 65, 885–899.e886
46. Lynch, C.D. *et al.* (2011) Filamin depletion blocks endoplasmic spreading and destabilizes force-bearing adhesions. *Mol. Biol. Cell* 22, 1263–1273
47. Urra, H. *et al.* (2018) IRE1 $\alpha$  governs cytoskeleton remodelling and cell migration through a direct interaction with filamin A. *Nat. Cell Biol.* 20, 942–953
48. Kai, F. *et al.* (2022) ECM dimensionality tunes actin tension to modulate endoplasmic reticulum function and spheroid phenotypes of mammary epithelial cells. *EMBO J.* 41, e109205
49. Ko, P. *et al.* (2021) Microtubule acetylation controls MDA-MB-231 breast cancer cell invasion through the modulation of endoplasmic reticulum stress. *Int. J. Mol. Sci.* 22, 6018
50. Chhabra, E.S. *et al.* (2009) INF2 is an endoplasmic reticulum-associated formin protein. *J. Cell Sci.* 122, 1430–1440
51. Chakrabarti, R. *et al.* (2018) INF2-mediated actin polymerization at the ER stimulates mitochondrial calcium uptake, inner membrane constriction, and division. *J. Cell Biol.* 217, 251–268
52. Shao, X. *et al.* (2015) Mechanical stimulation induces formin-dependent assembly of a perinuclear actin rim. *Proc. Natl. Acad. Sci. U. S. A.* 112, E2595–E2601
53. Pothapragada, S.P. *et al.* (2022) Matrix mechanics regulates epithelial defence against cancer by tuning dynamic localization of filamin. *Nat. Commun.* 13, 218
54. Jung, J. *et al.* (2022) Regulation of the COPII secretory machinery via focal adhesions and extracellular matrix signaling. *J. Cell Biol.* 221, e202110081
55. Farhan, H. *et al.* (2010) MAPK signaling to the early secretory pathway revealed by kinase/phosphatase functional screening. *J. Cell Biol.* 189, 997–1011
56. Simpson, J.C. *et al.* (2012) Genome-wide RNAi screening identifies human proteins with a regulatory function in the early secretory pathway. *Nat. Cell Biol.* 14, 764–774
57. Szafranski, J.D. *et al.* (2004) Chondrocyte mechanotransduction: effects of compression on deformation of intracellular organelles and relevance to cellular biosynthesis. *Osteoarthritis. Cartil.* 12, 937–946
58. Yadav, S. and Linstedt, A.D. (2011) Golgi positioning. *Cold Spring Harb. Perspect. Biol.* 3, a005322
59. Khuntia, P. *et al.* (2022) Actin-driven Golgi apparatus dispersal during collective migration of epithelial cells. *Proc. Natl. Acad. Sci. U. S. A.* 119, e2204808119
60. Fourniere, L. *et al.* (2019) RAB6 and microtubules restrict protein secretion to focal adhesions. *J. Cell Biol.* 218, 2215–2231
61. Romani, P. *et al.* (2019) Extracellular matrix mechanical cues regulate lipid metabolism through Lipin-1 and SREBP. *Nat. Cell Biol.* 21, 338–347
62. Bertolio, R. *et al.* (2019) Sterol regulatory element binding protein 1 couples mechanical cues and lipid metabolism. *Nat. Commun.* 10, 1326
63. Singh, V. *et al.* (2018) Cell-matrix adhesion controls Golgi organization and function through Arf1 activation in anchorage-dependent cells. *J. Cell Sci.* 131, jcs215855
64. Joseph, J.G. and Liu, A.P. (2020) Mechanical regulation of endocytosis: new insights and recent advances. *Adv. Biosyst.* 4, e1900278
65. Wang, G. *et al.* (2018) Biomechanical control of lysosomal secretion via the VAMP7 hub: a tug-of-war between VARP and LRRK1. *iScience* 4, 127–143
66. Block, M.R. *et al.* (2020) The mechano-sensitive response of  $\beta$ 1 integrin promotes SRC-positive late endosome recycling and activation of Yes-associated protein. *J. Biol. Chem.* 295, 13474–13487
67. Chen, C.C. *et al.* (2020) TRPML2 is an osmo/mechanosensitive cation channel in endolysosomal organelles. *Sci. Adv.* 6, eabb5064
68. Thottacherry, J.J. *et al.* (2018) Mechanochemical feedback control of dynamin independent endocytosis modulates membrane tension in adherent cells. *Nat. Commun.* 9, 4217
69. Sinha, B. *et al.* (2011) Cells respond to mechanical stress by rapid disassembly of caveolae. *Cell* 144, 402–413
70. Del Pozo, M.A. *et al.* (2021) Caveolae: mechanosensing and mechanotransduction devices linking membrane trafficking to mechanoadaptation. *Curr. Opin. Cell Biol.* 68, 113–123
71. Palamidessi, A. *et al.* (2019) Unjamming overcomes kinetic and proliferation arrest in terminally differentiated cells and promotes collective motility of carcinoma. *Nat. Mater.* 18, 1252–1263
72. Giacomello, M. *et al.* (2020) The cell biology of mitochondrial membrane dynamics. *Nat. Rev. Mol. Cell Biol.* 21, 204–224
73. Kraus, F. and Ryan, M.T. (2017) The constriction and scission machineries involved in mitochondrial fission. *J. Cell Sci.* 130, 2953–2960
74. Kalia, R. *et al.* (2018) Structural basis of mitochondrial receptor binding and constriction by DRP1. *Nature* 558, 401–405
75. Mahecic, D. *et al.* (2021) Mitochondrial membrane tension governs fission. *Cell Rep.* 35, 108947
76. Khalifat, N. *et al.* (2008) Membrane deformation under local pH gradient: mimicking mitochondrial cristae dynamics. *Biophys. J.* 95, 4924–4933
77. Beltrán-Heredia, E. *et al.* (2019) Membrane curvature induces cardiolipin sorting. *Commun. Biol.* 2, 225
78. Korobova, F. *et al.* (2013) An actin-dependent step in mitochondrial fission mediated by the ER-associated formin INF2. *Science* 339, 464–467
79. De Vos, K.J. *et al.* (2005) Mitochondrial function and actin regulate dynamin-related protein 1-dependent mitochondrial fission. *Curr. Biol.* 15, 678–683
80. Manor, U. *et al.* (2015) A mitochondria-anchored isoform of the actin-nucleating spire protein regulates mitochondrial division. *eLife* 4, e08828
81. Osellame, L.D. *et al.* (2016) Cooperative and independent roles of the Drp1 adaptors Mif, MiD49 and MiD51 in mitochondrial fission. *J. Cell Sci.* 129, 2170–2181
82. Miyamoto, T. *et al.* (2021) Rapid manipulation of mitochondrial morphology in a living cell with iCMM. *Cell Rep. Methods* 1, 100052
83. Shimura, D. *et al.* (2021) Protective mitochondrial fission induced by stress-responsive protein GJA1-20k. *eLife* 10, e69207
84. Helle, S.C.J. *et al.* (2017) Mechanical force induces mitochondrial fission. *eLife* 6, e30292
85. Moore, A.S. *et al.* (2016) Dynamic actin cycling through mitochondrial subpopulations locally regulates the fission–fusion balance within mitochondrial networks. *Nat. Commun.* 7, 12886
86. Tharp, K.M. *et al.* (2021) Adhesion-mediated mechanosignaling forces mitohormesis. *Cell Metab.* 33, 1322–1341.e1313
87. Papalazarou, V. *et al.* (2020) The creatine-phosphagen system is mechanoresponsive in pancreatic adenocarcinoma and fuels invasion and metastasis. *Nat. Metab.* 2, 62–80
88. Guo, T. *et al.* (2023) Mitochondrial fission and bioenergetics mediate human lung fibroblast durotaxis. *JCI Insight* 8, e157348

89. Wang, W. *et al.* (2012) Mitochondrial fission triggered by hyperglycemia is mediated by ROCK1 activation in podocytes and endothelial cells. *Cell Metab.* 15, 186–200
90. Hong, S.G. *et al.* (2022) Flow pattern-dependent mitochondrial dynamics regulates the metabolic profile and inflammatory state of endothelial cells. *JCI Insight* 7, e159286
91. Coon, B.G. *et al.* (2022) A mitochondrial contribution to anti-inflammatory shear stress signaling in vascular endothelial cells. *J. Cell Biol.* 221, e202109144
92. Romani, P. *et al.* (2022) Mitochondrial fission links ECM mechanotransduction to metabolic redox homeostasis and metastatic chemotherapy resistance. *Nat. Cell Biol.* 24, 168–180
93. Guo, L. *et al.* (2020) PINCH-1 regulates mitochondrial dynamics to promote proline synthesis and tumor growth. *Nat. Commun.* 11, 4913
94. Bartolák-Suki, E. *et al.* (2015) Fluctuation-driven mechanotransduction regulates mitochondrial-network structure and function. *Nat. Mater.* 14, 1049–1057
95. Arbore, C. *et al.* (2019) Probing force in living cells with optical tweezers: from single-molecule mechanics to cell mechanotransduction. *Biophys. Rev.* 11, 765–782
96. Nunes Vicente, F. *et al.* (2023) Novel imaging methods and force probes for molecular mechanobiology of cytoskeleton and adhesion. *Trends Cell Biol.* 33, 204–220
97. McKenna, M.J. and Shao, S. (2023) The endoplasmic reticulum and the fidelity of nascent protein localization. *Cold Spring Harb. Perspect. Biol.* 15, a041249
98. Zeuschner, D. *et al.* (2006) Immuno-electron tomography of ER exit sites reveals the existence of free COP1-coated transport carriers. *Nat. Cell Biol.* 8, 377–383
99. Weigel, A.V. *et al.* (2021) ER-to-Golgi protein delivery through an interwoven, tubular network extending from ER. *Cell* 184, 2412–2429.e16
100. Phuyal, S. and Farhan, H. (2021) Want to leave the ER? We offer vesicles, tubules, and tunnels. *J. Cell Biol.* 220, e202104062
101. Farhan, H. *et al.* (2008) Adaptation of endoplasmic reticulum exit sites to acute and chronic increases in cargo load. *EMBO J.* 27, 2043–2054
102. Zacharogianni, M. *et al.* (2011) ERK7 is a negative regulator of protein secretion in response to amino-acid starvation by modulating Sec16 membrane association. *EMBO J.* 30, 3684–3700
103. Lippincott-Schwartz, J. and Phair, R.D. (2010) Lipids and cholesterol as regulators of traffic in the endomembrane system. *Annu. Rev. Biophys.* 39, 559–578
104. Sorre, B. *et al.* (2009) Curvature-driven lipid sorting needs proximity to a demixing point and is aided by proteins. *Proc. Natl. Acad. Sci. U. S. A.* 106, 5622–5626
105. Almanza, A. *et al.* (2019) Endoplasmic reticulum stress signaling – from basic mechanisms to clinical applications. *FEBS J.* 286, 241–278
106. Rambold, A.S. *et al.* (2011) Tubular network formation protects mitochondria from autophagosomal degradation during nutrient starvation. *Proc. Natl. Acad. Sci. U. S. A.* 108, 10190–10195
107. Gomes, L.C. *et al.* (2011) During autophagy mitochondria elongate, are spared from degradation and sustain cell viability. *Nat. Cell Biol.* 13, 589–598
108. Mitra, K. *et al.* (2009) A hyperfused mitochondrial state achieved at G<sub>1</sub>-S regulates cyclin E buildup and entry into S phase. *Proc. Natl. Acad. Sci. U. S. A.* 106, 11960–11965
109. König, T. *et al.* (2021) MIROs and DRP1 drive mitochondrial-derived vesicle biogenesis and promote quality control. *Nat. Cell Biol.* 23, 1271–1286
110. MacVicar, T.D. and Lane, J.D. (2014) Impaired OMA1-dependent cleavage of OPA1 and reduced DRP1 fission activity combine to prevent mitophagy in cells that are dependent on oxidative phosphorylation. *J. Cell Sci.* 127, 2313–2325
111. Aon, M.A. *et al.* (2004) Percolation and criticality in a mitochondrial network. *Proc. Natl. Acad. Sci. U. S. A.* 101, 4447–4452
112. Zamponi, N. *et al.* (2018) Mitochondrial network complexity emerges from fission/fusion dynamics. *Sci. Rep.* 8, 363
113. Zamponi, N. *et al.* (2022) Universal dynamics of mitochondrial networks: a finite-size scaling analysis. *Sci. Rep.* 12, 17074



# Mechanics and functional consequences of nuclear deformations

Yohalie Kalukula<sup>1</sup>, Andrew D. Stephens<sup>2</sup>, Jan Lammerding<sup>1,3,4</sup> and Sylvain Gabriele<sup>1</sup>

**Abstract** | As the home of cellular genetic information, the nucleus has a critical role in determining cell fate and function in response to various signals and stimuli. In addition to biochemical inputs, the nucleus is constantly exposed to intrinsic and extrinsic mechanical forces that trigger dynamic changes in nuclear structure and morphology. Emerging data suggest that the physical deformation of the nucleus modulates many cellular and nuclear functions. These functions have long been considered to be downstream of cytoplasmic signalling pathways and dictated by gene expression. In this Review, we discuss an emerging perspective on the mechanoregulation of the nucleus that considers the physical connections from chromatin to nuclear lamina and cytoskeletal filaments as a single mechanical unit. We describe key mechanisms of nuclear deformations in time and space and provide a critical review of the structural and functional adaptive responses of the nucleus to deformations. We then consider the contribution of nuclear deformations to the regulation of important cellular functions, including muscle contraction, cell migration and human disease pathogenesis. Collectively, these emerging insights shed new light on the dynamics of nuclear deformations and their roles in cellular mechanobiology.

As the largest and stiffest organelle of eukaryotic cells<sup>1</sup>, the nucleus is constantly subjected to intrinsic and extrinsic forces that can lead to small and large nuclear deformations. For example, cytoskeletal forces position the nucleus within polarized cells, and actomyosin forces are required to squeeze the nucleus of migrating cells through small constrictions such as interstitial spaces. Accumulating evidence suggests that the nucleus contributes to cellular perception of mechanical stimuli and the corresponding cellular response through dynamic changes of its structure and morphology<sup>2,3</sup>. Therefore, the nucleus must be considered not only as the primary site of gene replication and transcription but also as a fundamental mechanotransduction component of the cell, capable of mechanosensing and orchestrating key cellular functions in response to mechanical stimulation.

The mechanotransduction properties of the nucleus are now well recognized, including its ability to adapt to the physical microenvironment of the cell with changes in nuclear morphology or the expression of specific genes<sup>4,5</sup>. By contrast, the role of the nucleus as a mechanosensitive organelle — whereby physical deformations induced by forces transmitted to the nuclear envelope directly impact nuclear and cellular functions — has only recently begun to emerge (BOX 1). For example, several lines of evidence indicate that forces acting on the nucleus can induce sufficient nuclear deformations to modulate chromatin structure and trigger important

protein conformational changes, thereby activating or repressing mechanoresponsive genes<sup>6,7</sup>. In vivo, the impact of nuclear deformations has been highlighted by the observation that many human diseases are associated with abnormal nuclear shapes<sup>8</sup> and disturbed mechanotransduction processes<sup>9</sup> such as impaired activation of genes in response to mechanical stimulation or mechanically induced DNA damage (BOX 2).

In this Review, we discuss the current understanding of the physical properties of the nucleus, and how the different nuclear components affect its mechanics. We then review the physiological contexts of nuclear deformations and highlight the importance of physical connections between the nuclear envelope and the cytoskeleton in the transmission of forces to the nucleus and driving its deformations. We also consider the emerging role of nuclear deformations in cellular mechanosensing and mechanotransduction.

## Nuclear organization

The extent of nuclear deformations is determined by the balance between the mechanical properties of the nucleus and the mechanical forces acting on it. Nuclear mechanical properties are dependent on the various components constituting the nuclear structure. The forces acting on the nucleus are primarily derived from the cytoskeleton, which establishes physical connections with the nuclear envelope (FIG. 1), although some forces can also originate from the outside of the cell.

<sup>1</sup>University of Mons, Soft Matter & Biomaterials Group, Interfaces and Complex Fluids Laboratory, Research Institute for Biosciences, CIRMAP, Mons, Belgium.

<sup>2</sup>Biology Department, University of Massachusetts Amherst, Amherst, MA, USA.

<sup>3</sup>Weill Institute for Cell and Molecular Biology, Cornell University, Ithaca, NY, USA.

<sup>4</sup>Nancy E. and Peter C. Meinig School of Biomedical Engineering, Cornell University, Ithaca, NY, USA.

✉e-mail:

jan.lammerding@cornell.edu;  
sylvain.gabriele@umons.ac.be  
<https://doi.org/10.1038/s41580-022-00480-z>

## Box 1 | Nuclear mechanosensing

Although it is now well recognized that nuclear deformations have both rapid and long-lasting consequences on nuclear and cellular function, the precise mechanisms by which nuclear deformations are translated into biochemical signals, and to what degree the nucleus itself serves as a cellular mechanosensor, remain incompletely understood. As a note of caution, many nuclear changes in response to external mechanical stimuli (for example, altered nuclear shape, chromatin organization, gene expression) cited as indicators of nuclear mechanosensing may reflect, at least in part, downstream effects of signalling pathways initiated in the cytoplasm or cell surface, rather than direct nuclear mechanosensing. In the following, we highlight recent findings and novel insights into established and proposed nuclear mechanosensing mechanisms. For a more detailed discussion, we refer the reader to some excellent recent reviews<sup>32,194,218,219</sup>.

**Stretch-activated opening of channels in the nuclear membranes**

Nuclear pore complexes allow passage of small molecules while excluding larger molecules that do not contain nuclear localization sequences or are transported by other proteins. Recent live cell imaging, electron microscopy and cryo-electron tomography studies found that nuclear pore complexes are highly sensitive to nuclear membrane tension<sup>15,198,199</sup>, increasing their diameter in response to elevated nuclear membrane tension and thus facilitating nuclear import, including of the mechanoresponsive transcription factor<sup>198</sup>. The nuclear envelope and endoplasmic reticulum (ER) membranes (which are continuous with the nuclear envelope) contain various other stretch-sensitive ion channels such as Piezo1 and inositol triphosphate receptor (InsP3R). Increased nuclear membrane tension, in response to cell compression, osmotic swelling or stretching application, may trigger opening of these channels and calcium release from the ER and perinuclear space, which can lead to increased cell contractility<sup>17,127</sup> as well as to the uptake of calcium into the nucleus, resulting in changes in chromatin organization and nuclear softening driven by loss of heterochromatin<sup>16</sup>. However, it remains unclear whether opening of these ion channels in response to cellular deformation occurs at the nuclear envelope, ER or the plasma membrane. One interesting hypothesis is that all three locations contribute to cellular mechanotransduction, depending on the context. As such, spatial coordination between ion channels in the different membranes would allow cells to distinguish between different sources of nuclear membrane strain such as osmotic swelling and compression<sup>3,127</sup>.

**Mechanosensing by the nuclear membranes and nuclear envelope proteins**

Changes in the tension or curvature of the nuclear membranes can alter the packing and/or composition of nuclear membrane phospholipids, which, together with increased intranuclear calcium concentrations, promote binding of nucleoplasmic phospholipase A2 (cPLA2) to the inner nuclear membrane<sup>192–194</sup>, where it can initiate cell signalling events related to actomyosin contractility and inflammation.

Besides altering protein interactions with the nuclear membranes, forces acting on the nucleus can also lead to local unfolding, conformational changes and increased phosphorylation of lamins<sup>105,109,220–222</sup>, although the functional relevance of these changes remains to be fully characterized. Furthermore, force application to the nucleus via nesprins leads to phosphorylation of emerin via Src kinases, resulting in the recruitment of lamins to the nuclear envelope and nuclear stiffening<sup>223</sup>. Although it remains unclear whether the increased phosphorylation is due to mechanically induced activation of nuclear Src kinase or emerin becoming more accessible to the kinase, this study, which was conducted on isolated nuclei, provided some of the most direct evidence for nuclear mechanosensing.

**Force-induced changes in chromatin organization**

Several studies have demonstrated mechanically induced changes in chromatin organization that could affect gene expression, including in neutrophils that had migrated through tight constrictions<sup>208</sup>, macrophages under spatial confinement<sup>179</sup> and a 3D chemo-mechanical model of the nuclear interior and its connections to the cytoskeleton. However, these studies did not completely address whether the effects were nucleus-intrinsic or mediated by cytoplasmic signals. Support for direct involvement of chromatin remodelling in nuclear mechanosensing comes from two recent studies, which found that force application to the cell surface leads to near instantaneous chromatin deformation, visualized by tracking multiple GFP–LacI-labelled genomic loci, and rapid (<15 s) increase in transcription of the corresponding transgene and other genes<sup>204,205</sup>. The magnitude of the response was directly related to the extent of chromatin deformation and histone methylation status. Of note, the chromatin ‘stretching’ reported in these studies likely does not reflect stretching of the DNA itself but rather partial unpacking of the chromatin, which may promote access to transcriptional regulators or polymerases<sup>205</sup>. Depletion of lamins, emerin or linker of nucleoskeleton and cytoskeleton (LINC) complex components abolished the force-induced gene expression<sup>204</sup>, pointing to the importance of nucleo-cytoskeletal coupling in nuclear mechanosensing. The effect of LINC complex disruption on the activation of mechanoresponsive genes contrasts with a previous study in which LINC complex disruption did not alter the expression of several mechanoresponsive genes despite reducing nuclear deformation<sup>36</sup>, possibly reflecting differences in cell type, the mode of force application or the extent/type of nuclear deformation resulting from the applied force.

Another intriguing thought is that liquid–liquid phase separation, which is a central player in the assembly of membraneless compartments within the nucleus, could contribute to nuclear mechanosensing. Indeed, significant mechanical forces through attractive and repulsive interactions between protein droplets and chromatin can alter chromatin organization and rearrangements<sup>34,90</sup>. One could therefore speculate that externally applied forces and resulting nuclear deformation could affect intranuclear biomolecular condensates, which are highly dynamic structures that may condense or dissolve under specific nuclear deformations, and thereby regulate nuclear functions.

**The nuclear envelope.** The nuclear envelope serves multiple pivotal functions: it controls access of cytoplasmic proteins to the genome, provides structural stability to the nucleus, and physically connects the nuclear interior and cytoskeleton (FIG. 1; see next sub-section). The nuclear

envelope comprises nuclear membranes, the nuclear lamina and nuclear pore complexes (NPCs). The inner and outer nuclear membranes (INM and ONM, respectively) are two concentric lipid bilayers, each ~4 nm thick, separated by the ~20–50 nm-wide perinuclear space<sup>10</sup> (FIG. 1a).

**Mechanotransduction**

In its literal sense, mechanotransduction refers to the molecular process in which mechanical stimuli are converted (or transduced) into biochemical signals, that is, equivalent to the ‘mechanosensing’ defined below. However, mechanotransduction is commonly used to more broadly refer to cellular responses to changes in the mechanical environment, including forces, deformations or mechanical properties. In this article, we use this broader definition of mechanotransduction.

**Mechanosensing**

Molecular process through which cells or cellular components translate mechanical forces or deformations into biochemical signals.



**Stress**

Expression of the mechanical loading in terms of force applied per cross-sectional area of an object. Units of stress are  $\text{N m}^{-2}$  (or Pa).

**Rhabdomyosarcoma**

Highly aggressive form of cancer mostly observed in children and adolescents that usually develops in soft tissues, such as the muscles, from mesenchymal cells that have failed to fully differentiate.

**Segmental premature ageing disease**

Pathological condition that reflects some but not other phenotypes of the normal ageing process at a much earlier age. For example, children with Hutchinson–Gilford progeria syndrome develop severe cardiovascular disease (heart attacks and strokes) in their early teens but lack neurodegenerative defects such as dementia and are not more prone to cancer.

The ONM is contiguous with the endoplasmic reticulum (ER) and can expand by the addition of lipids from the ER, allowing the nuclear surface area to adapt in response to deformation (although membrane recruitment to the nuclear envelope may be limited by physical resistance from the ER). Furthermore, the nuclear membrane is wrinkled and folded at low tension, which provides an additional membrane reservoir for adjusting

nuclear shape<sup>11</sup>. NPCs are homogeneously distributed over the nuclear membrane surface<sup>12</sup> and regulate the active nuclear transport of macromolecules larger than ~50 kDa into and out of the nucleus<sup>12,13</sup>. The size of the NPCs can change in response to mechanical stress, which accounts for up to 10% of nuclear surface expansion during nuclear deformations<sup>14–16</sup>. The nuclear envelope and ER additionally contain mechanosensitive ion channels

**Box 2 | Human pathologies associated with nuclear deformations**

Abnormalities in nuclear and chromatin organization are hallmarks of many diseases, ranging from heart disease to premature ageing and cancer<sup>224</sup>, where they can indicate, for example, metastatic potential<sup>8,225,226</sup>. Hundreds of mutations and variants have been found in genes encoding nuclear envelope components, including inner and outer nuclear membrane proteins (for example, nesprins, emerin and SUN (Sad1p, UNC-84) proteins) and lamins, and the diseases resulting from these mutations<sup>227</sup> are collectively called nuclear envelopopathies. Mutations in the *LMNA* gene, which encodes lamins A/C, alone cause over 13 human diseases, including congenital dilated cardiomyopathy<sup>228,229</sup>, various types of muscular dystrophy<sup>230</sup> and progeria<sup>231</sup>, with altered nuclear stability and mechanotransduction thought to contribute, at least in part, to the disease mechanism.

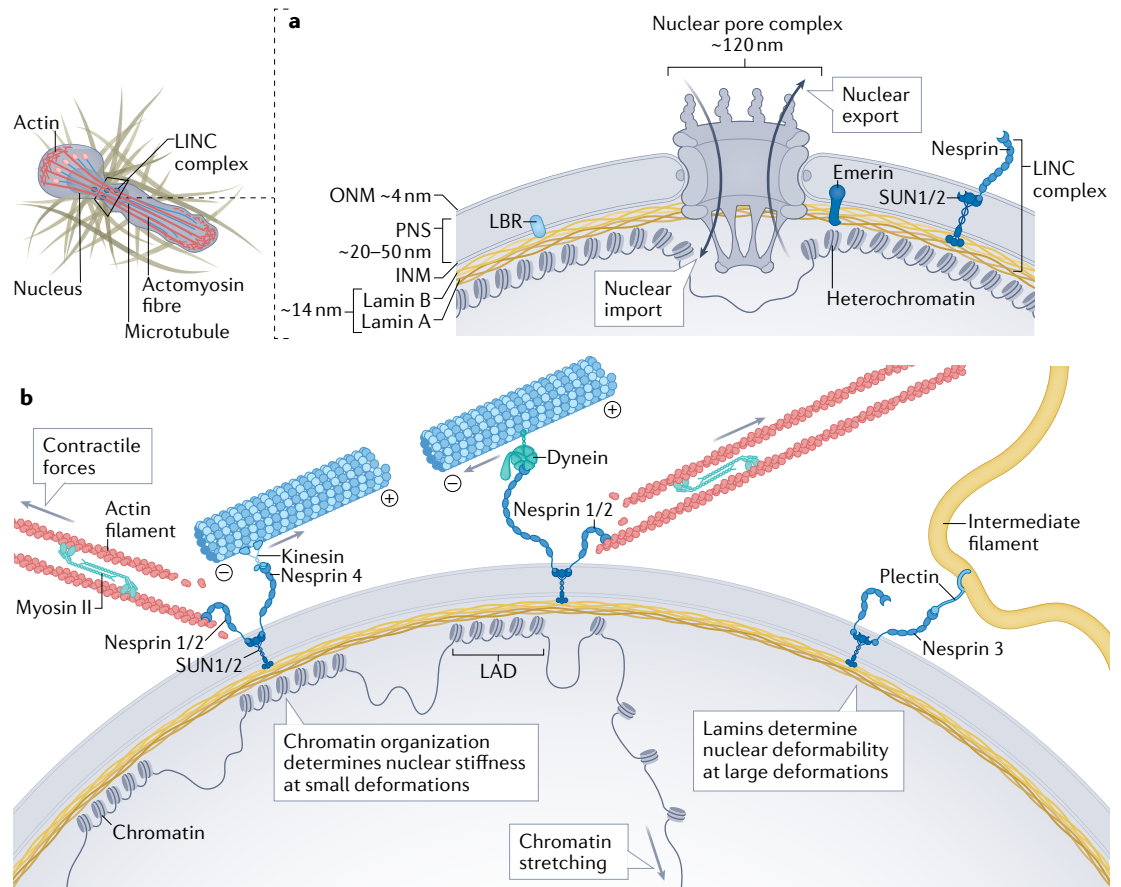
*LMNA* mutations associated with muscular dystrophy and dilated cardiomyopathy often result in more deformable and more fragile nuclei<sup>59</sup>. This leads to extensive nuclear envelope damage in skeletal muscle cells in vitro and in vivo, resulting from mechanical stress on the more fragile nuclei<sup>119</sup>. Lamin mutations associated with muscular dystrophy can also impair linker of nucleoskeleton and cytoskeleton (LINC) complex function<sup>55,232,233</sup> and other cellular processes. Furthermore, abnormal YAP (Yes-associated protein) activity, known to be responsive to nuclear deformation and lamin A levels<sup>27,198</sup>, has been reported in muscular dystrophy and rhabdomyosarcoma<sup>234</sup>. In *LMNA*-related congenital muscular dystrophy, lamin mutations increase YAP nuclear localization via increased nuclear import, implicating YAP as a potential pathogenic contributor in muscular dystrophies caused by nuclear envelope defects<sup>235</sup>.

Hutchinson–Gilford progeria syndrome (HGPS) is an exceptionally rare and severe segmental premature ageing disease caused by mutations in the *LMNA* gene. Most cases of HGPS result from a mutation that leads to alternative splicing, causing a truncated form of prelamin A (LA $\Delta$ 50) that remains farnesylated. Cells from patients with HGPS have irregular nuclear shapes<sup>236</sup>, increased nuclear stiffness and increased sensitivity to mechanical stress<sup>237–240</sup>, which may be responsible for the progressive loss of vascular smooth muscle cells in HGPS. The structural abnormalities of the mutant lamins and their stronger interaction with other lamins reduces the ability of the nuclear envelope to dissipate mechanical stress<sup>240</sup>. In addition to perturbing nuclear lamina organization, the mutant lamins also alter chromatin organization. Restoring the loss of heterochromatin alone in HeLa cells expressing LA $\Delta$ 50 and in cells from patients with HGPS is sufficient to restore normal nuclear shape, suggesting that heterochromatin loss may be responsible for many of the phenomena associated with HGPS<sup>64,91,241</sup>.

Deficiency of lamin B1 and lamin B2, but also increased expression of lamin B1, are associated with neurodevelopmental defects and distinct nuclear shape abnormalities in neurons<sup>242,243</sup>. Loss of B-type lamins interferes with proper nucleokinesis, a nuclear translation process required during neuronal development<sup>73</sup>. Lamin B1-deficient and lamin B2-deficient mouse embryos have defective migration of cortical neurons<sup>242–244</sup>, leading to neuronal layering abnormality in the cerebral cortex along with neonatal mortality. The neuronal migration abnormality may be explained by a weakened nuclear lamina (in particular as developing neural tissue lacks expression of A-type lamins) as preliminary work shows that B-type lamin depletion may affect nuclear mechanical properties<sup>245</sup>. Duplication of the gene encoding lamin B1 results in autosomal dominant leukodystrophy, which is characterized by widespread loss of myelin in the central nervous system<sup>246</sup>, although the molecular mechanisms underlying these defects remain unclear.

In addition to mutations in nuclear envelope proteins, cytoplasmic proteins can also result in nuclear defects and diseases. Tauopathies refer to a class of neurodegenerative diseases involving the aggregation of Tau protein, a neuronal microtubule-associated protein, into neurofibrillary or gliofibrillary tangles in the brain. Pathological accumulation of Tau, known as Tau nuclear rods or Tau-positive nuclear indentations<sup>247</sup>, have been identified in several neurodegenerative disorders, including Alzheimer disease, frontotemporal dementia and Huntington disease<sup>248,249</sup>. However, the mechanism underlying Tau-mediated pathogenesis is still unclear. Mutations in the Tau-encoding gene *MAPT* result in Tau mislocalization to the cell bodies rather than to the neuronal axon. This leads to abnormal microtubule organization, which can potentially deform the nuclear envelope via LINC complex-based coupling<sup>250</sup>, causing large nuclear lamin invaginations and defects in nucleocytoplasmic trafficking<sup>251,252</sup>.

Although the pathological mechanisms underlying the diverse envelopopathies are still not fully understood, various hypotheses have been put forward. The key role of the nuclear envelope in regulating nuclear mechanics and mechanotransduction suggests that defects in nuclear envelope/lamina proteins can result — directly (by changing nuclear physical properties) or indirectly (for example, via changes in chromatin organization or nucleo-cytoskeleton coupling) — in impaired nuclear stability, increased nuclear fragility and perturbations of mechanotransduction pathways, which could explain some of the tissue-specific phenotypes. This hypothesis is supported by numerous in vitro and in vivo observations of abnormalities in nuclear morphology (for example, wrinkling, irregularities, blebs and invaginations) associated with *LMNA* mutations linked to dilated cardiomyopathy, muscular dystrophy and HGPS as well as the increased DNA damage found in several laminopathies<sup>26,187,253</sup>. Besides their mechanical function, lamins have a key role in tethering and organizing chromatin as well as in signalling involved in transcriptional regulation. In support of this, laminopathic nuclei often display alterations in the organization of chromatin and signalling as well as broad alterations in gene expression<sup>7,254–258</sup>, which could contribute to tissue-specific phenotypes.



**Fig. 1 | The nuclear envelope and nucleo-skeletal interactions. a** | The nuclear envelope is composed of the inner (INM) and outer (ONM) nuclear membranes, which form a double lipid bilayer separated by the perinuclear space (PNS), and the proteinaceous nuclear lamina, which is attached to the INM and in close contact with condensed chromatin. The nuclear lamina meshwork is composed of A-type and B-type lamins. Nuclear pore complexes span the nuclear envelope and are surrounded by less condensed chromatin, and allow controlled nuclear import and export of large molecules. Lamins, along with other INM proteins, such as the lamin B receptor (LBR) and emerin, anchor chromatin to the nuclear envelope. Nesprins, ONM, SUN (Sad1p, UNC-84) domain proteins and INM together form the linker of nucleoskeleton and cytoskeleton (LINC) complex. The LINC complex provides a direct physical connection between cytoskeletal filaments and the nuclear interior, which allows the transfer of extracellular and cytoskeletal forces to the nucleus. **b** | The nuclear interior is connected to cytoskeletal filaments by nesprins and SUN domain proteins (SUN1/2). Nesprin 1 and nesprin 2 bind to actin filaments, whereas nesprin 3 interacts with intermediate filaments. Nesprins 1, 2 and 4 can interact with microtubules via kinesin and dynein molecular motors. Interactions between molecular motors and cytoskeletal filaments generate forces that are directly transmitted to the nucleus through LINC complexes. The genomic regions connected to the lamina are lamina-associated chromatin domains (LADs), which have low transcriptional activity.

such as Piezo1 (REF.<sup>16</sup>) and inositol triphosphate receptors (InsP3Rs)<sup>17</sup> that can respond to nuclear membrane tension (BOX 1). The nuclear lamina, a dense protein network underlying the INM, is primarily composed of lamins, a family of nuclear intermediate filaments. Lamins assemble into 300–400 nm-long and ~3.5 nm-thick nonpolar filaments, and form a ~14–30 nm-thick meshwork<sup>18,19</sup>. In mammalian somatic cells, the nuclear lamina is predominantly composed of four lamin isoforms: two A-type lamins (A and C) and two B-type lamins (B1 and B2)<sup>20</sup>. The *LMNA* gene encodes lamin A and lamin C and some rare isoforms that arise from alternative splicing, and the *LMNB1* and *LMNB2* genes encode lamin B1 and lamin B2, respectively<sup>20</sup>. Each lamin isoform forms separate but interacting meshworks<sup>21,22</sup>. B-type lamins are permanently modified by farnesylation and are thus primarily located at the nuclear membranes (FIG. 1a), whereas

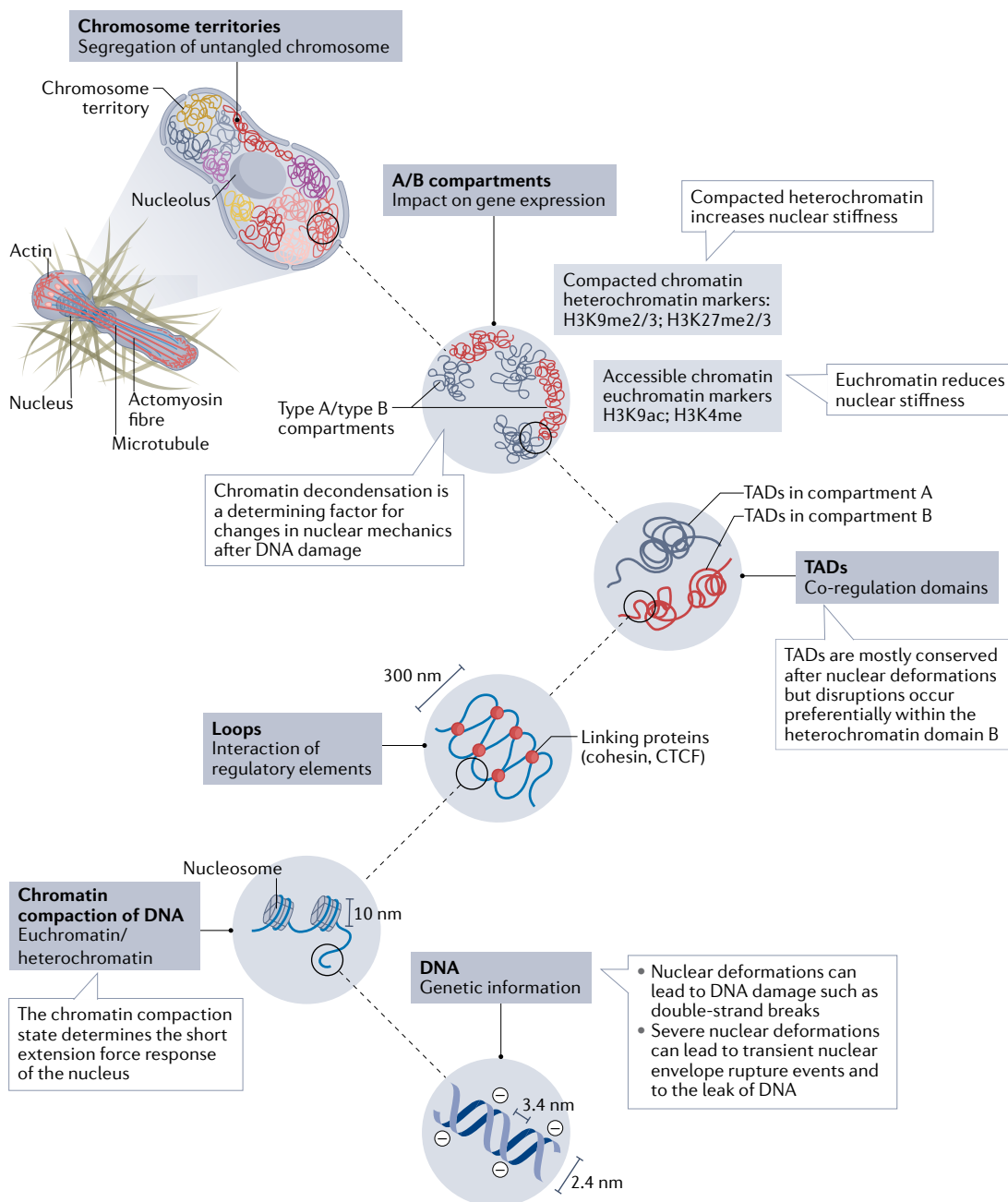
A-type lamins either lack farnesylation sites completely (lamin C) or have their farnesylated C terminus removed post-transcriptionally (lamin A) and can be localized at both the nuclear lamina and the nuclear interior<sup>23</sup>, with the intranuclear distribution of lamins mediated by lamina-associated polypeptide 2α (LAP2α) and other proteins<sup>24</sup>. Lamins interact with various binding partners, including NPC proteins, INM proteins, chromatin and various transcription regulators<sup>20</sup>. Accordingly, the lamina has many structural and other functions, including contributing to nuclear shape, mechanical stability, nucleo-cytoskeletal coupling, nuclear positioning, genome organization and mechanosensing<sup>25–27</sup>.

**The nuclear interior.** The nuclear interior primarily consists of chromatin and nuclear bodies such as nucleoli, Cajal bodies and promyelocytic leukaemia bodies, which

**Intermediate filaments**  
Large family of nuclear and cytoskeletal filaments that includes keratins (types I and II), desmin and vimentin (type III), neurofilaments (type IV) and lamins (type V). Intermediate filaments form dimers that then assemble into larger nonpolar filament structures that are characterized by their ability to extend substantially under mechanical stress.

are membraneless structures with specific signalling and processing functions<sup>28</sup>. Chromatin is composed of DNA and DNA-binding proteins, particularly histones

(FIG. 2). Chromatin can be classified into two categories, depending on its level of compaction, transcriptional activity and histone modifications. The loosely packed



**Fig. 2 | Chromatin organization and its impact on nuclear mechanics.** The figure illustrates the different levels of chromatin organization, from bottom to top. Chromosomal DNA is packaged inside the cell nucleus with the help of histones. At the simplest level, chromatin is a double-stranded helical structure of DNA. The negatively charged DNA double helix is complexed with histones, which are positively charged proteins, to form nucleosomes. Inside the interphase nucleus, chromosomes occupy distinct territories (highlighted by different colours). Within each chromosome territory, the chromatin is folded into multiple loops and segregated into two distinct compartments: compartment A, clustered around the nucleolus and nuclear bodies (permissive region), and compartment B (repressive region), located at the nuclear periphery. Within compartments, chromatin is further partitioned into topologically associating domains (TADs), which have preferential intradomain interactions compared with interdomain interactions with the neighbouring *cis* chromatin domains. Histone methylation, particularly at residues H3K9 and H3K27, is often associated with heterochromatin, whereas histone acetylation, particularly at residue H3K9, or histone methylation at residue H3K4, is typically associated with euchromatin. In addition to lamins, chromatin is a major mechanical component that determines nuclear size and stiffness. Chromatin is particularly important in resisting small nuclear deformations, whereas lamins dominate for large nuclear deformations. Chromatin modifications associated with euchromatin typically lead to reduced nuclear stiffness, while chromatin modifications associated with the more compacted heterochromatin increase nuclear stiffness.

**Farnesylation**

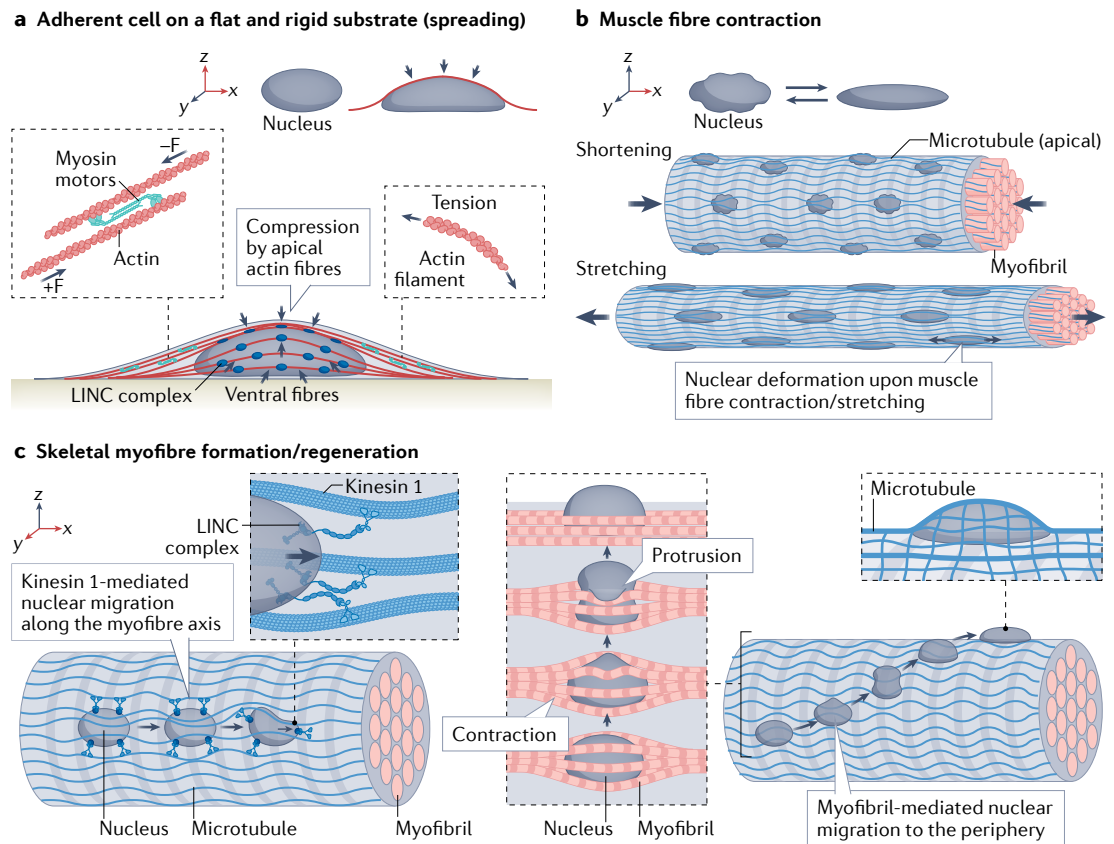
Post-translational modification of proteins catalysed by the enzyme farnesyltransferase, which adds a 15-carbon unsaturated hydrocarbon chain to a cysteine residue via a thioether linkage, thus anchoring the protein to a lipid membrane.

**Lamina-associated polypeptide 2α**

One of six alternatively spliced isoforms of the mammalian *LAP2* gene that is functionally and structurally different. *LAP2α* shares only the  $NH_2$  terminus with the other isoforms and contains a unique  $COOH$  terminus. *LAP2α* is localized throughout the nucleus and is a specific binding partner of nucleoplasmic A-type lamins.

euchromatin is transcriptionally accessible and mostly localized in the nuclear interior and near NPCs. Densely packed heterochromatin is considered transcriptionally repressed and tends to be located at the nuclear periphery and around the nucleoli, with likely connections in between<sup>29</sup>. The physical connections between chromatin and the nuclear envelope not only provide control over gene expression but also increase nuclear stiffness and stability, akin to the mechanical reinforcement used in composite materials or cross-linked polymer networks<sup>30–32</sup>. Although chromatin generally displays solid-like properties at the mesoscale, at high cation concentrations it can undergo liquid–liquid phase separation (LLPS) and, locally, chromatin can behave like a phase-separated condensate<sup>33,34</sup>. These physical properties of chromatin need to be considered when studying the contribution of chromatin to the mechanical properties of the nucleus (see next section).

**Physical connections between the nucleus and the cytoskeleton.** Force transmission between the cytoskeleton and the nucleus is required for nuclear movement and positioning, for example, during cell migration, nucleokinesis and muscle fibre regeneration<sup>35</sup> (FIG. 3). Cytoskeletal connections to the large and rigid nucleus are also important for cytoskeletal organization, affecting the organization of stress fibres, focal adhesions and cell–cell adhesions<sup>36,37</sup>. The physical coupling between the cytoskeleton and the nuclear interior is achieved by linker of nucleoskeleton and cytoskeleton (LINC) complexes that span the nuclear envelope<sup>35,36</sup> (FIG. 1a), although additional mechanisms, such as molecular motors binding to NPCs<sup>38</sup> or microtubules connecting to emerin and other nuclear envelope proteins<sup>39</sup>, may further contribute to nucleo-cytoskeletal coupling. LINC complexes are composed of nesprins (nuclear envelope spectrin repeat proteins) localized within the ONM



**Fig. 3 | Physiological sources of nuclear deformations.** **a** | Actomyosin contraction produces tension in actin fibres (in red) spanning the nucleus, which are connected to the nuclear envelope via linker of nucleoskeleton and cytoskeleton (LINC) complexes (in blue). In polarized adherent cells, such as epithelial cells, the contact to the basement membrane through transmembrane integrins defines a basal membrane, whereas the apical side has an exposed surface corresponding, for instance, to the lumen of internal cavities. Tension in apical actomyosin fibres generates vertical compressive forces that result in nuclear flattening. **b** | Contraction and stretching of myofibres induce nuclear deformations, including nuclear envelope wrinkling and expansion, respectively. Microtubules (in light blue) form a cage-like structure around nuclei that may provide additional mechanical support to the nuclei in contracting muscle fibres. **c** | Formation and regeneration of skeletal muscle fibre require migration of nuclei along the muscle fibre axis to the muscle fibre periphery, which involves LINC complexes and microtubule-associated motors, such as kinesin 1, that can pull on the nucleus, causing its movement and deformation. In addition, myofibril contraction drives nuclear movement to the fibre periphery during muscle fibre maturation. This process requires myofibrils to exert contractile forces on the nucleus, resulting in large nuclear deformations. This process is highly dependent on nuclear stiffness and lamin A/C levels.



**Topologically associating domains**

(TADs). Self-interacting megabase-scale genomic blocks in which DNA sequences exhibit significantly higher interaction frequencies with other DNA sequences within the domain than with those outside of the block.

**Liquid–liquid phase separation**

Physicochemical process leading to the formation of membraneless compartments or cell structures. This process is based on multivalent macromolecular interactions, including  $\pi$ – $\pi$  interactions, cation–anion interactions, dipole–dipole interactions and  $\pi$ –cation interactions, that drive the transition of some proteins into another phase with different physicochemical properties to induce the formation of membraneless organelles or cell structures.

**Nucleokinesis**

Translation of the nucleus within the cell, often neurons, which may or may not be coupled to cell migration.

**Stress fibres**

Actin filament assembly resulting from the interaction and merging of pre-existing radial fibres and transverse arcs (10–30 filaments). Stress fibres can reach a diameter of several hundreds of nanometres and are under constant tensile stress owing to actomyosin contractility.

**Focal adhesions**

Integrin-mediated cell–substrate adhesion structure anchoring the ends of stress fibres. Focal adhesions mediate strong attachments to substrates and function as an integrin-based signalling platform.

**Tensile force**

Pulling force resulting in the extension of an object.

**Viscoelastic**

Rheological property of natural or synthetic materials with viscous and elastic properties that allows for timescale-dependent deformation when subjected to mechanical stress.

that bind across the perinuclear space to SUN (Sad1p, UNC-84) domain-containing proteins located on the INM via their C-terminal KASH (Klarsicht, ANCl, Syne homology) domain<sup>35,40,41</sup>. This interaction appears to be at least in part responsible for controlling the spacing between the INM and the ONM<sup>35</sup>. On the cytoplasmic side, nesprin 1 and nesprin 2 bind to actin filaments<sup>42</sup> and — via kinesins<sup>43</sup> and dynein<sup>44</sup> — to microtubules (FIG. 1b). Nesprin 3 binds to intermediate filaments via plectin<sup>45</sup>. Nesprin 4, which is found in polarized epithelial cells, plays an important role in nuclear positioning via kinesin 1 (REF.<sup>46</sup>). KASH5 is a germ cell-specific KASH-domain protein required for proper meiosis<sup>47</sup>. On the nucleoplasmic side, SUN-domain proteins bind to the nuclear lamina, NPCs and chromatin. The current model considers that LINC complexes balance part of the cytoskeletal tensile force exerted on the ONM, with maximal stress values at nuclear extremities/poles<sup>48</sup>.

LINC complex localization at the nuclear envelope is associated with specific cellular functions. For example, LINC complex proteins are organized along apical stress fibres interacting with the cell nucleus<sup>49,50</sup> and at the front of the nucleus as cells squeeze their nuclei through small pores<sup>51</sup>. Although our current understanding of how LINC complex localization and force transmission are regulated is still incomplete, recent findings indicate that disulfide bonds between the SUN and KASH domains can serve as a crucial modulator of nucleo-cytoskeletal coupling<sup>35,41</sup>. Several additional components have been identified that mediate LINC complex function and force transmission, including FHOD1 (Formin homology 2 domain-containing protein 1)<sup>52</sup>, torsinA<sup>53</sup>, Samp<sup>54</sup> and lamins A/C<sup>55</sup>. Nesprins can also contribute to nucleo-cytoskeletal coupling independently of their actin and KASH domains via their spectrin repeats<sup>56</sup>. Nonetheless, many questions remain regarding the precise regulation of LINC complex formation and function.

**Nuclear mechanics**

The mechanical properties of the nucleus, including its size and stiffness, are one of the key pieces of information for understanding how nuclear deformations mediate cellular functions: the stiffer the nucleus, the more resistant to deformations it becomes.

**The physical properties of the nucleus.** Insights from various experimental assays<sup>57</sup> indicate that the nucleus behaves as a viscoelastic material, that is, it exhibits both elastic and viscous behaviour when subjected to external forces<sup>58</sup>. Elastic materials are defined by an instantaneous and reversible deformation, like a spring that extends under an applied force and snaps back to its original length when the force is removed. By contrast, viscous materials are liquid-like, exhibit flow and undergo irreversible deformation when subjected to force.

Numerous assays have been developed to quantitatively capture the rheological properties of the nucleus, ranging from micropipette aspiration and microindentation to stretching intact cells or isolated nuclei<sup>59</sup>. A major challenge lies in the fact that the viscoelastic response of the nucleus reflects a complex coupling between

chromatin, lamins and other nuclear components, and thus the exact behaviour can vary depending on the nature of the applied force/deformation and the molecular composition and organization of the cells being examined. Illustrating this challenge, some studies using micropipette aspiration found that the nucleus gradually deformed under an applied pressure before reaching a plateau, whereas, in other cases, the nucleus continued to deform under applied pressure, exhibiting a fluid-like behaviour<sup>58,60–63</sup>. Stretching isolated nuclei at physiologically relevant strain rates revealed that, for small deformations (<30% of the original length of the nucleus), the nuclear resistance is dominated by chromatin organization, whereas resistance to larger deformations is dominated by the expression levels of lamins A/C<sup>64</sup>. Furthermore, the nucleus undergoes strain stiffening, that is, it becomes stiffer and more difficult to deform upon larger deformations<sup>64,65</sup>.

After the removal of a mechanical strain, the elongated nucleus can relax with a nearly elastic response<sup>66–69</sup> or with a delayed response and even exhibit residual plastic deformation, characteristic of viscoelastic material properties<sup>70,71</sup>. The elastic response requires the presence of lamins A/C, SUN-domain protein linkages and vimentin<sup>67</sup>. Overall, these nuclear shape change dynamics may be explained by variations in nuclear lamina composition, chromatin organization, and cytoskeletal structure, composition and remodelling.

**Regulation of mechanical properties of the nucleus by its components.** Although A-type and B-type lamins share similar biochemical properties and filament structure, it is primarily the levels and assembly status of A-type lamins that determine nuclear stiffness and viscoelastic properties. Nonetheless, B-type lamins also contribute to nuclear stiffness and stability<sup>72,73</sup>, and loss of either lamin type results in abnormal nuclear shape and an increased propensity for nuclear envelope rupture<sup>66,74–77</sup>. Besides lamins, chromatin histone modification state and composition are major determinants of the mechanical properties of the nucleus, particularly for low nuclear deformation regimes<sup>64,78</sup>. Increasing the euchromatin content with histone deacetylase inhibitors, decreasing heterochromatin with histone methyltransferase inhibitors or disrupting dynamics of the linker histone H1, all lead to softer nuclei and more nuclear blebbing events — indicative of disturbed nuclear stability — independently of lamin levels<sup>64,78</sup>. New evidence also suggests that chromatin-associated proteins, such as HP1a, WDR5, BAF and NuMa, provide mechanical support to chromatin and regulate nuclear shape<sup>79–82</sup>. Interactions between chromatin and the nuclear envelope further contribute to nuclear stiffness by forming an interconnected network.

Furthermore, the physical properties of chromatin itself need to be considered when studying the contribution of chromatin to the mechanical properties of the nucleus. Although chromatin behaves as a solid at larger length scales, locally, chromatin can behave like a fluid<sup>33,34</sup>. It is now increasingly recognized that LLPS of nucleoplasmic components may serve as a key principle governing nuclear organization, with several nuclear

## Elastic

Property of a material that instantaneously deforms in response to a stress and recovers its size and shape after deformation. It is usually represented by a spring that stores energy in the form of elastic potential energy. Units of an elastic modulus are Pa (or  $\text{N m}^{-2}$ ).

## Viscous

Property of liquid of high viscosity, which corresponds to the resistance of a fluid to deform under either shear or extensional stress, defined as the ratio of shear stress to shear flow. Viscous fluids are usually depicted by a dashpot, which represents the internal friction within the fluid that dissipates energy over time. Units of viscosity are Pa s (or  $\text{N s m}^{-2}$ ).

## Strain

Geometric measure of the amount of deformation in the direction of the applied force divided by the initial length of the object (unitless number).

## Strain stiffening

Mechanical material property corresponding to a sudden increase of the elastic modulus under strain, that is, an increase in resistance to further deformation.

## Plastic deformation

Ability of a solid material to undergo permanent deformation (that is, irreversible change of shape) without rupture in response to applied forces.

## Linker histone H1

Histone protein family responsible for DNA compaction, whose members are located at the base of a nucleosome adjacent to the DNA entry/exit site to regulate the higher-order chromatin structure.

## Blebbing

Dynamic protrusion of the plasma or nuclear membrane, often characterized by a spherical morphology. At the cytoplasm, blebbing results from actomyosin contraction of the cortex that causes either transient detachment of the cell membrane from the actin cortex or a rupture in the actin cortex. The cytosol streams out and inflates the bleb. Nuclear blebs arise from increased intranuclear pressure and detachment of the nuclear membranes from the nuclear lamina.

components, such as the nucleolus or heterochromatin, showing properties of biomolecular condensates<sup>83–87</sup>. The propensity to form liquid droplets is enhanced in the vicinity of regions of low chromatin density because the higher mechanical energy required to deform the dense chromatin to create space for a growing protein droplet would generate an energetic penalty<sup>88</sup>. The growth of liquid droplets within the low chromatin density areas can lead to two distinct mechanical effects<sup>89</sup>. First, chromatin can be repelled as the drops grow by creating an effective repulsive interaction. In this case, the formation of protein condensates can be inhibited by the forces generated by the elastic chromatin network. A second effect can be driven by the tendency of neighbouring droplets to merge to minimize their surface energy. Indeed, regions of chromatin initially far apart and in separate droplets can be brought into proximity when the droplets merge, creating an effective attractive interaction that brings together different chromatin regions. The different types of interaction (repulsive versus attractive) between LLPS and chromatin are thus able to generate significant mechanical forces that can result in the structural rearrangement of chromatin<sup>90</sup>. Nonetheless, the relative contributions of LLPS versus other molecular mechanisms in determining the static and dynamic organization of chromatin within the nucleus remain to be fully elucidated. Additionally, the contribution of condensed chromatin to the mechanical integrity of the nucleus and its ability to respond to extranuclear forces are difficult to reconcile with a liquid state. Indeed, nuclear chromatin is mechanically responsive and can resist significant applied forces<sup>91</sup>, which is more consistent with a solid or gel state. Further studies that consider chromatin fibres as viscoelastic filaments that can behave as both a viscoelastic solid and a viscous liquid at different time and length scales may reconcile some of the apparently contradictory observations and ultimately provide a physical framework for understanding genome organization in space and time.

**Determinants of nuclear volume.** The initial observation that the ratio between cellular and nuclear volumes is largely constant across various cell sizes was made over 100 years ago<sup>92</sup>, and it is now well recognized that nuclear volume changes with chromatin organization and DNA content. Interestingly, accumulating evidence shows a close relationship between changes in cell morphology and nuclear deformations that often leads to modifications of nuclear volume, which can affect DNA synthesis<sup>93</sup>, gene transcription<sup>94,95</sup> and downstream signalling<sup>27</sup>. Yet, the precise mechanisms underlying nuclear volume regulation remain incompletely understood. The nuclear volume is determined by the balance between outward pressures that originate from the nucleoplasm and tend to expand the nucleus and inward pressures that originate from the cytoplasm and compress the nucleus. The outward pressures include contributions from both the chromatin and the fluid inside the nucleus. Notably, despite the presence of NPCs that facilitate flow of fluid either into or out of the nucleus, cells can establish hydrostatic pressure differences between the nucleoplasmic and cytoplasmic compartments<sup>96,97</sup>.

On the basis of the concept that the interior of living cells is ‘crowded’, changes in nuclear volume, such as inflating the nucleus, can be explained by the differences in colloid osmotic pressure between the nucleoplasm and cytoplasm<sup>98</sup>. Preliminary, theoretical works indeed suggest that the dominant pressure within the nucleus and cytoplasm originates from the osmotic pressure of the macromolecules preferentially localized to these compartments rather than from the effects of the mechanical properties of large complexes such as chromatin and the cytoskeleton<sup>99,100</sup>. The nuclear to cell volume ratio is determined by the number of macromolecules in the nucleoplasm and cytoplasm and increases when nuclear export is impaired owing to the accumulation of macromolecules in the nucleus<sup>100</sup>, demonstrating an active role of nucleo-cytosolic transport in the regulation of the osmotic pressure that controls nuclear size. More sensitive subcellular osmometers<sup>100</sup>, such as genetically encoded biosensors, are needed to establish definitive physiological values of colloid osmotic pressure and to determine how crowding inside cells is regulated as a function of the subcellular localization of macromolecules and physiological inputs.

**Adaptive changes in nuclear mechanics.** Deformation of cells and the nucleus can lead to changes in chromatin organization and compaction. These changes alter the mechanical properties of the nucleus as discussed above, providing a mechanism to prevent further deformations and protect the nucleus from mechanical stress<sup>16,101</sup>. Furthermore, mechanical force application can lead to the phosphorylation of emerin and subsequent recruitment of lamins to the nuclear envelope, causing rapid stiffening of the nucleus. In addition to binding to lamins, emerin is a recognized actin-binding protein that promotes actin polymerization at the nuclear envelope<sup>102</sup>. Emerin has also been recognized as a force sensor, relocating from the INM to the ONM in response to nuclear strain, leading to increased perinuclear actin polymerization<sup>103</sup>, which could alter nuclear deformability and protect it from damage<sup>104</sup>. By contrast, reducing cytoskeletal tension can soften the nucleus by increasing lamin phosphorylation and turnover<sup>105</sup>, highlighting the importance of the interplay between the nucleus and the cytoskeleton.

The difference in lamin expression between various cell types and tissues affects the deformability and mechanical stability of nuclei and may indicate tissue-specific adaptations to particular mechanical demands of the local microenvironment<sup>26,106–111</sup>. For example, nuclei in neutrophils have a particular lobulated morphology with characteristic low lamin A/C levels and elevated condensed chromatin level<sup>112</sup>; this nuclear organization promotes transit through tight spaces<sup>113</sup> such as lung capillaries that are only a few microns in diameter or even smaller gaps between endothelial cells. However, it is still under debate whether individual cells can dynamically adapt their nuclear stiffness on short timescales to promote migration through tight spaces. Confocal Brillouin microscopy revealed nuclear softening during transendothelial migration of breast cancer cells<sup>114</sup>. However, the origin and timing of such nuclear

**BAF**

Barrier-to-autointegration factor is an essential 10 kDa chromatin-binding protein that is highly conserved in metazoa and helps DNA anchoring to the nuclear envelope. BAF is involved in multiple pathways, including nuclear envelope reassembly (after mitosis and nuclear envelope rupture), chromatin epigenetics and DNA damage response. BAF function is controlled by phosphorylation/dephosphorylation waves that drive nuclear envelope disassembly.

**Biomolecular condensates**

Micron-scale compartments often formed by liquid–liquid phase separation that lack surrounding membranes and concentrate functionally related components such as proteins and nucleic acids.

**Colloid osmotic pressure**

Pressure generated by solutions of macromolecules in contact with pores that are permeable to water and ions but not to macromolecules. Colloid osmotic pressure generates depletion forces that push macromolecules together in crowded solutions and thus promotes aggregation and phase separation.

**Confocal Brillouin microscopy**

Optical technique combining Brillouin spectroscopy with confocal microscopy to provide a non-contact and direct read-out of the mechanical properties of a material (that is, stiffness, temperature or strain) at the micrometre scale. Spontaneous Brillouin light scattering arises from the interaction between photons and acoustic phonons (that is, propagation of thermodynamic fluctuations) and permits quantification of the intracellular longitudinal modulus without disturbing the cell.

softening remain poorly understood. Interestingly, inhibition of metalloproteinases that remodel the extracellular matrix (ECM) leads to nuclear softening via lamin A/C phosphorylation, which is essential for migration through ECM pores with a subnuclear diameter (confined cell migration; see also next section)<sup>115,116</sup>. This response requires an intact connection between the nucleus and the centrosome via the LINC complex protein nesprin 2 and the dynein adaptor Lis1 (REF.<sup>116</sup>). Chromatin remodelling can further modulate nuclear stiffness and cell migration in 3D environments<sup>80</sup>. These findings suggest that dynamic chromatin modification and changes in lamin levels and organization can mediate nuclear mechanics and promote cell migration in confined 3D environments<sup>117,118</sup>. However, reducing lamin A/C levels below a critical threshold may reduce cell survival under mechanical stress<sup>75,119,120</sup>.

**Sources of nuclear deformations**

The nucleus is constantly exposed to forces from the surrounding cytoskeleton, including from active positioning of the nucleus during cell polarization<sup>121</sup>, migration<sup>121</sup> or differentiation<sup>122</sup>. Recent advances in intravital imaging and modelling physiological microenvironments *in vitro* have documented large-scale nuclear deformations related to contraction and relaxation of striated muscle<sup>123,124</sup> and during confined cell migration<sup>75,76,125,126</sup>, although similar nuclear deformations and functional consequences are expected to also occur during numerous other important situations, including developmental cell migration<sup>127,128</sup> and nucleokinesis events<sup>129</sup>. Here, we discuss several physiological and pathological situations associated with nuclear deformations and how these deformations arise.

**Nuclear deformations in cells adhering to flat and rigid substrates.** Actin stress fibres and actomyosin contractility can impose vertical and lateral inward compressive forces on the nucleus. Lateral actin fibres can lead to nuclear deformations when cells migrate or are stretched<sup>130,131</sup>. Vertical compressive forces are exerted by apical actin stress fibres that form a dome-like structure across the nucleus and that are physically attached to the nuclear lamina through LINC complexes<sup>132</sup>. On flat rigid substrates, these forces flatten the nucleus during cell spreading (FIG. 3a) and can cause nuclear envelope rupture<sup>133–135</sup>. By contrast, the nucleus remains more rounded in cells on soft substrates<sup>136</sup> that are associated with lower cytoskeletal tension and fewer actin stress fibres<sup>137</sup>, or when the actin cytoskeleton or LINC complexes are disrupted<sup>135</sup>. Indeed, ventral actin fibres, which are thick actomyosin bundles connected from both ends to focal adhesions at the bottom of the cell, can exert lateral compressive forces on both nuclear sides<sup>93</sup>. The high level of tension in ventral actin stress fibres can lead to nuclear indentations. These indentations can measure a few microns and are characterized by local enrichment of LINC complexes and segregated domains of condensed chromatin, indicating that the nucleus responds to compression by adjusting its architecture<sup>50,138</sup>. Collectively, these findings suggest that the amount of tension within the perinuclear actin

fibres is an important source of nuclear deformations and nuclear mechanotransduction.

**Nuclear deformations in skeletal and cardiac muscle.**

Actomyosin contractility also has an important role in nuclear deformations in striated muscle cells (FIG. 3b). Large nuclear deformations were recently visualized in cardiac and skeletal muscle contraction in living fly larvae<sup>124</sup>. Increased expression of lamins A/C in muscle cells is essential to protect their nuclei from mechanical damage caused by muscle contraction<sup>26</sup> and during nuclear movement associated with myoblast elongation<sup>139</sup>. Another, more surprising mechanism responsible for mechanical stress on the nucleus are the cytoskeletal forces required to position muscle nuclei along the length of the muscle fibre and the nuclear periphery during myotube maturation<sup>140,141</sup>. LINC complex proteins such as nesprin 1, together with the microtubule associated motors kinesin 1 and dynein as well as other nuclear envelope proteins such as emerin, have been implicated in this process<sup>140,142,143</sup>. Generally, the physical stress associated with the motors pulling on the nucleus results in nuclear rotation and nuclear deformations<sup>144–146</sup> (FIG. 3c, left). In lamin A/C-deficient or mutant cells, which have mechanically weaker nuclei, the kinesin-mediated forces can result in large-scale nuclear deformations and damage<sup>119</sup>. In addition to the role of motor proteins, myofibril contraction was shown as a mechanism to move skeletal muscle nuclei to the periphery of muscle fibres, also incurring nuclear deformations in the process (FIG. 3c, right). Both a reduction and an increase in nuclear stiffness (by lamin A/C depletion or overexpression) perturbed the nuclear repositioning. Additionally, lamin A/C deficiency was associated with particularly pronounced nuclear deformations, suggesting an important role of nuclear mechanical properties in regulating this nuclear repositioning event<sup>143</sup>. Intriguingly, in lamin A/C-deficient and mutant mouse models that develop severe muscular dystrophy and dilated cardiomyopathy (BOX 2), reducing the cytoskeletal forces acting on the fragile muscle cell nuclei by disrupting the LINC complex prevents nuclear damage and results in improved muscle function and muscle cell viability *in vitro* and *in vivo*<sup>119,147</sup>, pointing to promising new therapeutic approaches for these devastating diseases. However, given that mutations in nesprins and SUN proteins can lead to muscular dystrophy and heart disease<sup>148</sup>, further studies will need to evaluate the long-term risks and consequences of LINC complex disruption using, for example, inducible LINC complex disruption models<sup>149</sup>.

**Nuclear deformations in developing tissues.** In early *Drosophila* embryo, pronounced nuclear deformations occur during cellularization — a process during which somatic nuclei at the periphery of the syncytial embryo move as the plasma membrane invaginates to form membranes around each nucleus. The nuclear deformations are caused by the formation of microtubules into bundles that run across the nuclear envelope<sup>150</sup>. These nuclear deformations may be particularly pronounced because A-type lamins are not expressed in *Drosophila* during cellularization, leading to more



deformable nuclei<sup>151</sup>. Nuclear movement during development also results in substantial nuclear deformations in the nematode *Caenorhabditis elegans*, which require cytoskeletal force transmission to the nucleus via the LINC complex<sup>152</sup>.

In epithelial systems, cellular intercalation is a common process occurring throughout development, whereby neighbouring cells exchange their place to maintain epithelium integrity. Depending on the cell density, cellular intercalation can lead to transient squeezing and nuclear deformations in the intercalating cell (FIG. 4a) likely due to compression by neighbouring cells and cytoskeletal remodelling that transmits forces onto the nucleus<sup>153,154</sup>.

Another phenomenon occurring during development that is associated with nuclear deformations is nucleokinesis in the central nervous system. One such nucleokinesis event is interkinetic nuclear migration in neural progenitor/stem cells around cell divisions<sup>155,156</sup> as is nucleokinesis of newborn neurons that migrate to their final destination in the tissue<sup>157</sup>. Both actin and microtubules have been involved in these nucleokinesis processes, depending on the system and cell type<sup>158</sup>. Microtubules can exert pulling forces on the nuclear lamina through LINC complexes that move the nucleus towards the centrosome, whereas actomyosin could push the nucleus from behind (FIG. 4b). Neuroepithelia are densely packed with cells, necessitating the nuclei to squeeze through narrow spaces. Thus, these cytoskeletal forces, together with the need for the nucleus to navigate the dense neuroepithelial tissue, result in nuclear deformations. Notably, developing neural tissues lack the expression of lamins A/C, which makes the nuclei less rigid, thereby supporting nuclear deformability<sup>159</sup>. At the same time, developing neural tissue requires lamin B to maintain nuclear integrity during nucleokinesis. For example, in the developing brain, loss of either lamin B1 or lamin B2 causes defective migration of cortical neurons and leads to severe nuclear architectural abnormalities (for example, chromatin protrusions) and nuclear membrane ruptures, likely explaining the severe brain development defects and reduced neuronal survival associated with B-type lamin deficiency<sup>73</sup>. It remains to be determined whether these defects are caused by disrupted transmission of force during nuclear movement or by a more fragile nucleus unable to bear the stress generated during nucleokinesis.

Besides nucleokinesis, live imaging studies have found remarkable nuclear deformations and rotation during the migration of cerebellar granule cells through narrow intercellular spaces in neural tissues<sup>144</sup>. During this process, microtubules steer the nucleus and drive its rotation and deformation through a dynamic interaction of nesprins with kinesin 1 and dynein. Given the apparent diversity of cytoskeletal organization in neuron species, further studies are needed to obtain a better understanding of nuclear dynamics and nuclear shape regulatory mechanisms in neural tissues.

**Nuclear deformations during confined migration.** Nuclear deformation is a hallmark of important physiological and pathological situations involving cell migration.

For instance, immune cells or invasive cancer cells must navigate through small interstitial spaces ranging from 1 to 20  $\mu\text{m}$  in diameter<sup>160,161</sup>, which requires cells to deform their nucleus to squeeze through the available spaces (FIG. 4c,d). In the absence of matrix metalloproteinases that digest the ECM and widen migratory tracks, the nucleus is often the main physical hindrance to cell migration through confined spaces<sup>75,125</sup>. Leukocytes can insert basolateral protrusion within (paracellular) or between (transcellular) endothelial cells to breach the endothelial barrier (FIG. 4c) and use actomyosin forces to push the nucleus through the pore, resulting in substantial nuclear deformations.

Tumour cells face similar challenges when invading tissues and intravasating and extravasating blood vessels to metastasize to distant tissues<sup>114</sup> (FIG. 4d). One of the primary sources of cytoskeletal forces to translocate and deform the nucleus is actomyosin contractility. This contractility can cause both tension and compression of the nucleus by actin stress fibres pulling or pushing on the nucleus<sup>146,162,163</sup>. However, build-up of actomyosin contractility can also increase the cytoplasmic hydrostatic pressure, which results in the influx of cytoplasmic content into the nucleus causing its volume expansion and blebbing, which hinders motility<sup>97</sup>. An additional, actin-based mechanism has been observed in dendritic cells, whose nuclei are rigid owing to high expression of lamina A/C. These cells use Arp2/3 complex, a central actin nucleator, to generate a perinuclear actin network. These perinuclear actin filaments accumulate around the constriction site and exert a lateral pushing force on the nucleus, facilitating migration through narrow ECM pores<sup>164</sup>. Alternatively to actomyosin contractility, mechanisms for propelling the nucleus may involve microtubule-associated motors, kinesins and dyneins<sup>165</sup>, which directly attach to the nucleus via nesprins and other proteins at the nuclear envelope, dragging the nucleus along the microtubule tracks. Whether the nucleus is pulled and/or pushed during confined migration is still debated<sup>166</sup>, although it is likely that cells can use multiple independent mechanisms, depending on the particular context (FIG. 4d). Hence, the nuclear deformation pattern can be expected to vary in different *in vivo* scenarios of confined migration.

Nuclear deformations during confined migration may also involve dynamic or persistent changes in nuclear mechanical properties. For example, transient nuclear softening has been reported during transendothelial migration of cancer cells<sup>114</sup>; neutrophils develop highly lobulated and deformable nuclei during granulopoiesis, which facilitates passages through tight spaces<sup>113</sup>; and highly invasive breast cancer cells are characterized by increased nuclear deformability and low lamin A/C levels<sup>115</sup>. Notably, the physical properties of the large nucleus can directly influence confined migration. The microtubule-mediated ‘frontward’ positioning of the nucleus in amoeboid cell migration was shown to allow cells to use their nucleus as a mechanical gauge to determine the path of least resistance when encountering bifurcations of the path with pores of different sizes<sup>162</sup>. This provides an example of how deformation of the nucleus aids cells in their ‘decision-making’ during migration through confined environments.

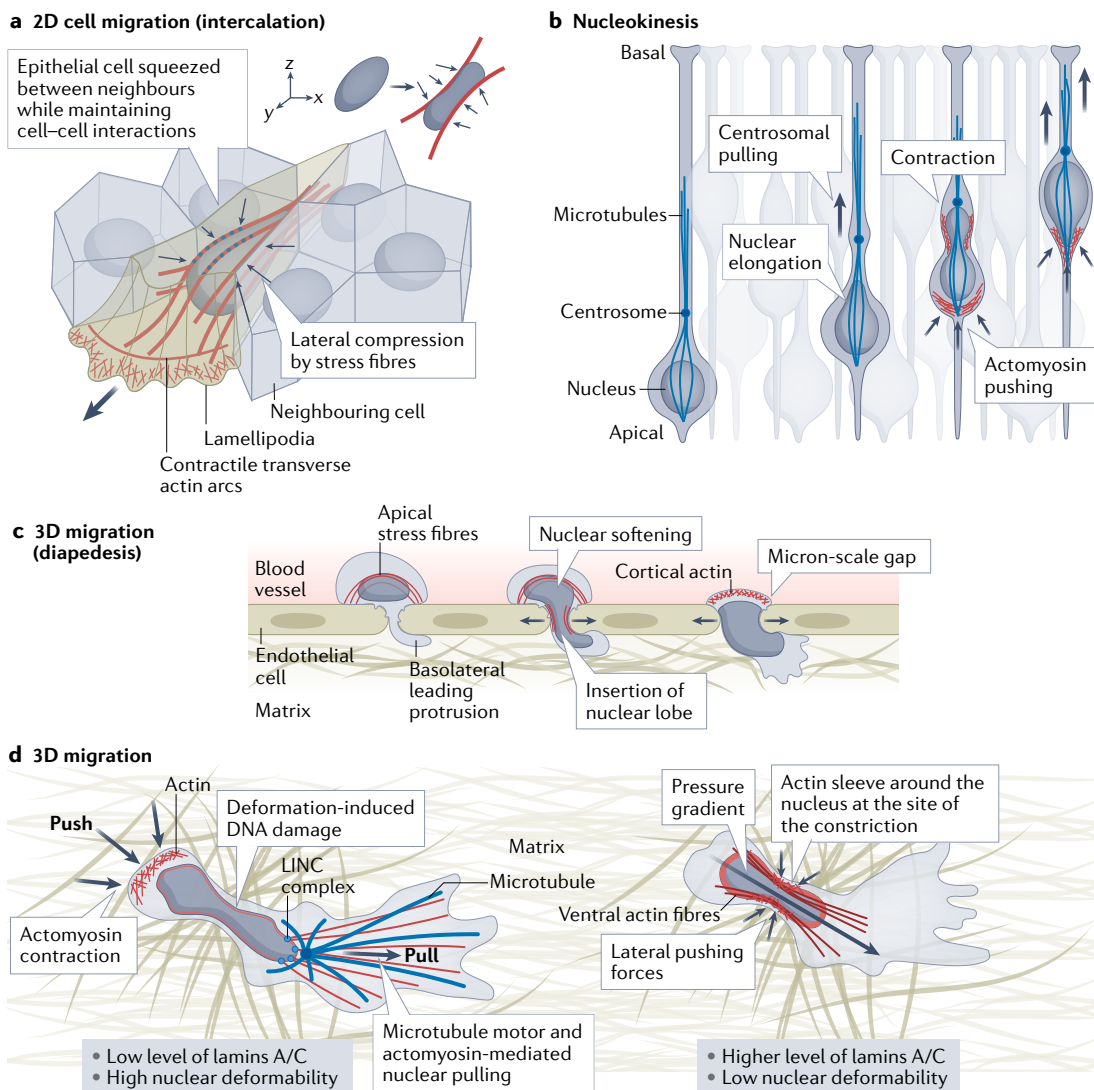
#### Interkinetic nuclear migration

Periodic movement of the nucleus between apical and basal surfaces of neuroepithelial progenitor cells as they progress through the cell cycle. Interkinetic nuclear migration results in all mitoses taking place at the apical side of the neuroepithelium. As a consequence, most newborn neurons resulting from division of neuroepithelial progenitors must move their soma from the apical side to more basal locations where they function.

#### Cerebellar granule cells

Among the smallest and the most numerous neuron type that form dense and distinct layers of the cerebellar cortex.





**Fig. 4 | Migration-associated nuclear deformations.** **a** | Epithelial cell intercalation within dense bidimensional tissues requires cellular elongation and nuclear deformations. Lateral compressive forces are exerted on both nuclear sides by ventral fibres, which are thick actomyosin bundles connected from both ends to focal adhesions. **b** | Nucleokinesis events are observed during development of the neuroepithelium of the central nervous system and are accompanied by considerable nuclear deformations. One of these events is the migration of early-born neurons, which reposition their soma from the apical to the basal side of the neuroepithelium to reach their final destination. This event occurs in densely packed, neuroepithelial tissue and involves pulling forces on the nucleus exerted by a microtubule cage towards the centrosome and pushing forces at the cell rear generated by actomyosin contraction, depending on the system and the neuronal cell type. In mammals, microtubules exert pulling forces on the nuclear lamina through linker of nucleoskeleton and cytoskeleton (LINC) complexes that move the nucleus towards the centrosome. Together with stresses instigated by neighbouring cells, these cytoskeletal forces deform the nucleus. **c** | Immune cells and tumour cells can breach the endothelial barrier of blood vessels by inserting protrusion between or inside endothelial cells. Migration through the small openings in the endothelium (a few micrometres in diameter) is associated with large nuclear deformations and may be linked to nuclear softening. **d** | Migration of cells *in vivo* requires them to squeeze through narrow spaces, navigating often very complex and dense structures of the extracellular matrix as well as moving in between cells. Actomyosin contraction leads to pushing and pulling forces and cooperates with microtubule motors at the front, which are recruited to generate pulling forces. Together, the balance of forces results in the forward movement of the nucleus through the narrow constriction. Nuclear deformations result from the balance between the amount and direction of the applied cytoskeletal force, the mechanical properties of the nucleus and the degree of external confinement. Cells with low levels of lamin A/C expression, and thus more deformable nuclei, can more easily move through tight spaces as seen in neutrophils and some metastatic cancer cells. However, reduced lamin expression makes cells more prone to nuclear envelope rupture (FIG. 5). Cells that express high levels of lamins A/C (for example, dendritic cells) can use a perinuclear, actin 'sleeve' that is recruited at the site of the constriction to locally deform the stiffer nucleus.

### Laminopathy

Over 450 mutations have been reported in the genes encoding lamins, in particular the *LMNA* gene, causing diseases termed laminopathies. The number of identified laminopathies has steadily increased in recent years, currently including 13 known conditions. Most of these diseases are rare but *LMNA* mutations are the second most common cause of congenital dilated cardiomyopathy. Although lamins are nearly ubiquitously expressed, many of the laminopathies exhibit tissue-specific phenotypes, for example, primarily affecting striated muscles and tendons, hence the suggestion of a mechanical connection.

### LEM-domain proteins

The LAP2, emerin and MAN1 (LEM) domain is a ~40-residue helix–loop–helix fold conserved both in eukaryotes and in prokaryotic DNA/RNA-binding proteins. Except for LAP2 proteins, which have a second LEM domain that binds DNA, the function of a eukaryotic LEM domain is to directly bind the conserved chromatin protein BAF.

### TREX1

Three prime repair exonuclease 1 is the major 3' → 5' DNA exonuclease in mammalian cells and metabolizes preferentially single-stranded DNA. It cleans the cytosol from DNA fragments coming from endogenous elements. Unless degraded, the accumulation of these DNA fragments can activate innate immune signalling.

### ATR kinase

Serine/threonine protein kinase activated in S phase and involved in sensing DNA damage and activating DNA damage checkpoint upon genotoxic stresses (for example, ionizing radiation or ultraviolet light), thereby acting as a DNA damage sensor.

### Epithelial–mesenchymal transition

Transcriptionally governed process over which epithelial cells establish a front–rear polarity while acquiring a mesenchymal and motile phenotype.

### Consequences of nuclear deformations

Given the central role of the cell nucleus in cellular function, it is easy to imagine how nuclear deformations can lead to various transient or persistent consequences, including downstream signalling, altered nucleo-cytosolic transport and genome regulation as well as loss of nuclear envelope integrity and DNA damage (FIGS 5,6). Notably, although these outcomes of deforming the nucleus are now well established, the molecular mechanisms responsible and whether the nucleus itself senses mechanical signals and translates these into biological responses (BOX 1) often remain unresolved and a matter of active research.

### Deformation-associated nuclear envelope rupture and repair

Nuclear envelope rupture describes the transient loss of nuclear membrane integrity at localized sites rather than global breakdown of the nuclear envelope as occurs in vertebrate cells during mitosis. Spontaneous nuclear envelope rupture events, persisting typically for between a few minutes and several dozens of minutes, were first observed *in vitro* in cells expressing the HIV protein VPR<sup>167</sup> and subsequently in fibroblasts from patients diagnosed with laminopathy<sup>168</sup> and in cancer cells<sup>146</sup>. Since then, it has become apparent that physical stress on the nucleus and the associated nuclear deformations can lead to transient nuclear envelope rupture events, particularly during migration through confined environments, and that the probability of nuclear envelope rupture increases with the degree of confinement<sup>75,76,163,169</sup>.

Nuclear envelope rupture events have been documented *in vitro* and *in vivo*. These ruptures are often associated with loss of A-type or B-type lamins<sup>168,170,171</sup>, lamin mutations<sup>172–174</sup>, peripheral heterochromatin disruption<sup>91</sup> or high-level mechanical stress resulting from tensile or compressive forces on the nucleus<sup>74–76,164,175–178</sup>. On the basis of super-resolution imaging and computational modelling, the nuclear envelope rupture sites are estimated to be ~100 nm in diameter<sup>73,74,179</sup>. A current hypothesis proposes that nuclear envelope ruptures occur at pre-existing gaps or defects in the nuclear lamina, particularly where the lamin B meshwork is weaker and thus cannot sufficiently support the nuclear membranes. This mechanical fragility causes the membrane to form a bleb that expands under continued mechanical stress and ultimately ruptures<sup>169,180</sup> (FIG. 5). However, nuclear envelope ruptures and membrane blebs have also been observed in the absence of nuclear lamina gaps; they may thus generally arise when the nuclear membranes peel off the underlying nuclear lamina in response to increased nuclear pressure resulting from cytoskeletal forces<sup>75,181,182</sup>. A better understanding of the mechanisms that drive nuclear envelope rupture will require study of the dynamics of the heterogeneous lamina meshwork and its interaction with the nuclear membranes during nuclear deformations.

In line with the observations that most nuclear envelope rupture events are transient, cells have robust mechanisms to repair their nuclear membrane during interphase and even more persistent rupture events (a few hours) can eventually be repaired<sup>183</sup>. The mechanisms involved

in interphase nuclear membrane repair are largely shared with those during resealing of the nuclear envelope after mitosis. The nuclear membrane repair mechanism is based on the recruitment of specific proteins to the sites of nuclear envelope rupture, particularly BAF, LEM-domain proteins, A-type lamins and membrane remodelling proteins, including endosomal sorting complexes required for transport (ESCRT)-III remodelling complex and its binding recruiting factor CHMP7 (REFS<sup>76,177,183–185</sup>). The extent of rupture is correlated with the amount of cytoplasmic BAF entering the nucleus and accumulating at the rupture site<sup>76,183,185</sup>. A current model of nuclear membrane repair considers that the binding of cytosolic BAF to the exposed chromatin initializes recruitment of both new ER membranes to repair the membrane hole and the ESCRT-III complex to reseal the remaining gaps (FIG. 5). BAF also recruits cytoplasmic lamins A/C to the rupture site, further contributing to the restoration of nuclear envelope integrity. Interestingly, some nuclear processes, such as transcription and DNA replication, can be disturbed after nuclear envelope rupture events, leading, for instance, to aneuploidy or extensive DNA damage such as persistent double-stranded DNA breaks<sup>134</sup>.

**Mechanically induced DNA damage.** Severe nuclear deformations occurring, for example, during confined migration, external compression or nuclear repositioning in dense tissues can induce DNA damage upon nuclear envelope rupture<sup>72,73,171,186</sup> and even in the absence of rupture<sup>187</sup> (FIG. 5). Nuclear envelope rupture can cause DNA damage by allowing access of the ER-associated exonuclease TREX1 into the nucleus<sup>186</sup> or by loss of DNA damage repair factors from the nucleus via nuclear efflux<sup>171,182</sup>. Nuclear envelope rupture-associated DNA damage occurs throughout all phases of the cell cycle, more often in cells deficient for the DNA damage sensor ATR kinase<sup>188</sup>. By contrast, deformation-induced DNA damage (DNA damage in the absence of nuclear envelope rupture) occurs primarily in S/G2 phases, that is, during active DNA replication. This damage is linked to increased replication stress, possibly due to torsional stress on DNA resulting from the nuclear deformation during confined migration or mechanical compression of cells<sup>187</sup>. Interestingly, different cell lines exhibit different propensities for these modes of DNA damage<sup>186,187</sup>, but the exact molecular underpinnings for these cell type-specific differences remain to be elucidated.

What are the long-term consequences of DNA damage and nuclear envelope rupture for cells and tissues homeostasis? Repeated migration through tight constrictions can lead to the accumulation of DNA damage and changes in chromosome copy number, which may drive the emergence and evolution of malignant cells<sup>182</sup>. Furthermore, TREX1-dependent DNA damage following nuclear envelope rupture may favour tumour cell invasion by inducing a partial epithelial–mesenchymal transition, including increased degradation of collagen and increased invasive potential<sup>172</sup>. The precise mechanisms linking TREX1 and collagen degradation activity is still unknown but is believed to be downstream of the DNA damage response pathway initiated by ATM kinase<sup>189,190</sup>.

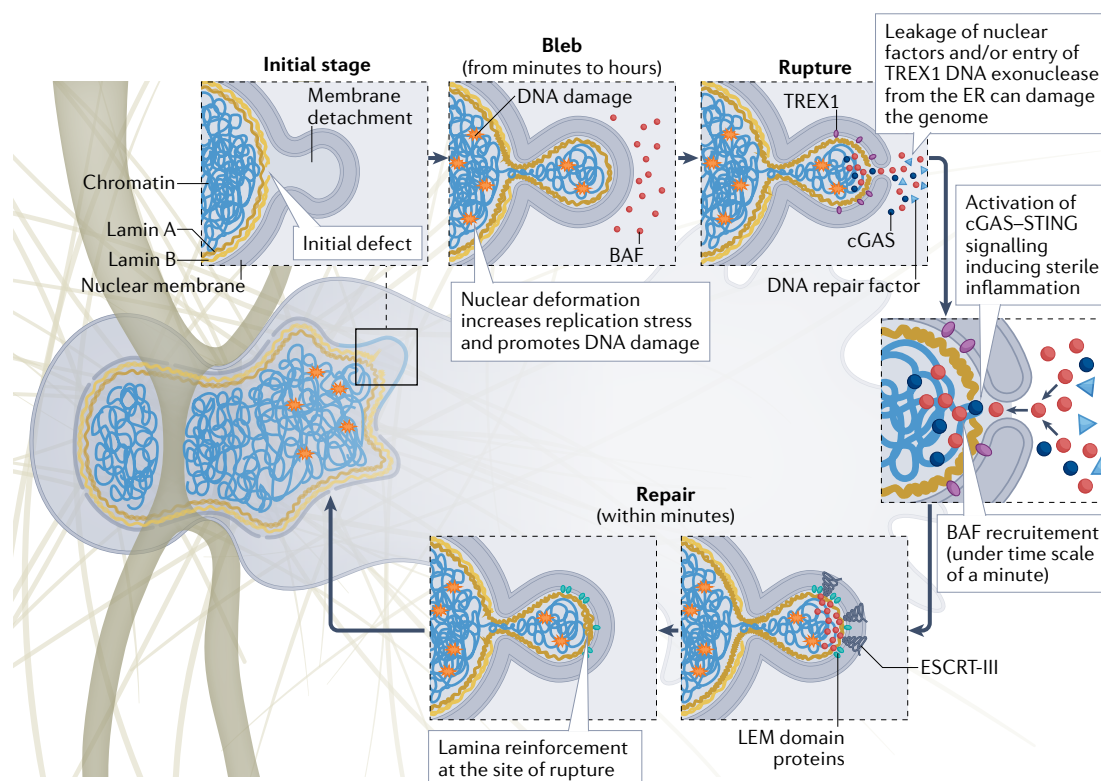
**ATM kinase**

Serine/threonine protein kinase that is recruited and activated to sites of DNA double-strand breaks and signals to various downstream targets to initiate cell cycle arrest and DNA repair.

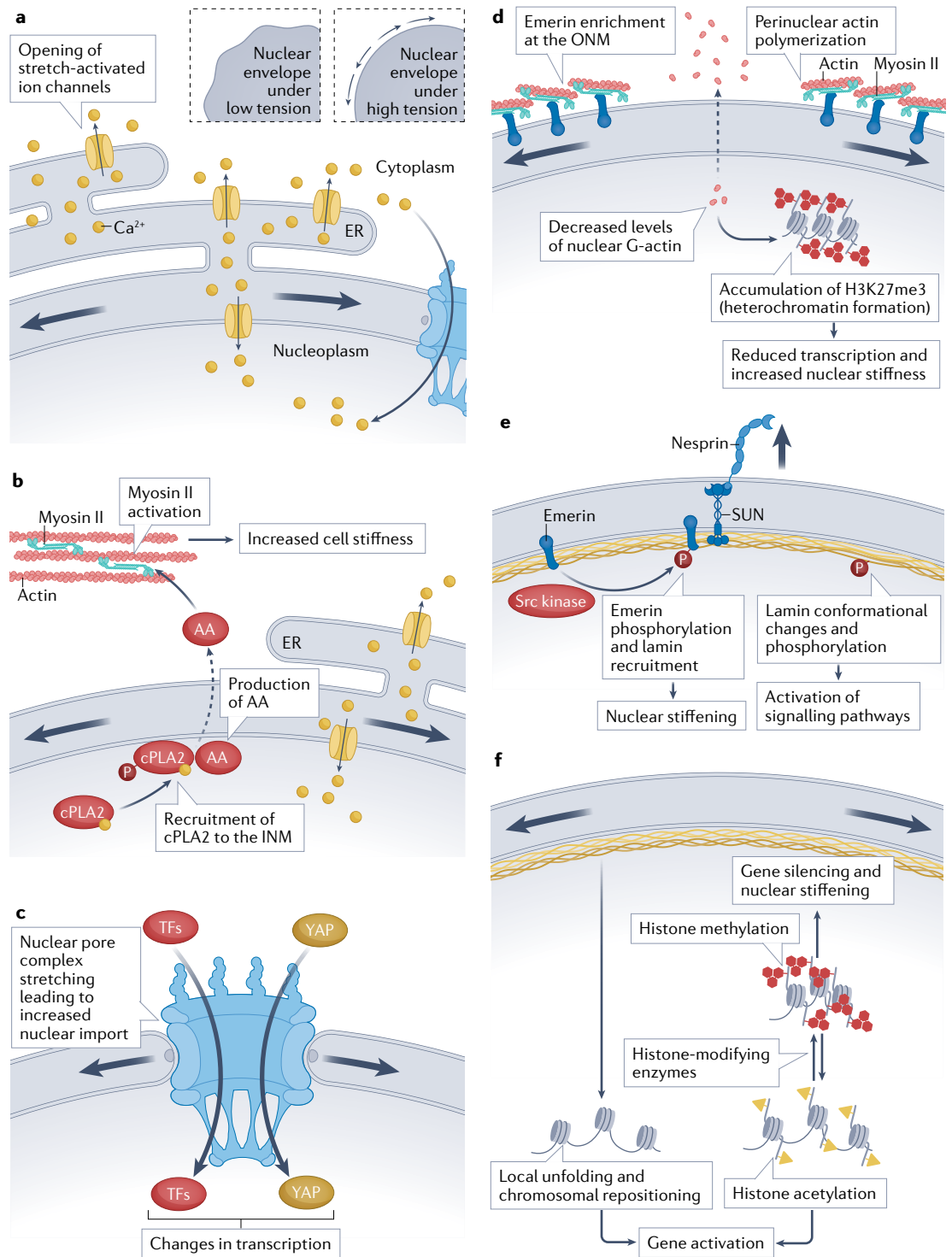
Nuclear envelope rupture can also lead to activation of the pro-inflammatory cGAS–STING DNA-sensing pathway, as it allows access of cytosolic cGAS to the genomic DNA at sites of rupture<sup>75,76,191</sup>. A recent study found that increased cGAS–STING signalling can drive cancer metastasis in a mouse breast cancer model<sup>191</sup>, although, in this case, cGAS–STING activation was primarily due to nuclear envelope rupture in micronuclei and not in primary nuclei.

**Nuclear deformation-associated signalling.** Confinement of cells below a critical threshold, typically a fraction of the uncompressed nuclear height, results in nuclear flattening, an increase in nuclear membrane tension and

unfolding of nuclear membrane invaginations<sup>17,127</sup>. Unfolding of the nuclear envelope under increasing membrane tension allows the nucleus to deform without exceeding critical membrane tension in the nuclear membranes<sup>186</sup> but may also trigger downstream signalling events. This nuclear mechanosensing of cellular confinement has been referred to as ‘cellular proprioception’. One example is the increased uptake of calcium into the nucleus, which is promoted by calcium release from the ER — an event that is also mechanically triggered, resulting from confinement, nuclear flattening and expansion of the nucleus/ER–plasma membrane contact area (FIG. 6a). Increased nuclear membrane tension, further amplified by the increased intranuclear



**Fig. 5 | Nuclear envelope rupture and repair.** Migration through confined environments or external compression of cells can result in nuclear envelope ruptures. The rupture process is typically initiated by the formation of a nuclear membrane extrusion, or bleb, where the nuclear membranes detach from the underlying lamina. Blebs are driven by increased hydrostatic pressure within the nucleus. Blebs form at sites with high nuclear membrane curvature and where an initial defect (weakening) in the nuclear lamina exists. Blebs can have varying size and can contain chromatin or can just be filled with fluid. They typically lack lamin B and nuclear pore complexes, whereas lamins A/C and chromatin can enter the bleb. Continued nuclear compression by confinement from the extracellular matrix, apical actin stress fibres, cell contractions or external compression results in bleb expansion until the nuclear membranes in the bleb exceed a critical strain threshold and rupture, leading to the leakage of soluble proteins from the nucleoplasm into the cytoplasm and uncontrolled influx of cytoplasmic proteins into the nucleus. The lifetime of blebs can range from minutes to hours, but the rupture itself is usually quite short, on the order of minutes. Following nuclear envelope rupture, BAF is rapidly (within minutes) recruited to initiate nuclear envelope repair. The recruitment of endosomal sorting complexes required for transport (ESCRT)-III complexes further contributes to resealing the nuclear membranes. The process of repair/rescue is typically completed within 10–15 min and often associated with recruitment of nucleoplasmic lamins A/C to the site of rupture. Although the rupture is resealed, the bleb/protrusion often persists and is not fully resorbed. Severe nuclear deformation can sensitize cells to DNA damage. This can be caused by nuclear envelope rupture, which has been linked to the translocation of exonuclease TREX1 from the endoplasmic reticulum (ER) to the inner nuclear membrane. Nuclear rupture may also cause depletion of DNA repair factors, promoting damage, and can also induce sterile inflammation by exposing nuclear DNA to the cytosolic DNA sensor cGAS–STING. In cells undergoing S phase, nuclear deformation can promote DNA damage even in the absence of nuclear envelope rupture, likely by inducing torsional stress and interfering with DNA replication.



**Sterile inflammation**  
Immune response that is typically associated with the recognition of intracellular contents released from damaged and necrotic cells by inflammatory signalling receptors or triggered by exogenous material that can injure cells. This process occurs in the absence of microorganisms.

calcium concentrations<sup>192–194</sup>, results in the recruitment of nucleoplasmic phospholipase A2 (cPLA2) to the INM, where it catalyses the production of arachidonic acid (an omega-6 polyunsaturated fatty acid) and lysophosphatidic acid, which are then released to the cytoplasm<sup>192,194,195</sup> (FIG. 6b). Recruitment of cPLA2 to the INM can be triggered by osmotic swelling associated with cell and tissue injury, inducing inflammatory signalling<sup>192</sup>, or by physical confinement of cells<sup>17,127</sup>. Arachidonic acid has been implicated in regulating myosin II activity, both directly<sup>196</sup> and indirectly via

protein phosphorylation<sup>197</sup>, which results in the increase in cortical actomyosin contractility (FIG. 6b). Thus, the higher nuclear membrane tension resulting from nuclear deformations modulates cell morphology and promotes migration through narrow constrictions<sup>17,127</sup>.

**Nuclear deformations and nuclear transport.** Recent structural evidence indicates that nuclear membrane tension is associated with an open state of the NPCs and that reduced tension causes NPC constriction<sup>198,199</sup>. Hence, it is conceivable that forces acting on the nucleus



◀ **Fig. 6 | Examples of functional consequences of nuclear deformations.** **a** | High tension exerted on the nuclear envelope during nuclear deformations induces straightening and unfolding of the wrinkled nuclear envelope, which may lead to the opening of stretch-activated ion channels. The nature of these channels remains to be established, but it was suggested that a key mechanosensitive calcium channel Piezo1 localizes to the nuclear envelope and the endoplasmic reticulum (ER) in addition to the plasma membrane. As the nuclear envelope is continuous with ER membranes, the stress on the nuclear envelope may also propagate to the ER, leading to the opening of mechanosensitive channels in that location. **b** | Increased nuclear membrane tension, coupled with calcium release from the ER that increases intranuclear calcium concentrations, promotes the recruitment of cytosolic phospholipase A2 (cPLA2) from the nucleoplasm to the inner nuclear membrane (INM), where cPLA2 synthesizes arachidonic acid (AA) that is subsequently released to the cytoplasm. The activation of the cPLA2–AA pathway leads to RhoA activation and increased myosin II recruitment to the cell cortex, increasing actomyosin contractility. **c** | Increased nuclear membrane tension promotes stretching of nuclear pore complexes, leading to increased nuclear import of transcription factors (TFs) and mechanoresponsive transcriptional activators, such as YAP (Yes-associated protein)/TAZ (transcriptional coactivator with PDZ-binding motif). **d** | Deformation of the nucleus induces enrichment of emerin at the outer nuclear membrane (ONM). Relocalization of emerin to the ONM promotes perinuclear actin polymerization that leads to decreased levels of intranuclear free monomeric actin (G-actin). This perturbs the activity of several chromatin modifiers that bind to G-actin, such as HDAC1/2, Tip60, INO80, SWI1, SWI/SNF and RSC100, resulting in increased heterochromatin formation (accumulation of histone H3 Lys27 and Lys9 trimethylation (H3K27me3, H3K9me3)). By increasing compaction of the genome, these epigenetic changes reduce global transcriptional activity and impact the mechanical properties of the nucleus. **e** | Nuclear deformations transduced by linker of nucleoskeleton and cytoskeleton (LINC) complexes induce phosphorylation of emerin, which is mediated by Src family kinases, and confer nuclear adaptation to force by promoting lamin recruitment, thereby causing nuclear stiffening. Nuclear deformations can also induce conformational changes in lamins A/C and/or modulate the phosphorylation status of lamins A/C, which can alter the interaction of lamins with their binding partners and influence lamin distribution, dynamics and degradation, initiating further signalling events and promoting changes in genome organization. **f** | Forces acting on the nucleus may reposition or locally unfold chromatin domains, altering their transcriptional activity, and modulate the methylation level of histones by methyltransferases and deacetylases, regulating transcriptional activity. SUN, Sad1p, UNC-84.

and the resulting nuclear deformations will have a considerable impact on nucleo-cytosolic transport, including import–export dynamics of important transcription and epigenetic regulators. For instance, nuclear deformations can modulate the balance of nuclear and cytoplasmic pools of two key mechanoresponsive transcription regulators, YAP (Yes-associated protein) and TAZ (transcriptional coactivator with PDZ-binding motif)<sup>200</sup> (FIG. 6c), which have crucial roles in regulating a wide range of key biological processes<sup>201</sup>. In mouse embryonic fibroblasts, mechanical signals from ECM rigidity are transmitted to the nucleus via LINC complexes. These forces cause nuclear envelope stretching, likely opening nuclear pores and promoting nuclear import of YAP<sup>198</sup>. By contrast, during differentiation of myoblasts into myotubes, nuclear elongation (see discussion above) promotes YAP nuclear export to drive cell differentiation<sup>139</sup>. More recently, YAP nuclear export was associated with substrate curvature changes that impose nuclear deformations. Nuclei located on convex zones (that is, crests) were flattened with an elevated nuclear presence of YAP and chromatin was less condensed, whereas nuclei on concave zones (that is, valleys) were highly elongated, contained more condensed chromatin, and YAP was predominantly cytoplasmic<sup>202</sup>. These findings support the notion of a control of YAP/TAZ by nuclear deformations and highlight the

importance of mechanical and cytoskeletal regulation of the nuclear shape in modulating YAP/TAZ signalling. Several lines of evidence indicate that similar effects can be observed by imposing nuclear deformations with higher cell density<sup>203</sup> or various external forces<sup>27,132,198,200</sup>, without changing the mechanical properties of ECM. However, precisely how the intracellular localization of YAP is modulated by nuclear shape and volume changes<sup>198</sup>, and how this observation relates to known regulators of YAP nuclear translocation, remain to be elucidated.

**Mechanically induced genome regulation.** Recent evidence suggests that the cytoskeleton can modify not only the physical state of the nucleus but also the chromatin state and gene expression. For example, local stresses applied to integrins can propagate to the LINC complex through the actin cytoskeleton and lead to chromatin unpacking<sup>204</sup> and epigenetic changes in chromatin (such as H3K9me3 demethylation)<sup>205</sup> that promote force-induced transcription in the nuclear interior. At the nuclear envelope periphery, local cytoskeletal forces, such as actin fibre-based indentation of the nucleus, can severely deform the nuclear envelope and trigger reversible formation of heterochromatin<sup>50,93,206</sup>. Nuclear deformations during confined migration can also induce increased activity of histone methylases and histone deacetylases (HDACs). This results in an increase in H3K9me3 and H3K27me3 heterochromatin marks (FIG. 6d) and promotes cell migration through mechanisms that are yet to be defined<sup>118,207</sup>. This increase in heterochromatin abundance can last from hours to days<sup>207</sup>. In addition to local changes in chromatin architecture and organization, dynamic nuclear deformation could be an underlying driving force of spatiotemporal genomic reorganization. Indeed, suppression of nuclear deformation in the mouse retinal photoreceptors results in impairment of heterochromatin clustering into chromocentres<sup>94</sup>. There is also evidence that confined cell migration leads to rearrangements in 3D genome organization in neutrophils and cancer cells<sup>208,209</sup>.

Chromatin modifications can also arise from changes in the nuclear actin pool. Increased perinuclear actin polymerization, mediated by re-localization of emerin to the ONM in response to nuclear deformations (FIG. 6e), can result in increased facultative heterochromatin formation by depleting monomeric actin from the nucleus, reducing transcription and activating Polycomb repressive complex 2 (PRC2)<sup>103</sup>. Mechanically induced depolymerization of actin can also lead to translocation of HDAC3 from the cytoplasm into the nucleus, resulting in increased heterochromatin formation<sup>210</sup>. Spatial confinement can similarly reduce actin polymerization, thereby reducing nuclear translocation of megakaryoblastic leukaemia 1 protein (MKL1), a mechanoresponsive coactivator of the serum response factor (SRF), which regulates many physiological processes, including pro-inflammatory macrophage differentiation<sup>179</sup>. Likewise, emerin-mediated actin polymerization can modulate nuclear translocation of MKL1 (REF.<sup>211</sup>). Sustained activity of MKL1 results in reduced nuclear volume and globally reduced chromatin accessibility<sup>49</sup>.

#### cGAS–STING DNA-sensing pathway

Cellular cytosolic double-stranded DNA sensor, allowing innate immune response to infections, inflammation and cancer.

#### Micronuclei

Small DNA-containing nuclear structures that are spatially isolated from the main nucleus. Micronuclei form from lagging chromosomes or chromosome fragments following mitotic errors or DNA damage, respectively.

#### Chromocentres

Dense aggregation of heterochromatin formed during interphase.

## Facultative heterochromatin

Condensed, transcriptionally silent chromatin region that can decondense and adapt to allow transcription within temporal and spatial contexts. Facultative heterochromatin is not characterized by repetitive sequences so, at the DNA sequence level, it is entirely different from constitutive heterochromatin.

## Polycomb repressive complex 2

(PRC2). Major repressive chromatin complex formed by Polycomb group (PcG) proteins.

## Serum response factor

(SRF). Transcription factor that plays a key role in the transduction of mechanical signals from cytoplasmic actin and extracellular matrix proteins to the nucleus. SRF is involved in various cellular processes, from cell proliferation to differentiation and development.

Mechanically induced epigenetic changes can have a functional impact on gene expression and cell fate regulation (FIG. 6f). For example, human mesenchymal cells respond to matrix stiffening by increasing nuclear membrane tension and histone acetylation via deactivation of HDACs, leading to osteogenic fate determination<sup>212</sup>. By contrast, LINC complex disruption, which presumably reduces nuclear membrane tension, leads to upregulation of HDACs and inhibits osteogenic differentiation<sup>212</sup>. Similarly, persistent differentiation of fibroblasts to myofibroblasts relies on increased chromatin compaction mediated by nuclear mechanosensing of cytoskeletal forces via LINC complexes (FIG. 6f) that results in increased activity of HDACs<sup>213</sup>. In macrophages, spatial confinement can suppress the acquisition of a pro-inflammatory phenotype and associated transcriptional programmes (for example, expression of IL-6, CXCL9, IL-1 $\beta$  and iNOS) by inducing epigenetic alterations (such as an increase in H3K36me2) and promoting chromatin compaction<sup>179</sup>. In cardiac myocytes, peripheral heterochromatin characterized by H3K9me3 marks, which closely correlates with intranuclear deformations and reducing nuclear deformations by LINC complex disruption, results in loss of peripheral H3K9me2/3 marks and reduced expression of cardiac developmental genes<sup>214</sup>.

The molecular details by which mechanical deformation of the cell and nucleus result in chromatin modification and reorganization remain incompletely understood, but two major contributors have emerged to date: an increase in intracellular cations (calcium and/or magnesium) by activation of stretch-activated ion channels and remodelling of the nuclear and/or perinuclear actin network. Repetitive stretching of mesenchymal stem cells activates mechanosensitive ion channels, such as Piezo1, leading to increased intracellular calcium levels and increased heterochromatin formation (marked by H3K9me2 and H3K9me3), ultimately promoting mesenchymal differentiation<sup>215,216</sup>. In epithelial cells, cyclic mechanical stretch triggers immediate nuclear deformation that leads to Piezo1-mediated calcium release from the ER, reducing lamina-associated heterochromatin (H3K9me3 marks) within a ~30 min window<sup>16</sup>. This results in nuclear softening that decreases stress and DNA damage in the stretched cells<sup>16</sup>. Long-term (8–12 h) cyclic uniaxial stretch application causes transcriptional repression, increased heterochromatin (H3K27me3) and silencing of differentiation gene expression<sup>16</sup>. Intriguingly, activation of mechanosensitive ion channels by increasing extracellular multivalent ion concentrations, even in the absence of cell stretching or compression, is sufficient to trigger a similar increase in heterochromatin<sup>91</sup>. The increased heterochromatin content mechanically strengthened the nucleus, rescued abnormal nuclear morphology in *LMNA*-mutant and breast cancer cells, reduced nuclear envelope ruptures and prevented DNA damage<sup>91</sup>. Collectively, these findings demonstrate that mechanosensitive ion channels respond to mechanical stimuli causing an increase in intracellular calcium that leads to chromatin modifications, which mechanically protect the nucleus and influence cell fate decisions. These stretch-sensitive ion channels can be found on the

plasma membrane, the ER and, potentially, the nuclear envelope itself, with the contribution of specific channels and their locations likely depending on the particular cellular context and the mechanical cue.

## Conclusions and perspectives

Considerable efforts in recent years have started to shed light on the fascinating roles of nuclear deformations in cell function, whereby chromatin organization, compaction, stretching and modifications that arise from nuclear deformations control the downstream expression of genes and cell fate decisions. Altogether, these discoveries have revealed the remarkable mechanoresponsive nature of the nucleus and the key role of nuclear proteins in the cellular response to mechanical stimuli. However, many open questions remain. For example, although potential mechanisms have been proposed (BOX 1), how the nucleus senses the different forces and deformations that it is subject to in different contexts and how it transduces this signal for specific responses remain elusive. Although substantial progress has been made in the understanding of nucleo-cytoskeletal coupling, the precise mechanisms for the spatiotemporal regulation of force transmission across the LINC complex required for many cellular functions has yet to be fully elucidated. Connections between the nucleus, other organelles and the plasma membrane have received far less attention and should be investigated in more detail. Inside the nucleus, a better understanding of the role of nuclear F-actin and associated motor proteins as well as LLPS processes in the maintenance of the nuclear structure, genomic organization and chromatin remodelling will require deeper investigation.

Deciphering the complex mechanical interplay between chromatin, the nuclear envelope, cytoskeletal filaments and the cell surface in mechanobiology will benefit from interdisciplinary and integrative approaches, combining live-cell imaging with high spatial and temporal resolution, genetic manipulation and precise mechanical manipulation. Much of our knowledge about nuclear mechanotransduction has come from innovative technologies. Addressing current challenges in this field will require further technological innovations, for instance, to visualize gene expression in live cells while exerting subcellular deformations, ideally on a genome-wide scale and yet with single-cell resolution. In addition to these experimental breakthroughs, mechanochemical models of the nucleus developed by theoretical modelling will be essential to explore how the cooperation between mechanical and biochemical parameters regulates feedback loops<sup>217</sup> in nuclear signalling pathways. A better understanding of the molecular mechanisms governing nuclear mechanobiology would not only clarify how the various cellular mechanotransduction pathways are combined to determine downstream cellular function but may also guide the development of novel therapeutic strategies to treat human diseases that arise from impaired nuclear mechanics, mechanotransduction and disturbed nucleo-cytoskeletal force transmission (BOX 2).

Published online 5 May 2022

1. Lammerding, J. Mechanics of the nucleus. *Compr. Physiol.* **1**, 783–807 (2013).
2. Szczesny, S. E. & Mauck, R. L. The nuclear option: evidence implicating the cell nucleus in mechanotransduction. *J. Biomech. Eng.* **139**, 0210061–02100616 (2017).
3. Long, J. T. & Lammerding, J. Nuclear deformation lets cells gauge their physical confinement. *Dev. Cell* **56**, 156–158 (2021).
4. Thomas, C. H., Collier, J. H., Sfeir, C. S. & Healy, K. E. Engineering gene expression and protein synthesis by modulation of nuclear shape. *Proc. Natl Acad. Sci. USA* **99**, 1972–1977 (2002).
5. Skinner, B. M. & Johnson, E. E. P. Nuclear morphologies: their diversity and functional relevance. *Chromosoma* **126**, 195–212 (2017).
6. Gupta, S., Marcel, N., Sarin, A. & Shivashankar, G. V. Role of actin dependent nuclear deformation in regulating early gene expression. *PLoS ONE* **7**, e53031 (2012).
7. Miroshnikova, Y. A., Nava, M. M. & Wickström, S. A. Emerging roles of mechanical forces in chromatin regulation. *J. Cell Sci.* **130**, 2243–2250 (2017).
8. Zink, D., Fischer, A. H. & Nickerson, J. A. Nuclear structure in cancer cells. *Nat. Rev. Cancer* **4**, 677–687 (2004).
9. Clippinger, S. R. et al. Disrupted mechanobiology links the molecular and cellular phenotypes in familial dilated cardiomyopathy. *Proc. Natl Acad. Sci. USA* **116**, 17831–17840 (2019).
10. Franke, W. W., Scheer, U., Krohne, G. & Jarasch, E. D. The nuclear envelope and the architecture of the nuclear periphery. *J. Cell Biol.* **91**, 39s–50s (1981).
11. Kim, D.-H. et al. Volume regulation and shape bifurcation in the cell nucleus. *J. Cell Sci.* **128**, 3375–3385 (2015).
12. Jevtić, P. et al. The nucleoporin ELYS regulates nuclear size by controlling NPC number and nuclear import capacity. *EMBO Rep.* **20**, e47283 (2019).
13. García-González, A. et al. The effect of cell morphology on the permeability of the nuclear envelope to diffusive factors. *Front. Physiol.* **9**, 925 (2018).
14. Donnalaja, F., Jacchetti, E., Soncini, M. & Raimondi, M. T. Mechanosensing at the nuclear envelope by nuclear pore complex stretch activation and its effect in physiology and pathology. *Front. Physiol.* **10**, 896 (2019).
15. Schuller, A. P. et al. The cellular environment shapes the nuclear pore complex architecture. *Nature* **598**, 667–671 (2021).
16. Nava, M. M. et al. Heterochromatin-driven nuclear softening protects the genome against mechanical stress-induced damage. *Cell* **181**, 800–817.e22 (2020).  
**This article demonstrates how cyclic strain application can induce transient chromatin modifications, which, together with slower realignment of cells perpendicular to the stretch direction, help protect the cells from mechanically induced DNA damage.**
17. Lomakin, A. J. et al. The nucleus acts as a ruler tailoring cell responses to spatial constraints. *Science* **370**, eaba2894 (2020).  
**This article, together with concurrently published work by Venturini et al. (ref. 127), demonstrates how physical confinement that compresses the nucleus triggers increased cell cortical contractility via recruitment of cPLA2 to the stretched nuclear membranes.**
18. Turgay, Y. et al. The molecular architecture of lamins in somatic cells. *Nature* **543**, 261–264 (2017).
19. Tenga, R. & Medalia, O. Structure and unique mechanical aspects of nuclear lamin filaments. *Curr. Opin. Struct. Biol.* **64**, 152–159 (2020).
20. de Leeuw, R., Gruenbaum, Y. & Medalia, O. Nuclear lamins: thin filaments with major functions. *Trends Cell Biol.* **28**, 34–45 (2018).
21. Shimi, T. et al. Structural organization of nuclear lamins A, C, B1, and B2 revealed by superresolution microscopy. *Mol. Biol. Cell* **26**, 4075–4086 (2015).
22. Kolb, T., Maass, K., Hergt, M., Aebi, U. & Herrmann, H. Lamin A and lamin C form homodimers and coexist in higher complex forms both in the nucleoplasmic fraction and in the lamina of cultured human cells. *Nucleus* **2**, 425–433 (2011).
23. Nmezi, B. et al. Concentric organization of A- and B-type lamins predicts their distinct roles in the spatial organization and stability of the nuclear lamina. *Proc. Natl Acad. Sci. USA* **116**, 4307–4315 (2019).
24. Naetar, N. et al. LAP2alpha maintains a mobile and low assembly state of A-type lamins in the nuclear interior. *eLife* **10**, e63476 (2021).
25. Pascual-Reguant, L. et al. Lamin B1 mapping reveals the existence of dynamic and functional euchromatin lamin B1 domains. *Nat. Commun.* **9**, 3420 (2018).
26. Cho, S. et al. Mechanosensing by the lamina protects against nuclear rupture, DNA damage, and cell-cycle arrest. *Dev. Cell* **49**, 920–935 (2019).
27. Koushki, N. et al. Lamin A redistribution mediated by nuclear deformation determines dynamic localization of YAP. *bioRxiv* <https://doi.org/10.1101/2020.03.19.998708> (2020).
28. Zhu, L. & Brangwynne, C. P. Nuclear bodies: the emerging biophysics of nucleoplasmic phases. *Curr. Opin. Cell Biol.* **34**, 23–30 (2015).
29. Buchwalter, A., Kaneshiro, J. M. & Hetzer, M. W. Coaching from the sidelines: the nuclear periphery in genome regulation. *Nat. Rev. Genet.* **20**, 39–50 (2019).
30. Bizhanova, A. & Kaufman, P. Close to the edge: heterochromatin at the nucleolar and nuclear peripheries. *Biochim. Biophys. Acta* **1864**, 194666 (2021).
31. Miron, E. et al. Chromatin arrangements in chains of mesoscale domains with nanoscale functional topography independent of cohesin. *Sci. Adv.* **6**, eaba8811 (2020).
32. Cho, S., Irianto, J. & Discher, D. E. Mechanosensing by the nucleus: from pathways to scaling relationships. *J. Cell Biol.* **216**, 305–315 (2017).
33. Maeshima, K., Tamura, S., Hansen, J. C. & Itoh, Y. Fluid-like chromatin: toward understanding the real chromatin organization present in the cell. *Curr. Opin. Cell Biol.* **64**, 77–89 (2020).
34. Hansen, J. C., Maeshima, K. & Hendzel, M. J. The solid and liquid states of chromatin. *Epigenetics Chromatin* **14**, 50 (2021).
35. Jahed, Z., Domkam, N., Ornowski, J., Yerima, G. & Mofrad, M. R. K. Molecular models of LINC complex assembly at the nuclear envelope. *J. Cell Sci.* **134**, jcs258194 (2021).
36. Lombardi, M. L. et al. The interaction between nesprins and SUN proteins at the nuclear envelope is critical for force transmission between the nucleus and cytoskeleton. *J. Biol. Chem.* **286**, 26743–26753 (2011).  
**First function demonstration of force transmission from the cytoskeleton to the nucleus via the LINC complex.**
37. Denis, K. B. et al. The LINC complex is required for endothelial cell adhesion and adaptation to shear stress and cyclic stretch. *Mol. Biol. Cell* **32**, 1654–1663 (2021).
38. Splinter, D. et al. Bicaudal D2, Tynein, and kinesin-1 associate with nuclear pore complexes and regulate centrosome and nuclear positioning during mitotic entry. *PLoS Biol.* **8**, e1000350 (2010).
39. Salpingidou, G., Smertenko, A., Hausmanowa-Petrucewicz, I., Hussey, P. J. & Hutchison, C. J. A novel role for the nuclear membrane protein emerin in association of the centrosome to the outer nuclear membrane. *J. Cell Biol.* **178**, 897–904 (2007).
40. Sosa, B. A., Rothbaler, A., Kutay, U. & Schwartz, T. U. LINC complexes form by binding of three KASH peptides to domain interfaces of trimeric SUN proteins. *Cell* **149**, 1035–1047 (2012).  
**First detailed structural characterization of the LINC complex, elucidating how forces can be transmitted across the nesprin–SUN protein interface.**
41. Cruz, V. E., Demircioglu, F. E. & Schwartz, T. U. Structural analysis of different LINC complexes reveals distinct binding modes. *J. Mol. Biol.* **432**, 6028–6041 (2020).
42. Rajgor, D. & Shanahan, C. M. Nesprins: from the nuclear envelope and beyond. *Expert Rev. Mol. Med.* **15**, e5 (2013).
43. Wilson, M. H. & Holzbaier, E. L. F. Nesprins anchor kinesin-1 motors to the nucleus to drive nuclear distribution in muscle cells. *Development* **142**, 218–228 (2014).
44. Fridolfsson, H. N., Ly, N., Meyerzon, M. & Starr, D. A. UNC-83 coordinates kinesin-1 and dynein activities at the nuclear envelope during nuclear migration. *Dev. Biol.* **338**, 237–250 (2010).
45. Wilhelmens, K. et al. Nesprin-3, a novel outer nuclear membrane protein, associates with the cytoskeletal linker protein plectin. *J. Cell Biol.* **171**, 799–810 (2005).
46. Roux, K. J. et al. Nesprin 4 is an outer nuclear membrane protein that can induce kinesin-mediated cell polarization. *Proc. Natl Acad. Sci. USA* **106**, 2194–2199 (2009).
47. Horn, H. F. et al. A mammalian KASH domain protein coupling meiotic chromosomes to the cytoskeleton. *J. Cell Biol.* **202**, 1023–1039 (2013).
48. Agrawal, A. & Lele, T. P. Mechanics of nuclear membranes. *J. Cell Sci.* **132**, jcs229245 (2019).
49. Hoffman, L. M. et al. Mechanical stress triggers nuclear remodeling and the formation of transmembrane actin nuclear lines with associated nuclear pore complexes. *Mol. Biol. Cell* **31**, 1774–1787 (2020).
50. Versaevael, M. et al. Super-resolution microscopy reveals LINC complex recruitment at nuclear indentation sites. *Sci. Rep.* **4**, 7362 (2015).
51. Davidson, P. M. et al. Nesprin-2 accumulates at the front of the nucleus during confined cell migration. *EMBO Rep.* **21**, e49910 (2020).
52. Lim, S. M., Cruz, V. E., Antoku, S., Gundersen, G. G. & Schwartz, T. U. Structures of FHOD1-nesprin1/2 complexes reveal alternate binding modes for the FH3 domain of formins. *Structure* **29**, 540–552 (2021).
53. Saunders, C. A. et al. TorsinA controls TAN line assembly and the retrograde flow of dorsal perinuclear actin cables during rearward nuclear movement. *J. Cell Biol.* **216**, 657–674 (2017).
54. Gudise, S., Figueroa, R. A., Lindberg, R., Larsson, V. & Hallberg, E. Samp1 is functionally associated with the LINC complex and A-type lamina networks. *J. Cell Sci.* **124**, 2077–2085 (2011).
55. Zverger, M. et al. Myopathic lamin mutations impair nuclear stability in cells and tissue and disrupt nucleo-cytoskeletal coupling. *Hum. Mol. Genet.* **22**, 2335–2349 (2013).
56. Hao, H. et al. The nesprin-1/2 ortholog ANC-1 regulates organelle positioning in *C. elegans* independently from its KASH or actin-binding domains. *eLife* **10**, e61069 (2021).
57. Versaevael, M., Riaz, M., Grevesse, T. & Gabriele, S. Cell confinement: putting the squeeze on the nucleus. *Soft Matter* **9**, 6665–6676 (2013).
58. Guilak, F., Tedrow, J. R. & Burgkart, R. Viscoelastic properties of the cell nucleus. *Biochem. Biophys. Res. Commun.* **269**, 781–786 (2000).
59. Hobson, C. M., Falvo, M. R. & Superfine, R. A survey of physical methods for studying nuclear mechanics and mechanobiology. *Appl. Bioeng.* **5**, 041508 (2021).
60. Pajerowski, J. D., Dahl, K. N., Zhong, F. L., Sammak, P. J. & Discher, D. E. Physical plasticity of the nucleus in stem cell differentiation. *Proc. Natl Acad. Sci. USA* **104**, 15619–15624 (2007).
61. Davidson, P. M. et al. High-throughput microfluidic micropipette aspiration device to probe time-scale dependent nuclear mechanics in intact cells. *Lab. Chip* **19**, 3652–3663 (2019).
62. Rowat, A. C., Foster, L. J., Nielsen, M. M., Weiss, M. & Ipsen, J. H. Characterization of the elastic properties of the nuclear envelope. *J. R. Soc. Interface* **2**, 63–69 (2005).
63. Dahl, K. N., Engler, A. J., Pajerowski, J. D. & Discher, D. E. Power-law rheology of isolated nuclei with deformation mapping of nuclear substructures. *Biophys. J.* **89**, 2855–2864 (2005).
64. Stephens, A. D., Banigan, E. J., Adam, S. A., Goldman, R. D. & Marko, J. F. Chromatin and lamin A determine two different mechanical response regimes of the cell nucleus. *Mol. Biol. Cell* **28**, 1984–1996 (2017).
65. Grevesse, T., Dabiri, B. E., Parker, K. K. & Gabriele, S. Opposite rheological properties of neuronal microcompartments predict axonal vulnerability in brain injury. *Sci. Rep.* **5**, 9475 (2015).
66. Lammerding, J. et al. Lamins A and C but not lamin B1 regulate nuclear mechanics. *J. Biol. Chem.* **281**, 25768–25780 (2006).  
**This report establishes the important role of lamins A/C in determining nuclear deformability.**
67. Neelam, S. et al. Direct force probe reveals the mechanics of nuclear homeostasis in the mammalian cell. *Proc. Natl Acad. Sci. USA* **112**, 5720–5725 (2015).
68. Lammerding, J. et al. Lamin A/C deficiency causes defective nuclear mechanics and mechanotransduction. *J. Clin. Invest.* **113**, 370–378 (2004).  
**First report of a role of lamins A/C in mediating nuclear mechanotransduction and providing nuclear stability to allow cells to withstand mechanical stress.**
69. Lammerding, J. et al. Abnormal nuclear shape and impaired mechanotransduction in emerin-deficient cells. *J. Cell Biol.* **170**, 781–791 (2005).



70. Cao, X. et al. A chemomechanical model for nuclear morphology and stresses during cell transendothelial migration. *Biophys. J.* **111**, 1541–1552 (2016).
71. Versaevael, M. et al. Probing cytoskeletal pre-stress and nuclear mechanics in endothelial cells with spatiotemporally controlled (de-)adhesion kinetics on micropatterned substrates. *Cell Adhes. Migr.* **11**, 98–109 (2017).
72. Ferrera, D. et al. Lamin B1 overexpression increases nuclear rigidity in autosomal dominant leukodystrophy fibroblasts. *FASEB J.* **28**, 3906–3918 (2014).
73. Chen, N. Y. et al. An absence of lamin B1 in migrating neurons causes nuclear membrane ruptures and cell death. *Proc. Natl Acad. Sci. USA* **116**, 25870–25879 (2019).
74. Hatch, E. M. & Hetzer, M. W. Nuclear envelope rupture is induced by actin-biased nucleus confinement. *J. Cell Biol.* **215**, 27–36 (2016).
75. Raab, M. et al. ESCRT III repairs nuclear envelope ruptures during cell migration to limit DNA damage and cell death. *Science* **352**, 359–362 (2016).
76. Denais, C. M. et al. Nuclear envelope rupture and repair during cancer cell migration. *Science* **352**, 353–358 (2016).
- Together with the work by Raab et al. (ref. 75), this is the first report of migration-induced nuclear envelope rupture and DNA damage, while also identifying a role of ESCRT proteins in interphase nuclear envelope resealing.**
77. Zhang, Q. et al. Local, transient tensile stress on the nuclear membrane causes membrane rupture. *Mol. Biol. Cell* **30**, 899–906 (2019).
78. Furusawa, T. et al. Chromatin decompaction by the nucleosomal binding protein HMGN5 impairs nuclear sturdiness. *Nat. Commun.* **6**, 6138 (2015).
79. Samwer, M. et al. DNA cross-bridging shapes a single nucleus from a set of mitotic chromosomes. *Cell* **170**, 956–972 (2017).
80. Wang, P. et al. WDR5 modulates cell motility and morphology and controls nuclear changes induced by a 3D environment. *Proc. Natl Acad. Sci. USA* **115**, 8581–8586 (2018).
81. Serra-Marques, A. et al. The mitotic protein NuMA plays a spindle-independent role in nuclear formation and mechanics. *J. Cell Biol.* **219**, e202004202 (2020).
82. Tamashunas, A. C. et al. High-throughput gene screen reveals modulators of nuclear shape. *Mol. Biol. Cell* **31**, 1392–1402 (2020).
83. Larson, A. G. et al. Liquid droplet formation by HP1 $\alpha$  suggests a role for phase separation in heterochromatin. *Nature* **547**, 236–240 (2017).
84. Welsh, T. J., Shen, Y., Levin, A. & Knowles, T. P. J. Mechanobiology of protein droplets: force arises from disorder. *Cell* **175**, 1457–1459 (2018).
85. Gibson, B. A. et al. Organization of chromatin by intrinsic and regulated phase separation. *Cell* **179**, 470–484 (2019).
86. Zidovska, A. The rich inner life of the cell nucleus: dynamic organization, active flows, and emergent rheology. *Biophys. Rev.* **12**, 1093–1106 (2020).
87. Jord, A. A. et al. Cytoplasmic forces functionally reorganize nuclear condensates in oocytes. *bioRxiv* <https://doi.org/10.1101/2021.03.15.434387> (2021).
88. Bracha, D. et al. Mapping local and global liquid phase behavior in living cells using photo-oligomerizable seeds. *Cell* **175**, 1467–1480 (2018).
89. Shin, Y. et al. Spatiotemporal control of intracellular phase transitions using light-activated optoDroplets. *Cell* **168**, 159–171 (2017).
90. Shin, Y. et al. Liquid nuclear condensates mechanically sense and restructure the genome. *Cell* **175**, 1481–1491 (2018).
91. Stephens, A. D. et al. Physicochemical mechanotransduction alters nuclear shape and mechanics via heterochromatin formation. *Mol. Biol. Cell* **30**, 2320–2330 (2019).
92. Cantwell, H. & Nurse, P. Unravelling nuclear size control. *Curr. Genet.* **65**, 1281–1285 (2019).
93. Versaevael, M., Grevesse, T. & Gabriele, S. Spatial coordination between cell and nuclear shape within micropatterned endothelial cells. *Nat. Commun.* **3**, 671 (2012).
- This work demonstrates that actomyosin stress fibres regulate nuclear deformations in response to cell shape changes and report a drastic condensation of chromatin in deformed nuclei.**
94. Seirin-Lee, S. et al. Role of dynamic nuclear deformation on genomic architecture reorganization. *PLoS Comput. Biol.* **15**, e1007289 (2019).
95. Alisafaei, F., Jokhun, D. S., Shivashankar, G. V. & Shenoy, V. B. Regulation of nuclear architecture, mechanics, and nucleocytoplasmic shuttling of epigenetic factors by cell geometric constraints. *Proc. Natl Acad. Sci. USA* **116**, 13200–13209 (2019).
96. Petrie, R. J., Koo, H. & Yamada, K. M. Generation of compartmentalized pressure by a nuclear piston governs cell motility in a 3D matrix. *Science* **345**, 1062–1065 (2014).
97. Mistriotis, P. et al. Confinement hinders motility by inducing RhoA-mediated nuclear influx, volume expansion, and blebbing. *J. Cell Biol.* **218**, 4093–4111 (2019).
98. Mitchison, T. J. Colloid osmotic parameterization and measurement of subcellular crowding. *Mol. Biol. Cell* **30**, 173–180 (2019).
99. Deviri, D. & Safran, S. A. Balance of osmotic pressures determines the volume of the cell nucleus. *bioRxiv* <https://doi.org/10.1101/2021.10.01.462771v1> (2021).
100. Lemi re, J., Real-Calderon, P., Holt, L. J., Fai, T. G. & Chang, F. Control of nuclear size by osmotic forces in *Schizosaccharomyces pombe*. *bioRxiv* <https://doi.org/10.1101/2021.12.05.471221> (2021).
101. Takata, H. et al. Chromatin compaction protects genomic DNA from radiation damage. *PLoS ONE* **8**, e75622 (2013).
102. Holaska, J. M., Kowalski, A. K. & Wilson, K. L. Emerin caps the pointed end of actin filaments: evidence for an actin cortical network at the nuclear inner membrane. *PLoS Biol.* **2**, e231 (2004).
103. Le, H. Q. et al. Mechanical regulation of transcription controls Polycomb-mediated gene silencing during lineage commitment. *Nat. Cell Biol.* **18**, 864–875 (2016).
- This study identifies how mechanical stretch can result in emerin translocation to the ONM, where it facilitates perinuclear actin polymerization that results in depletion of intranuclear actin and changes in chromatin organization.**
104. Davidson, P. M. & Cadot, B. Actin on and around the nucleus. *Trends Cell Biol.* **31**, 211–223 (2020).
105. Buxboim, A. et al. Matrix elasticity regulates lamin-a, c phosphorylation and turnover with feedback to actomyosin. *Curr. Biol.* **24**, 1909–1917 (2014).
106. Mattout, A. et al. An EDMD mutation in *C. elegans* lamin blocks muscle-specific gene relocation and compromises muscle integrity. *Curr. Biol.* **21**, 1603–1614 (2011).
107. Solovei, I. et al. LBR and lamin A/C sequentially tether peripheral heterochromatin and inversely regulate differentiation. *Cell* **152**, 584–598 (2013).
108. Shin, J.-W. et al. Lamins regulate cell trafficking and lineage maturation of adult human hematopoietic cells. *Proc. Natl Acad. Sci. USA* **110**, 18892–18897 (2013).
109. Swift, J. et al. Nuclear lamin-a scales with tissue stiffness and enhances matrix-directed differentiation. *Science* **341**, 1240104 (2013).
- Detailed proteomic analysis linking higher levels of lamins A/C (and collagen) to cells residing in stiff tissues, suggesting a mechano-adaptive role of lamin A/C expression.**
110. Zuela, N., Dorfman, J. & Gruenbaum, Y. Global transcriptional changes caused by an EDMD mutation correlate to tissue specific disease phenotypes in *C. elegans*. *Nucleus* **8**, 60–69 (2017).
111. Iyer, K. V. et al. Apico-basal cell compression regulates lamin A/C levels in epithelial tissues. *Nat. Commun.* **12**, 1756 (2021).
112. Olins, D. E. & Olins, A. L. Granulocyte heterochromatin: defining the epigenome. *BMC Cell Biol.* **6**, 39 (2005).
113. Rowat, A. C. et al. Nuclear envelope composition determines the ability of neutrophil-type cells to passage through micron-scale constrictions. *J. Biol. Chem.* **288**, 8610–8618 (2013).
- One of the first reports providing functional evidence that nuclear envelope composition and deformability determine the ability of cells to transit through tight constrictions.**
114. Roberts, A. B. et al. Tumor cell nuclei soften during transendothelial migration. *J. Biomech.* **121**, 110400 (2021).
115. Bell, E. S. et al. Low lamin A levels enhance confined cell migration and metastatic capacity in breast cancer. *bioRxiv* <https://doi.org/10.1101/2021.07.12.451842v1> (2021).
116. Infante, E. et al. LINC complex–Lis1 interplay controls MT1–MMP matrix digest-on-demand response for confined tumor cell migration. *Nat. Commun.* **9**, 2443 (2018).
117. Krause, M. et al. Cell migration through three-dimensional confining pores: speed accelerations by deformation and recoil of the nucleus. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **374**, 20180225 (2019).
118. Gerlitz, G. The emerging roles of heterochromatin in cell migration. *Front. Cell Dev. Biol.* **8**, 394 (2020).
119. Earle, A. J. et al. Mutant lamins cause nuclear envelope rupture and DNA damage in skeletal muscle cells. *Nat. Mater.* **19**, 464–473 (2020).
120. Mitchell, M. J. et al. Lamin A/C deficiency reduces circulating tumor cell resistance to fluid shear stress. *Am. J. Physiol. Cell Physiol.* **309**, C736–C746 (2015).
121. Gundersen, G. G. & Worman, H. J. Nuclear positioning. *Cell* **152**, 1376–1389 (2013).
122. Roman, W. & Gomes, E. R. Nuclear positioning in skeletal muscle. *Semin. Cell Dev. Biol.* **82**, 51–56 (2018).
123. Collins, M. A. et al. Enscosin-dependent changes in microtubule organization and LINC complex-dependent changes in nucleus-nucleus interactions result in quantitatively distinct myonuclear positioning defects. *Mol. Biol. Cell* **32**, ar27 (2021).
124. Lorber, D., Rotkopf, R. & Volk, T. A minimal constraint device for imaging nuclei in live *Drosophila* contractile larval muscles reveals novel nuclear mechanical dynamics. *Lab. Chip* **20**, 2100–2112 (2020).
125. Davidson, P. M., Denais, C., Bakshi, M. C. & Lammerding, J. Nuclear deformability constitutes a rate-limiting step during cell migration in 3-D environments. *Cell. Mol. Bioeng.* **7**, 293–306 (2014).
- This report, along with work by Harada et al. (ref. 175) and Wolf et al. (ref. 126), presents some of the first evidence that increased nuclear deformability caused by reduced lamin A/C expression enhances cell migration through confined environments.**
126. Wolf, K. et al. Physical limits of cell migration: control by ECM space and nuclear deformation and tuning by proteolysis and traction force. *J. Cell Biol.* **201**, 1069–1084 (2013).
- This study demonstrates that nuclear deformability presents a rate-limiting factor in the ability of cells to migrate through constrictions smaller than ~10% of the undeformed nuclear cross-section.**
127. Venturini, V. et al. The nucleus measures shape changes for cellular proprioception to control dynamic cell behavior. *Science* **370**, eaba2644 (2020).
- As presented in work by Lomakin et al. (ref. 17), it was shown that the nuclear envelope provides a gauge of cell deformation and activates a mechanotransduction pathway that controls actomyosin contractility via mechanically induced recruitment of cPLA2 to the INM.**
128. Smith, E. R. et al. Nuclear envelope structural proteins facilitate nuclear shape changes accompanying embryonic differentiation and fidelity of gene expression. *BMC Cell Biol.* **18**, 8 (2017).
129. Spear, P. C. & Erickson, C. A. Interkinetic nuclear migration: a mysterious process in search of a function. *Dev. Growth Differ.* **54**, 306–316 (2012).
130. Tsukamoto, S. et al. Compressive forces driven by lateral actin fibers are a key to the nuclear deformation under uniaxial cell-substrate stretching. *Biochem. Biophys. Res. Commun.* **597**, 37–43 (2022).
131. Alam, S. G. et al. The nucleus is an intracellular propagator of tensile forces in NIH 3T3 fibroblasts. *J. Cell Sci.* **128**, 1901–1911 (2015).
132. Aureille, J. et al. Nuclear envelope deformation controls cell cycle progression in response to mechanical force. *EMBO Rep.* **20**, e48084 (2019).
133. Lammerding, J. & Wolf, K. Nuclear envelope rupture: actin fibers are putting the squeeze on the nucleus. *J. Cell Biol.* **215**, 5–8 (2016).
134. Hatch, E. M. Nuclear envelope rupture: little holes, big openings. *Curr. Opin. Cell Biol.* **52**, 66–72 (2018).
135. Khatau, S. B. et al. A perinuclear actin cap regulates nuclear shape. *Proc. Natl Acad. Sci. USA* **106**, 19017–19022 (2009).
136. Lovett, D. B., Shekhar, N., Nickerson, J. A., Roux, K. J. & Lele, T. P. Modulation of nuclear shape by substrate rigidity. *Cell. Mol. Bioeng.* **6**, 230–238 (2013).
137. Gupta, M. et al. Adaptive rheology and ordering of cell cytoskeleton govern matrix rigidity sensing. *Nat. Commun.* **6**, 7525 (2015).
138. Nagayama, K., Yahiro, Y. & Matsumoto, T. Apical and basal stress fibers have different roles in mechanical regulation of the nucleus in smooth muscle cells cultured on a substrate. *Cell. Mol. Bioeng.* **6**, 473–481 (2013).



139. Bruyère, C. et al. Actomyosin contractility scales with myoblast elongation and enhances differentiation through YAP nuclear export. *Sci. Rep.* **9**, 15565 (2019).
140. Azevedo, M. & Baylies, M. K. Getting into position: nuclear movement in muscle cells. *Trends Cell Biol.* **30**, 303–316 (2020).
141. Roman, W. et al. Muscle repair after physiological damage relies on nuclear migration for cellular reconstruction. *Science* **374**, 355–359 (2021).
142. Gimpel, P. et al. Nesprin-1 $\alpha$ -dependent microtubule nucleation from the nuclear envelope via Akap450 is necessary for nuclear positioning in muscle cells. *Curr. Biol.* **27**, 2999–3009.e9 (2017).
143. Roman, W. et al. Myofibril contraction and crosslinking drive nuclear movement to the periphery of skeletal muscle. *Nat. Cell Biol.* **19**, 1189–1201 (2017).
144. Wu, Y. K., Umeshima, H., Kurisu, J. & Kengaku, M. Nesprins and opposing microtubule motors generate a point force that drives directional nuclear motion in migrating neurons. *Development* **145**, dev158782 (2018).
145. Picariello, H. S. et al. Myosin IIA suppresses glioblastoma development in a mechanically sensitive manner. *Proc. Natl Acad. Sci. USA* **116**, 15550–15559 (2019).
146. Vargas, J. D., Hatch, E. M., Anderson, D. J. & Hetzer, M. W. Transient nuclear envelope rupturing during interphase in human cancer cells. *Nucleus* **3**, 88–100 (2012).
147. Chai, R. J. et al. Disrupting the LINC complex by AAV mediated gene transduction prevents progression of lamin induced cardiomyopathy. *Nat. Commun.* **12**, 4722 (2021).
- First report showing that LINC complex disruption can improve disease progression in a laminopathy mouse model of dilated cardiomyopathy.**
148. Piccus, R. & Brayson, D. The nuclear envelope: LINCing tissue mechanics to genome regulation in cardiac and skeletal muscle. *Biol. Lett.* **16**, 20200302 (2020).
149. Razafsky, D., Potter, C. & Hodzic, D. Validation of a mouse model to disrupt LINC complexes in a cell-specific manner. *J. Vis. Exp.* **106**, e53318 (2015).
150. Hampoelz, B. et al. Microtubule-induced nuclear envelope fluctuations control chromatin dynamics in *Drosophila* embryos. *Development* **138**, 3377–3386 (2011).
151. Schulte, S. R. et al. A comparative study of *Drosophila* and human A-type lamins. *PLoS ONE* **4**, e7564 (2009).
152. Bone, C. R., Chang, Y.-T., Cain, N. E., Murphy, S. P. & Starr, D. A. Nuclei migrate through constricted spaces using microtubule motors and actin networks in *C. elegans* hypodermal cells. *Development* **143**, 4193–4202 (2016).
153. Driver, E. C., Northrop, A. & Kelley, M. W. Cell migration, intercalation and growth regulate mammalian cochlear extension. *Development* **144**, 3766–3776 (2017).
154. Mohammed, D. et al. Substrate area confinement is a key determinant of cell velocity in collective migration. *Nat. Phys.* **15**, 858–866 (2019).
155. Yanakieva, I., Erzberger, A., Matejić, M., Modes, C. D. & Norden, C. Cell and tissue morphology determine actin-dependent nuclear migration mechanisms in neuroepithelia. *J. Cell Biol.* **218**, 3272–3289 (2019).
156. Norden, C., Young, S., Link, B. A. & Harris, W. A. Actomyosin is the main driver of interkinetic nuclear migration in the retina. *Cell* **138**, 1195–1208 (2009).
157. Tsai, L.-H. & Gleeson, J. G. Nucleokinesis in neuronal migration. *Neuron* **46**, 383–388 (2005).
158. Cooper, J. A. Mechanisms of cell migration in the nervous system. *J. Cell Biol.* **202**, 725–734 (2013).
159. Young, S. G., Jung, H.-J., Lee, J. M. & Fong, L. G. Nuclear lamins and neurobiology. *Mol. Cell Biol.* **34**, 2776–2785 (2014).
160. Wolf, K. et al. Collagen-based cell migration models in vitro and in vivo. *Semin. Cell Dev. Biol.* **20**, 931–941 (2009).
161. Yamada, K. M. & Sixt, M. Mechanisms of 3D cell migration. *Nat. Rev. Mol. Cell Biol.* **20**, 738–752 (2019).
162. Renkawitz, J. et al. Nuclear positioning facilitates amoeboid migration along the path of least resistance. *Nature* **568**, 546–550 (2019).
163. Maciejowski, J. & Hatch, E. M. Nuclear membrane rupture and its consequences. *Annu. Rev. Cell Dev. Biol.* **36**, 85–114 (2020).
164. Thiam, H.-R. et al. Perinuclear Arp2/3-driven actin polymerization enables nuclear deformation to facilitate cell migration through complex environments. *Nat. Commun.* **7**, 10997 (2016).
165. Fridolfsson, H. N. & Starr, D. A. Kinesin-1 and dynein at the nuclear envelope mediate the bidirectional migrations of nuclei. *J. Cell Biol.* **191**, 115–128 (2010).
166. Marks, P. C. & Petrie, R. J. Push or pull: how cytoskeletal crosstalk facilitates nuclear movement through 3D environments. *Phys. Biol.* **19**, 021003 (2022).
167. de Noronha, C. M. C. et al. Dynamic disruptions in nuclear envelope architecture and integrity induced by HIV-1 Vpr. *Science* **294**, 1105–1108 (2001).
168. Vos, W. H. D. et al. Repetitive disruptions of the nuclear envelope invoke temporary loss of cellular compartmentalization in laminopathies. *Hum. Mol. Genet.* **20**, 4175–4186 (2011).
- First report of spontaneous nuclear envelope rupture in laminopathy cells.**
169. Srivastava, N. et al. Nuclear fragility, blaming the blebs. *Curr. Opin. Cell Biol.* **70**, 100–108 (2021).
170. Pfeifer, C. R. et al. Gaussian curvature dilutes the nuclear lamina, favoring nuclear rupture, especially at high strain rate. *Nucleus* **13**, 129–143 (2022).
171. Pfeifer, C. R., Vashisth, M., Xia, Y. & Discher, D. E. Nuclear failure, DNA damage, and cell cycle disruption after migration through small pores: a brief review. *Essays Biochem.* **63**, 569–577 (2019).
172. Goldman, R. D. et al. Accumulation of mutant lamin A causes progressive changes in nuclear architecture in Hutchinson-Gilford progeria syndrome. *Proc. Natl Acad. Sci. USA* **101**, 8963–8968 (2004).
173. Muchir, A. et al. Nuclear envelope alterations in fibroblasts from patients with muscular dystrophy, cardiomyopathy, and partial lipodystrophy carrying lamin A/C gene mutations. *Muscle Nerve* **30**, 444–450 (2004).
174. Karoutas, A. et al. The NSL complex maintains nuclear architecture stability via lamin A/C acetylation. *Nat. Cell Biol.* **21**, 1248–1260 (2019).
175. Harada, T. et al. Nuclear lamin stiffness is a barrier to 3D migration, but softness can limit survival. *J. Cell Biol.* **204**, 669–682 (2014).
- Study demonstrating that increased nuclear deformability caused by reduced lamin A/C expression enhances cell migration through confined environments but renders cells more susceptible to mechanically induced damage and cell death.**
176. Rowat, A. C., Lammerding, J. & Ipsen, J. H. Mechanical properties of the cell nucleus and the effect of emerin deficiency. *Biophys. J.* **91**, 4649–4664 (2006).
177. Berre, M. L., Aubertin, J. & Piel, M. Fine control of nuclear confinement identifies a threshold deformation leading to lamina rupture and induction of specific genes. *Integr. Biol.* **4**, 1406–1414 (2012).
178. Takaki, T. et al. Actomyosin drives cancer cell nuclear dysmorphia and threatens genome stability. *Nat. Commun.* **8**, 16013 (2017).
179. Jain, N. & Vogel, V. Spatial confinement downsizes the inflammatory response of macrophages. *Nat. Mater.* **17**, 1134–1144 (2018).
180. Isermann, P. & Lammerding, J. Consequences of a tight squeeze: nuclear envelope rupture and repair. *Nucleus* **8**, 268–274 (2017).
181. Deviri, D. et al. Scaling laws indicate distinct nucleation mechanisms of holes in the nuclear lamina. *Nat. Phys.* **15**, 823–829 (2019).
182. Irianto, J. et al. DNA damage follows repair factor depletion and portends genome variation in cancer cells after pore migration. *Curr. Biol.* **27**, 210–223 (2017).
- This study reports increasing DNA damage and chromosomal abnormalities in tumour cells after repeated migration through small constrictions.**
183. Halfmann, C. T. et al. Repair of nuclear ruptures requires barrier-to-autointegration factor. *J. Cell Biol.* **218**, 2136–2149 (2019).
184. Penfield, L. et al. Dynein pulling forces counteract lamin-mediated nuclear stability during nuclear envelope repair. *Mol. Biol. Cell* **29**, 852–868 (2018).
185. Young, A. M., Gunn, A. L. & Hatch, E. M. BAF facilitates interphase nuclear membrane repair through recruitment of nuclear transmembrane proteins. *Mol. Biol. Cell* **31**, 1551–1560 (2020).
186. Nader, G. P. et al. Compromised nuclear envelope integrity drives TREX1-dependent DNA damage and tumor cell invasion. *Cell* **184**, 5230–5246 (2021).
187. Shah, P. et al. Nuclear deformation causes DNA damage by increasing replication stress. *Curr. Biol.* **31**, 753–765 (2021).
188. Kidiyoor, G. R. et al. ATR is essential for preservation of cell mechanics and nuclear integrity during interstitial migration. *Nat. Commun.* **11**, 4828 (2020).
189. Jiang, Y. N. et al. Interleukin 6-triggered ataxia-telangiectasia mutated kinase activation facilitates epithelial-to-mesenchymal transition in lung cancer by upregulating vimentin expression. *Exp. Cell Res.* **381**, 165–171 (2019).
190. Peng, B., Ortega, J., Gu, L., Chang, Z. & Li, G.-M. Phosphorylation of proliferating cell nuclear antigen promotes cancer progression by activating the ATM/Akt/GSK3 $\beta$ /Snail signaling pathway. *J. Biol. Chem.* **294**, 7037–7045 (2019).
191. Bakhoun, S. F. et al. Chromosomal instability drives metastasis through a cytosolic DNA response. *Nature* **553**, 467–472 (2018).
192. Enyedi, B., Jelcic, M. & Niethammer, P. The cell nucleus serves as a mechanotransducer of tissue damage-induced inflammation. *Cell* **165**, 1160–1170 (2016).
- First demonstration that increased nuclear membrane tension can trigger recruitment of cPLA2 to the, where it can induce further downstream signalling.**
193. Shen, Z. et al. A synergy between mechanosensitive calcium- and membrane-binding mediates tension-sensing by C2-like domains. *Proc. Natl Acad. Sci. USA* **119**, e2112390119 (2022).
194. Niethammer, P. Components and mechanisms of nuclear mechanotransduction. *Annu. Rev. Cell Dev. Biol.* **37**, 233–256 (2021).
195. Shen, Z. & Niethammer, P. A cellular sense of space and pressure. *Science* **370**, 295–296 (2020).
196. Katayama, T. et al. Stimulatory effects of arachidonic acid on myosin ATPase activity and contraction of smooth muscle via myosin motor domain. *Am. J. Physiol. Heart Circ. Physiol.* **298**, H505–H514 (2010).
197. Brown, M., Roulson, J.-A., Hart, C. A., Tawadros, T. & Clarke, N. W. Arachidonic acid induction of Rho-mediated transendothelial migration in prostate cancer. *Br. J. Cancer* **110**, 2099–2108 (2014).
198. Elosegui-Artola, A. et al. Force triggers YAP nuclear entry by regulating transport across nuclear pores. *Cell* **171**, 1397–1410 (2017).
- First report of mechanically induced opening of NPCs, mediating import of the mechanoresponsive transcription factor YAP.**
199. Zimmerli, C. E. et al. Nuclear pores dilate and constrict in cellulose. *Science* **374**, eabd9776 (2021).
200. Driscoll, T. P., Cosgrove, B. D., Heo, S.-J., Shurden, Z. E. & Mauck, R. L. Cytoskeletal to nuclear strain transfer regulates YAP signaling in mesenchymal stem cells. *Biophys. J.* **108**, 2783–2793 (2015).
201. Moya, I. M. & Halder, G. Hippo-YAP/TAZ signalling in organ regeneration and regenerative medicine. *Nat. Rev. Mol. Cell Biol.* **20**, 211–226 (2019).
202. Luciano, M. et al. Cell monolayers sense curvature by exploiting active mechanics and nuclear mechanoadaptation. *Nat. Phys.* **17**, 1382–1390 (2021).
203. Aragona, M. et al. A mechanical checkpoint controls multicellular growth through YAP/TAZ regulation by actin-processing factors. *Cell* **154**, 1047–1059 (2013).
204. Tajik, A. et al. Transcription upregulation via force-induced direct stretching of chromatin. *Nat. Mater.* **15**, 1287–1296 (2016).
- First report of mechanically induced chromosome stretching and increased gene expression.**
205. Sun, J., Chen, J., Mohagheghian, E. & Wang, N. Force-induced gene up-regulation does not follow the weak power law but depends on H3K9 demethylation. *Sci. Adv.* **6**, eaay9095 (2020).
206. Almonacid, M. et al. Active fluctuations of the nuclear envelope shape the transcriptional dynamics in oocytes. *Dev. Cell* **51**, 145–157.e10 (2019).
207. Hsia, C.-R. et al. Confined migration induces heterochromatin formation and alters chromatin accessibility. *bioRxiv* <https://doi.org/10.1101/2021.09.22.461293> (2021).
208. Jacobson, E. C. et al. Migration through a small pore disrupts inactive chromatin organization in neurophil-like cells. *BMC Biol.* **16**, 142 (2018).
209. Gollosi, R. et al. Constricted migration is associated with stable 3D genome structure differences in melanoma cells. *bioRxiv* <https://doi.org/10.1101/856583> (2020).

210. Damodaran, K. et al. Compressive force induces reversible chromatin condensation and cell geometry-dependent transcriptional response. *Mol. Biol. Cell* **29**, 3039–3051 (2018).
211. Ho, C. Y., Jaalouk, D. E., Vartiainen, M. K. & Lammerding, J. Lamin A/C and emerin regulate MKL1-SRF activity by modulating actin dynamics. *Nature* **497**, 507–511 (2013).
212. Killaars, A. R., Walker, C. J. & Anseth, K. S. Nuclear mechanosensing controls MSC osteogenic potential through HDAC epigenetic remodeling. *Proc. Natl Acad. Sci. USA* **117**, 21258–21266 (2020).
213. Walker, C. J. et al. Nuclear mechanosensing drives chromatin remodelling in persistently activated fibroblasts. *Nat. Biomed. Eng.* **5**, 1485–1499 (2021).
214. Seelbinder, B. et al. Nuclear deformation guides chromatin reorganization in cardiac development and disease. *Nat. Biomed. Eng.* **5**, 1500–1516 (2021).
215. Heo, S.-J. et al. Mechanically induced chromatin condensation requires cellular contractility in mesenchymal stem cells. *Biophys. J.* **111**, 864–874 (2016).
216. Heo, S.-J. et al. Biophysical regulation of chromatin architecture instills a mechanical memory in mesenchymal stem cells. *Sci. Rep.* **5**, 16895 (2015).
217. Hannezo, E. & Heisenberg, C.-P. Mechanochemical feedback loops in development and disease. *Cell* **178**, 12–25 (2019).
218. Kirby, T. J. & Lammerding, J. Emerging views of the nucleus as a cellular mechanosensor. *Nat. Cell Biol.* **20**, 373–381 (2018).
219. Miroshnikova, Y. A. & Wickström, S. A. Mechanical forces in nuclear organization. *Cold Spring Harb. Perspect. Biol.* **14**, a039685 (2022).
220. Swift, J. & Discher, D. E. The nuclear lamina is mechano-responsive to ECM elasticity in mature tissue. *J. Cell Sci.* **127**, 3005–3015 (2014).
221. Ihalaïnen, T. O. et al. Differential basal-to-apical accessibility of lamin A/C epitopes in the nuclear lamina regulated by changes in cytoskeletal tension. *Nat. Mater.* **14**, 1252–1261 (2015).
222. Sapra, K. T. et al. Nonlinear mechanics of lamin filaments and the meshwork topology build an emergent nuclear lamina. *Nat. Commun.* **11**, 6205 (2020).
223. Guilluy, C. et al. Isolated nuclei adapt to force and reveal a mechanotransduction pathway in the nucleus. *Nat. Cell Biol.* **16**, 376–381 (2014).
- First report of mechano-adaptive changes in isolated nuclei, indicating a nucleus-intrinsic ability to respond to mechanical forces.**
224. Zwerger, M., Ho, C. Y. & Lammerding, J. Nuclear mechanics in disease. *Annu. Rev. Biomed. Eng.* **13**, 397–428 (2011).
225. Nyirenda, N., Farkas, D. L. & Ramanujan, V. K. Preclinical evaluation of nuclear morphometry and tissue topology for breast carcinoma detection and margin assessment. *Breast Cancer Res. Treat.* **126**, 345–354 (2011).
226. Mueller, J. L. et al. Rapid staining and imaging of subnuclear features to differentiate between malignant and benign breast tissues at a point-of-care setting. *J. Cancer Res. Clin. Oncol.* **142**, 1475–1486 (2016).
227. Somech, R., Shaklai, S., Amariglio, N., Rechavi, G. & Simon, A. J. Nuclear envelopathies — raising the nuclear veil. *Pediatr. Res.* **57**, 8–15 (2005).
228. Hershsberger, R. E., Hedges, D. J. & Morales, A. Dilated cardiomyopathy: the complexity of a diverse genetic architecture. *Nat. Rev. Cardiol.* **10**, 531–547 (2013).
229. Wong, X. & Stewart, C. L. The laminopathies and the insights they provide into the structural and functional organization of the nucleus. *Annu. Rev. Genomics Hum. Genet.* **21**, 263–288 (2020).
230. Bonne, G. et al. Mutations in the gene encoding lamin A/C cause autosomal dominant Emery–Dreifuss muscular dystrophy. *Nat. Genet.* **21**, 285–288 (1999).
231. Sandre-Giovannoli, A. D. et al. Lamin A truncation in Hutchinson-Gilford progeria. *Science* **300**, 2055 (2003).
232. Folker, E. S., Ostlund, C., Luxton, G. W. G., Worman, H. J. & Gundersen, G. G. Lamin A variants that cause striated muscle disease are defective in anchoring transmembrane actin-associated nuclear lines for nuclear movement. *Proc. Natl Acad. Sci. USA* **108**, 131–136 (2011).
233. Méjat, A. & Misteli, T. LINC complexes in health and disease. *Nucleus* **1**, 40–52 (2010).
234. Fischer, M., Rikeit, P., Knaus, P. & Coirault, C. YAP-mediated mechanotransduction in skeletal muscle. *Front. Physiol.* **7**, 41 (2016).
235. Owens, D. J. et al. Lamin mutations cause increased YAP nuclear entry in muscle stem cells. *Cells* **9**, 816 (2020).
236. Eriksson, M. et al. Recurrent de novo point mutations in lamin A cause Hutchinson–Gilford progeria syndrome. *Nature* **425**, 293–298 (2003).
237. Verstraeten, V. L. R. M., Ji, J. Y., Cummings, K. S., Lee, R. T. & Lammerding, J. Increased mechanosensitivity and nuclear stiffness in Hutchinson-Gilford progeria cells: effects of farnesyltransferase inhibitors. *Aging Cell* **7**, 383–393 (2008).
238. Booth, E. A., Spagnol, S. T., Alcoser, T. A. & Dahl, K. N. Nuclear stiffening and chromatin softening with progerin expression leads to an attenuated nuclear response to force. *Soft Matter* **11**, 6412–6418 (2015).
239. Kim, P. H. et al. Disrupting the LINC complex in smooth muscle cells reduces aortic disease in a mouse model of Hutchinson-Gilford progeria syndrome. *Sci. Transl. Med.* **10**, eaat7163 (2018).
240. Dahl, K. N. et al. Distinct structural and mechanical properties of the nuclear lamina in Hutchinson-Gilford progeria syndrome. *Proc. Natl Acad. Sci. USA* **103**, 10271–10276 (2006).
241. Columbaro, M. et al. Rescue of heterochromatin organization in Hutchinson-Gilford progeria by drug treatment. *Cell. Mol. Life Sci.* **62**, 2669–2678 (2005).
242. Coffinier, C. et al. Deficiencies in lamin B1 and lamin B2 cause neurodevelopmental defects and distinct nuclear shape abnormalities in neurons. *Mol. Biol. Cell* **22**, 4683–4693 (2011).
243. Young, S. G., Jung, H.-J., Coffinier, C. & Fong, L. G. Understanding the roles of nuclear A- and B-type lamins in brain development. *J. Biol. Chem.* **287**, 16103–16110 (2012).
244. Coffinier, C., Fong, L. G. & Young, S. G. LINCing lamin B2 to neuronal migration: growing evidence for cell-specific roles of B-type lamins. *Nucleus* **1**, 407–411 (2010).
245. Vortmeyer-Krause, M. et al. Lamin B2 follows lamin A/C-mediated nuclear mechanics and cancer cell invasion efficacy. *bioRxiv* <https://doi.org/10.1101/2020.04.07.028969> (2020).
246. Padiath, Q. S. et al. Lamin B1 duplications cause autosomal dominant leukodystrophy. *Nat. Genet.* **38**, 1114–1123 (2006).
247. Ballatore, C., Lee, V. M.-Y. & Trojanowski, J. Q. Tau-mediated neurodegeneration in Alzheimer’s disease and related disorders. *Nat. Rev. Neurosci.* **8**, 663–672 (2007).
248. Sergeant, C., Baillet, S. & Dehaene, S. Timing of the brain events underlying access to consciousness during the attentional blink. *Nat. Neurosci.* **8**, 1391–1400 (2005).
249. Fernández-Nogales, M. et al. Huntington’s disease is a four-repeat tauopathy with tau nuclear rods. *Nat. Med.* **20**, 881–885 (2014).
250. Crisp, M. et al. Coupling of the nucleus and cytoplasm: role of the LINC complex. *J. Cell Biol.* **172**, 41–53 (2006).
- First description of the LINC complex and its role in connecting the cytoskeleton and nuclear interior.**
251. Paonessa, F. et al. Microtubules deform the nuclear membrane and disrupt nucleocytoplasmic transport in tau-mediated frontotemporal dementia. *Cell Rep.* **26**, 582–593.e5 (2019).
252. Fernández-Nogales, M. & Lucas, J. J. Altered levels and isoforms of tau and nuclear membrane invaginations in Huntington’s disease. *Front. Cell. Neurosci.* **13**, 574 (2020).
253. von Appen, A. et al. LEM2 phase separation promotes ESCRT-mediated nuclear envelope reformation. *Nature* **582**, 115–118 (2020).
254. Stephens, A. D., Banigan, E. J. & Marko, J. F. Separate roles for chromatin and lamins in nuclear mechanics. *Nucleus* **9**, 119–124 (2018).
255. Bronshtein, I. et al. Loss of lamin A function increases chromatin dynamics in the nuclear interior. *Nat. Commun.* **6**, 8044 (2015).
256. Corne, T. D. J. et al. Deregulation of focal adhesion formation and cytoskeletal tension due to loss of A-type lamins. *Cell Adhes. Migr.* **11**, 447–463 (2017).
257. Nikolova, V. et al. Defects in nuclear structure and function promote dilated cardiomyopathy in lamin A/C-deficient mice. *J. Clin. Invest.* **113**, 357–369 (2004).
258. Puckelwartz, M. J. et al. Nesprin-1 mutations in human and murine cardiomyopathy. *J. Mol. Cell. Cardiol.* **48**, 600–608 (2010).

**Acknowledgements**

The authors apologize to all authors whose work could not be included owing to space constraints. A.D.S. is supported by the Pathway to Independence Award (R00GM123195) and 4D Nucleome 2 centre grant (1UM1HG011536). S.G. acknowledges funding from FEDER Prostem Research Project no. 1510614 (Wallonia DG06), the F.R.S.-FNRS Epifore Project no. T.0092.21 and the Interreg MAT(TJ)SSE project, which is financially supported by Interreg France-Wallonie-Vlaanderen (Fonds Européen de Développement Régional, FEDER-ERDF). Y.K. is financially supported by FRIA (F.R.S.-FNRS) and FRMH (Fonds pour la Recherche Médicale dans le Hainaut). J.L. is supported by awards from the National Institutes of Health (R01HL082792, R01GM137605, U54CA210184), the National Science Foundation (URoL-2022048) and the VolkswagenStiftung (Az. 96733).

**Author contributions**

Y.K., J.L. and S.G. conceptualized the article. J.L. and S.G. contributed equally to the editing of the text. Figure designs were generated by Y.K. and S.G. and further edited by J.L., Y.K. and S.G. All authors contributed substantially to the discussion of the content and approved the final content.

**Competing interests**

The authors declare no competing interests.

**Peer review information**

*Nature Reviews Molecular Cell Biology* thanks Matthieu Piel, Pere Roca-Cusachs who co-reviewed with Zanetta Kechagia, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Publisher’s note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© Springer Nature Limited 2022

## Review

## Tumor Innervation: Cancer Has Some Nerve

Hunter D. Reavis,<sup>1,2,3</sup> H. Isaac Chen,<sup>4,5,6</sup> and Ronny Drapkin <sup>1,2,7,\*</sup>

Over the past decade, several landmark reports have demonstrated that the nervous system plays an active role in cancer initiation and progression. These studies demonstrate that ablation of specific nerve types (parasympathetic, sympathetic, or sensory) abrogates tumor growth in a tissue-specific manner. Further, many tumor types are more densely innervated than their normal tissues of origin. These striking results raise fundamental questions regarding tumor innervation, how it is initiated, and how it molecularly contributes to disease. In this review, we aim to address what is currently known about the origin of tumor-infiltrating nerves, how they may be recruited to tumors, and how their presence may give rise to aggressive disease.

## Nerve Dependence across Cancer Types

Neural circuitry provides a means of communication between the brain and the rest of the human body. Canonical somatosensory nerves relay signals to the central nervous system for processing stimuli, while efferent motor and autonomic nerves relay signals from the central nervous system to target tissues throughout the body in order to respond to stimuli (Figure 1). However, it has recently become appreciated that normal neuron development and communication cues may be altered in the context of disease, including, but not limited to tumorigenesis.

While preliminary reports of nerves infiltrating tumors surfaced in the early 20th century, it was long thought that these components of the tumor microenvironment (TME) were merely passive bystanders [1,2]. However, the tumor innervation field has recently gained attention following the observation that cancer cells intimately interact with nerves infiltrating the tumor and that they do so with a purpose [3,4]. In 2001, cocultures of mouse neurons with prostate cancer cells revealed that neural processes actively extend towards cancer cells and ultimately stimulate their growth [5]. This study inspired the idea that nerves may not simply be bystanders within the TME, but instead may be functionally promoting tumorigenesis; cancer cells preferentially recruit these nerves to exploit nerve-mediated effects that fuel their expansion. What began as an observation has since transformed into an entire research focus dedicated to understanding when, how, and why nerve signals influence cancer progression.

Initial *in vitro* coculture experiments that rationalized studying tumor innervation led to the next logical question: What happens to tumors in the absence of nerves? A 2013 study set out to answer this, finding that prostate tumors are innervated by both adrenergic sympathetic and cholinergic parasympathetic nerves and that ablation of these nerves impairs *in vivo* tumorigenesis [6]. This paper was succeeded by other important nerve ablation studies in breast and gastric cancers that demonstrated a similar dependence on autonomic innervation [7,8]. However, it is important to note that there appears to be a tropic relationship between the tumor type and the innervation source. For example, stimulation of sympathetic nerves is associated with breast and pancreatic cancer progression, while stimulation of parasympathetic nerves is associated with a reduction in tumor growth [7,9,10] (Table 1). Further, models of head and neck squamous cell carcinoma, pancreatic ductal adenocarcinoma, cervical carcinoma, basal cell carcinoma, melanoma, and high-grade serous ovarian carcinoma demonstrate dependence on sensory

## Highlights

Tumors may promote their innervation by recruiting locoregional nerves, neural progenitor cells (NPCs), and/or reprogramming other cell types.

Nerves may be attracted to the tumor microenvironment (TME) by both soluble secreted factors as well as functional cargo transported within tumor-derived small extracellular vesicles (sEVs).

Nerves contribute to disease progression, in both tumor cell-dependent and tumor cell-independent manners.

<sup>1</sup>Penn Ovarian Cancer Research Center, Department of Obstetrics and Gynecology, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA

<sup>2</sup>Graduate Program in Cell and Molecular Biology, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA

<sup>3</sup>Department of Cancer Biology, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA

<sup>4</sup>Department of Neurosurgery, University of Pennsylvania, Philadelphia, PA, USA

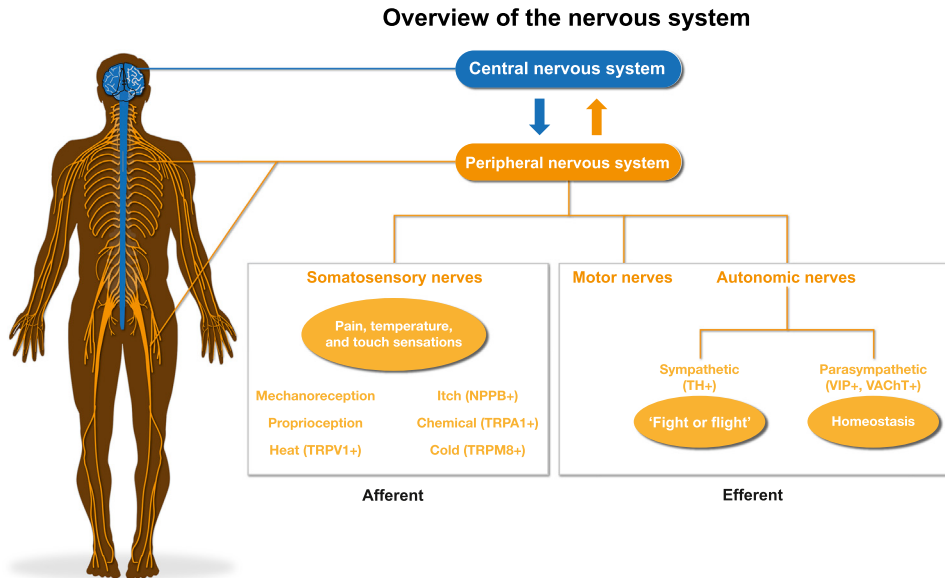
<sup>5</sup>Institute for Regenerative Medicine, University of Pennsylvania, Philadelphia, PA, USA

<sup>6</sup>Corporal Michael J. Crescenz Veterans Affairs Medical Center, Philadelphia, PA, USA

<sup>7</sup>Basser Center for BRCA, Abramson Cancer Center, University of Pennsylvania School of Medicine, Philadelphia, PA, USA

\*Correspondence: [rdrapkin@pennmedicine.upenn.edu](mailto:rdrapkin@pennmedicine.upenn.edu) (R. Drapkin).





**Trends In Cancer**

**Figure 1. Overview of the Nervous System.** The central nervous system is comprised of the brain and the spinal cord (blue), that integrates afferent signals from the peripheral nervous system (orange) and consequently triggers a response by means of efferent signals to the body. Somatosensory nerves in the peripheral nervous system may be stimulated by activation of mechanoreceptors, proprioceptors, natriuretic peptide B (NPPB+) itch, transient receptor potential cation channel, subfamily A member 1 (TRPA1+) chemical, transient receptor potential vanilloid (TRPV1+) heat, or transient receptor potential cation channel subfamily M member 8 (TRPM8+) cold receptors. Efferent signals consist of both motor and autonomic nerves. Autonomic nerves may either be tyrosine hydroxylase (TH+) sympathetic nerves that control 'fight or flight' responses, or vasoactive intestinal peptide/vesicular acetylcholine transporter (VAcHT+/VIP+) parasympathetic nerves that maintain homeostasis of the body.

**Table 1. Summary of Innervation Studies across Different Cancer Types<sup>a</sup>**

	General innervation	Sympathetic	Parasympathetic	Sensory	Relevant guidance molecules	Refs
Basal cell	Undefined	Undefined	Undefined	+	Undefined	[15]
Breast	+	+	-	+	NGF, VEGFA <sup>b</sup>	[7,33,48,49]
Cervical	Undefined	Undefined	Undefined	+	sEVs	[13]
Colon	+	Undefined	-	Undefined	Undefined	[50-52]
Gastric	Undefined	-	+	Undefined	NGF	[8,53]
Glioma	+	Undefined	Undefined	Undefined	Undefined	[41,42,54,55]
Head and neck	+	Undefined	Undefined	+	EphrinB1 (sEV), miR-34a (sEV)	[12,14]
Melanoma	Undefined	Undefined	Undefined	+	Undefined	(S. Talbot, unpublished)
Ovary	Undefined	+	Undefined	+	BDNF	[16,56]
Pancreatic	+	+	-	+	SEMA3D	[9-11,35,37]
Prostate	+	+	+	Undefined	SEMA4F, proNGF, GCSF	[6,31,32,34,47]
Thyroid	+	Undefined	Undefined	Undefined	proNGF	[57]

<sup>a</sup>Innervation associated with protumorigenic processes are annotated '+', and antitumorigenic processes are annotated '-'. Relevant neurotrophic factors regulating tumor innervation are listed when applicable. sEV indicates that small extracellular vesicle cargo is associated with the observed phenotype.

<sup>b</sup>Abbreviations: BDNF, brain-derived neurotrophic factor; SEMA4F, semaphorin 4F; VEGFA, vascular endothelial growth factor A.



nerves [11–16] (S. Talbot, unpublished). Altogether, these findings highlight the importance of studying tumor innervation in a tissue-specific manner; it is not simply true that all innervation types promote disease progression in all cancer types.

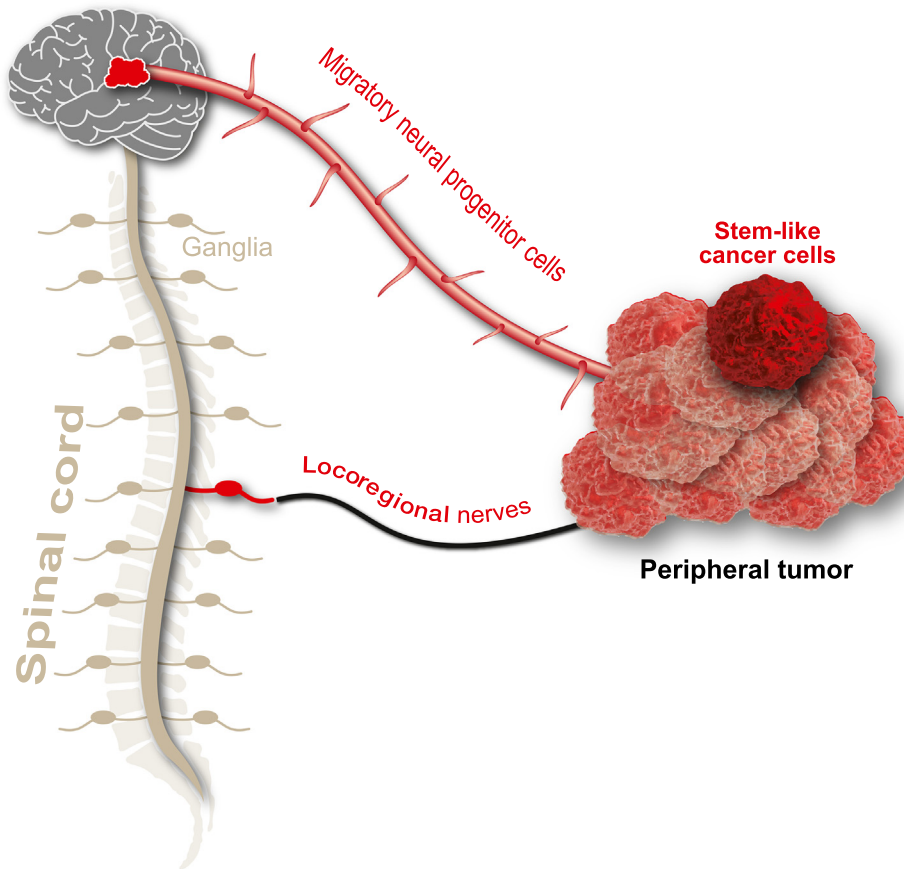
## Tumor Nerve Origins

### Locoregional Nerves

It seems logical that cancers recruit local peripheral nerves that normally innervate their tissue of origin. This hypothesis has recently been tested using uniquely engineered adeno-associated viral (AAV) vectors. Specifically, this study took advantage of retroAAVs that are retrogradely transported from nerve terminals to the nerve soma. By incorporating a neuron-specific promoter driving expression of diphtheria toxin (DTA) into the retroAAV, the authors were able to ablate specific subpopulations of tumor-infiltrating nerves. For instance, intratumoral injection of a retroAAV-TH-DTA results in retrograde transport of DTA, where expression is driven by the tyrosine hydroxylase promoter (TH, sympathetic); such expression specifically ablates sympathetic tumor-infiltrating nerves. Selective ablation of these TH+ nerves led to drastic reductions in metastasis and primary tumor volumes, highlighting the fact that tumor growth is regulated by nerves that directly interact with the tumor, as opposed to an artifact of systemic genetic or pharmacological nerve ablation [7]. Alternatively, other groups have employed mechanical severing of locoregional nerves to study breast cancer progression [8]. However, whether or not these nerve terminals extend from the ganglia in the spinal cord that typically provide innervation to normal tissue remains unknown (Figure 2). Another recent study suggests that locoregional nerves may be reprogrammed in the TME upon cancer cell transformation. Evidence suggests that *TP53* mutant oral cavity squamous cell carcinoma cells reprogram trigeminal sensory nerves to restore TH+ adrenergic nerves that are ablated by pharmacological sympathectomy with 6-OHDA [14]. Furthermore, severing of the trigeminal nerve decreased tumor growth in *TP53* mutant xenograft models. Therefore, it is indeed plausible that tumors may be exploiting locoregional neural plasticity to fuel their own growth.

### Remote Neural Progenitor Cells

Curiously, the subtypes of nerves that are found in tumors are not always identical to the subtypes of resident nerves that infiltrate the tissue of origin. For example, transient receptor potential vanilloid (TRPV)1+ nociceptive pain sensory nerve ‘twigs’ are abundant in high grade serous ovarian carcinomas, yet absent from the normal fallopian tube and ovaries [16]. While it is possible that these *de novo* ‘twigs’ could be derived from existing sensory nerves in neighboring tissues, their origins may also be explained by an alternative, more provocative origin of tumor innervation. A recent prostate cancer study utilized lineage-tracing of doublecortin-positive (DCX+) neural progenitor cells (NPCs) that originate in the subventricular zone (SVZ) of the brain to examine whether tumor innervation can originate from the central nervous system. Importantly, the authors showed that SVZ-derived NPCs can escape the blood–brain barrier and travel through the circulation to colonize the prostate tumor. Upon arrival in the TME, these NPCs mature to form TH+ sympathetic nerves, which were previously shown to promote prostate tumorigenesis [6]. In addition, they show that depletion of these DCX+ NPCs abrogates prostate tumor initiation and progression and that transplantation of DCX+ cells from the SVZ enhances tumor xenograft growth and metastasis *in vivo* [17]. This model provides important insight into how cancer may be taking advantage of normal systemic developmental processes like neurogenesis and further indicates that complex signals exist between peripheral tumors and the central nervous system. NPCs maintain their quiescent, undifferentiated state at least in part by cell–cell interactions with vascular endothelial cells [18]. Therefore, cancer-derived factors in the circulation could be recruiting these cells to the TME. In support of this idea, a 1998 study showed that NPCs in the SVZ of adult human cancer patients actively incorporate the thymidine-analog BrDU into their DNA, indicating that

**Trends in Cancer**

**Figure 2. Schematic of Potential Sources of Tumor Innervation.** Densely innervated tumors may be supported by neurons derived from many different origins. In addition to canonical locoregional innervation extending from spinal cord ganglia, neural progenitor cells from the subventricular zone may be able to migrate through the vasculature to peripheral tumor sites. Alternatively, plastic or stem-like cancer cells may transdifferentiate within the tumor microenvironment and support disease progression.

under pathological conditions, these cells are indeed capable of escaping their quiescent state and re-entering the cell cycle [19]. Given that NPCs from the SVZ are thought to be tumor progenitor cells in glioblastoma, this concept of atypical NPC activation is further strengthened [20]. However, the scale of SVZ neurogenesis in the adult human is highly controversial. While several groups have established the proliferative potential of NPCs in the SVZ, it is still heavily debated whether or not these NPCs are capable of noncanonical migration and/or persist into adulthood [21–23]. Further studies into the migratory capacity of adult NPCs will be necessary to determine the potential of these cell populations to contribute to tumor infiltrating nerves.

#### Cancer Stem Cells and Plastic Cancer Cells

Stem cells are classically considered to be cells that can both self-renew and differentiate to generate progeny of different lineages [24]. In the context of cancer stem cells, this means that tumor heterogeneity may be informed by a single cell with the capacity for ‘stemness.’ By this notion, it is not unreasonable to assume that fate determination during the expansion of cancer stem cell progeny could, in fact, give rise to neural populations within the tumor that are capable of differentiating into fully functional nerves in the TME. Lineage tracing of *ex vivo*-tagged human gastric and colorectal carcinoma stem cells revealed that these cells

not only produce tumors in xenograft mouse models, but also that the tagged human cancer stem cells differentiate and express neural proteins upon engraftment. Moreover, inhibition of this differentiation capacity abrogated xenograft tumor growth [25]. Taken together, these data strongly support the idea that cancer stem cells can expand and mature into heterogeneous cell populations, including neuron-like cells, that are found in the TME. Identification of stem-like progenitors in prostate cancer revealed that a subset of these cells exhibit neurogenic gene expression profiles, consistent with the cancer stem cell nerve origin hypothesis [26]. While it is appreciated that both cancer stemness and nerve density are independently associated with aggressive disease, further investigation is required to determine the relative contribution of cancer stem cells to tumor innervation.

Cancer cell plasticity has also been implicated in the generation of neuron-like cells in the TME. By contrast to cancer stem cells, 'plastic' cancer cells are not necessarily multipotent, self-renewing cells, but rather cancer cells whose identity is capable of being reversibly reprogrammed [27]. For example, *in vitro* studies demonstrate that a cocktail of epigenetic inhibitors (HDACi + EZH2i + LSD1i + DNMTi) can induce expression of neural gene programs in a variety of cancer cell lines [28]. With the notion that cancer cells are plastic and capable of gaining neural identity, these data raise the question as to whether a neural reprogramming phenomenon may be occurring naturally *in vivo* during tumorigenesis. 3D cocultures of rat SVZ neural stem cells with human pancreatic cancer cells have revealed that cancer cells aggregate on the surface of neurospheres and begin to mimic the morphology of differentiated neurons. Further, implantation of these hybrid cancer cells into nude mice accelerates tumor progression, indicated by decreased survival of mice harboring hybrid xenografts when compared with the parental pancreatic cancer cell line [29]. These data not only underscore the correlation between neural cells and aggressive disease, but also suggest that these populations may arise from reprogrammed cancer cells themselves.

## Mechanisms of Nerve Recruitment

### Growth Factors and Axon Guidance Molecules

Under normal physiological conditions, nerve growth is dictated by changes in intracellular neuronal signaling in response to local biochemical signals, including neuropeptides, neurotransmitters, growth factors, and morphogens, that recruit axons to target tissues. As cancer is known to exploit normal developmental processes to fuel tumorigenesis, evidence suggests that secretion of these proneural signals, particularly growth factors, in the TME may facilitate nerve recruitment [30] (Table 1). For example, prostate cancer cells secrete a precursor to neural growth factor (proNGF) that induces axonogenesis *in vitro* [31]. Granulocyte-colony stimulating factor (GCSF) has also been implicated in prostate tumor innervation, indicating that this process is multifaceted and may be mediated by multiple secreted factors [32]. Further, inhibition of NGF in breast cancer cells partially abrogates neurite outgrowth, bolstering the idea that these factors secreted by the tumor can alter the behavior of nerves and highlighting the tissue-dependent nature of tumor innervation [33].

In addition to circulating growth factors, axon guidance molecules may either be secreted or present on the surface of tumor cells in the TME to promote directional cues that ensure tumor infiltration. Functional testing reveals that overexpression of semaphorin 4F (S4F) increases both the frequency of neurite outgrowth as well as neurite length in *in vitro* cocultures with mouse neuroblastoma cells [34]. Another recent study has further implicated semaphorin overexpression in the enhancement of nerve-dependent tumorigenesis, reporting on another member of the semaphorin family, SEMA3D. In this study, the authors found that SEMA3D secreted by pancreatic cancer cells can engage and activate plexin D1 (PLXND1) on the neural

surface. The resulting innervation observed in PLXND1 active pancreatic cancer models was abrogated upon SEMA3D knockdown and/or inhibition of this paracrine signal with a PLXND1 neutralizing antibody [35]. Taken together, these results suggest a functional role of semaphorin-mediated axon guidance in recruitment of nerves to the TME. However, it is important to note that while these neurotrophic factors have been implicated in axonogenesis, whether they are directly secreted into the extracellular space or released as extracellular vesicle cargo remains under investigation.

### Small Extracellular Vesicles

While it has been long appreciated that small extracellular vesicles (sEVs) released from cancer cells increase tumor growth [36], recent evidence suggests that these sEVs also have axonogenic properties. In *in vitro* coculture experiments with rat pheochromocytoma (PC12) cells and sEVs isolated from cervical carcinoma cell cultures, robust neurite outgrowth is observed [13]. In accordance with these data, sEVs isolated from high-grade serous carcinomas (HGSCs) and oncogene-transformed fallopian tube cell lines similarly induce PC12 neurite outgrowth *in vitro*, whereas sEVs isolated from normal fallopian tube cell lines do not [16]. Not only does this imply that sEVs play an important role in signaling with neural cells in the TME, but also that newly transformed cancer cells may be initiating this conversation early on in tumorigenesis. These data are further supported by similar reports in head and neck squamous cell carcinoma, where genetic and pharmacological inhibition of sEV release not only abrogate tumor growth, but also decrease tumor innervation [12]. Further analysis of sEV cargo revealed that this innervation phenotype is enhanced by the membrane protein EphrinB1. Another group recently employed miRNA profiling of sEV cargo to find that miR-34a is transported within sEVs from TP53-WT (wild type) head and neck cancer cells to suppress neuritogenesis of local sensory nerves by disrupting the expression of genes involved in sensory nerve differentiation [14]. However, aside from these two studies, the functional components of the sEV cargo that mediate tumor innervation remain largely unknown.

## Effects of Tumor Innervation on Disease Progression

### Direct Effects on Cancer Cell Proliferation and Migration

Cancer-mediated recruitment of nerves to the TME triggers reciprocal signaling whereby interactions with nerves support the progression of disease. For example, cocultures of pancreatic cancer cells with dorsal root ganglia (DRG) exhibit increased proliferation, as indicated by strong Ki-67 staining, as well as decreased apoptosis in the presence of these sensory neurons [37]. Similarly, in prostate cancer, cells that are closer to the nerve exhibit increased Ki-67 staining and decreased apoptosis when compared with cancer cells that are further away from the site of the nerve, which is believed to be mediated by increasing NF- $\kappa$ B signaling in the cancer cells [38]. From a paracrine perspective, several cancer types exhibit neuropeptide receptors on their cell surface, providing a plausible mechanism for nerve-to-tumor communication. In support of this idea, inhibition of neuropeptide receptors on esophageal squamous cell carcinoma cells decreases cell proliferation and viability in a PI3K/AKT/NF- $\kappa$ B-dependent mechanism [39]. When evaluating cancer cell surface receptors that may be responsive to nerve-derived factors, it is also important to mention that several tumor types express neurotrophic growth factor receptors, indicating that signaling molecules may also act on the cancer cells in an autocrine manner [40]. Alternatively, two pivotal studies in glioma have recently revealed that cancer cells may be capable of forming direct, functional synapses with neurons [41,42]. Not only are postsynaptic signals detectable in glioma cells, but direct activation of the presynaptic neurons stimulates glioma proliferation. These studies are the first of their kind, showing that electrical stimulation of neurons mediates direct communication with cancer cells in a synaptic manner. Altogether, these data raise questions as to whether direct interactions of neurons and cancer

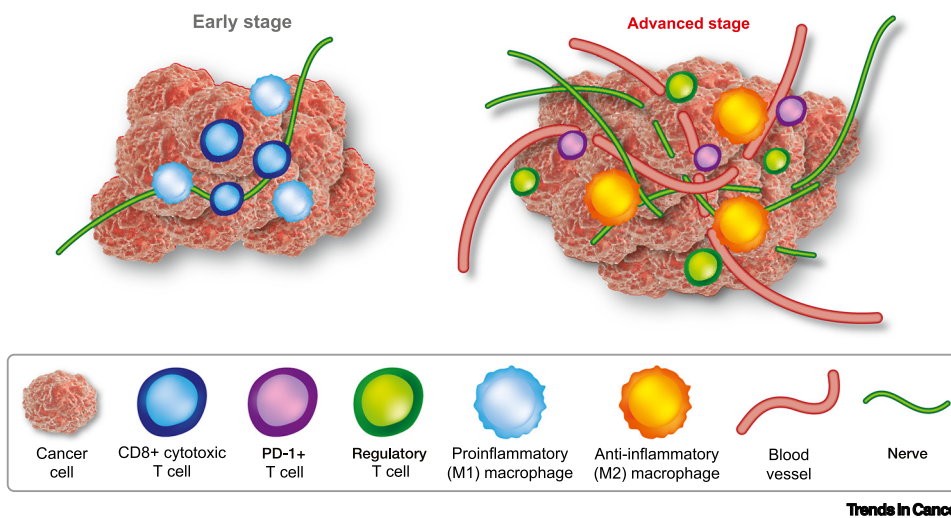


cells are primarily mediated by electrical and/or biochemical signaling processes and will need to be further investigated in peripheral tumor types.

In addition to the nerve-mediated pro-growth signaling, *de novo* innervation may also stimulate intracellular migratory programs. *In vitro* pancreatic cancer/DRG cocultures [37] support this, where increased proliferation and decreased cell death were accompanied by retrograde extension and migration of cancer cells towards the sensory nerves. It is important to distinguish that this process is distinct from the clinical process of perineural invasion (PNI); PNI is a process whereby tumor cells migrate along existing nerve bundles that serve as a 'highway' for tumor dissemination. Conversely, *de novo* tumor innervation refers to the infiltration of nerve 'twigs' that lack the same structure and organization seen in resident tissue nerves. Additional investigation will be required to determine the mechanisms of *de novo* tumor innervation on cancer cell migration.

#### Indirect Effects on Non-cancer Cells within the TME

Several recent reports suggest a role for activated nerves in modulating other non-tumor cell populations within the TME. Importantly, immune responses are susceptible to modulation by the nervous system [43,44]. For example, stimulation of the vagus nerve induces the release of acetylcholine, which inhibits tumor necrosis factor (TNF)-mediated macrophage activation in the presence of endotoxins [45]. With this in mind, it is reasonable to hypothesize that prolonged exposure to secreted neural factors may sculpt the immune landscape to allow a tumor to evade immunosurveillance (Figure 3). In the context of the innervated TME, recent reports suggest that the adaptive immune landscape changes quite significantly upon sympathectomy. In analyzing the T cell landscape, ablation of sympathetic nerves decreases expression of the immune checkpoint molecule programmed cell death-1 (PD-1) on T cells, indicating that these lymphocyte populations should be more effective at mounting an antitumor immune response. Moreover, they report that sympathetic innervation increases immunosuppressive FOXP3+ regulatory T cell populations [7]. Gene expression analysis within



**Figure 3. Dense Innervation Influences Immune Cell Populations and Vasculature within the Tumor Microenvironment.** During early tumorigenesis, tumors are likely sparsely innervated and vascularized, with active immunosurveillance managed by proinflammatory macrophages and active cytotoxic T cells. As the cancer progresses, dense innervation yields increased angiogenesis and preferentially supports the survival of immunosuppressive cell populations [anti-inflammatory M2 macrophages, exhausted/PD-1 (programmed cell death-1) expressing T cells, regulatory T cells].

triple-negative breast cancer patients revealed that patients with a high neurogenic signature also exhibit a more immunosuppressive profile when compared with other subtypes, marked particularly by M2-macrophage polarization [46]. In a more recent melanoma study, TRPV1+ sensory nerve ablation restores cytotoxic CD8+ T cell activity, as indicated by decreases in the number of exhausted cytotoxic T cells present in the TME. This phenotype is mediated by neuropeptides released from TRPV1+ sensory nerves that are enriched in the melanoma TME, suggesting a direct role for innervation in tumorigenesis by suppressing the immune response (S. Talbot, unpublished).

During development, nerves extend alongside endothelial cells that construct the vasculature. Importantly, recent reports show that early angiogenesis, another hallmark of cancer, is regulated by sympathetic adrenergic nerves during tumor development [47]. While this study has broadened the implications of tumor innervation on remodeling the vasculature of the TME (Figure 3), it is one of the first of its kind and will need to be investigated in other contexts.

### Concluding Remarks

While it has become increasingly clear that nerves play a functional role in cancer, there is still much to be uncovered. With several investigations into tumor innervation in different tumor types, we can appreciate that ‘one size’ does not fit all. Further studies into the role of each nerve type in different cancer contexts will help to elucidate potential therapeutic approaches. Understanding the balance between stemness and natural innervation will provide further biological insight into how cancer hijacks normal developmental processes (see [Outstanding Questions](#)). Identifying the relevant players that participate in attracting these nerves could identify targetable growth factors and/or sEV-enclosed neurotrophic factors that could be successful in staving off nerve-mediated tumor onset. Lastly, understanding the interplay between these tumor infiltrating nerves, tumor cells, and non-tumor cells may inform potential combination therapeutics to further weaken the TME barrier that protects these malignancies. Altogether, the identification of nerve dependence in cancer has illuminated the field and should prove to be a very exciting new therapeutic opportunity.

### Acknowledgments

We thank Dr Paola Vermeer for her suggestions and critical reading of the manuscript and Cooper Graphics (<http://cooper247.com>) for the illustrations. This work was supported by the National Cancer Institute SPORE in ovarian cancer P50 CA228991, a Department of Defense Impact Award W81XWH-19-1-0232, the Dr Miriam and Sheldon G. Adelson Medical Research Foundation, The Honorable Tina Brozman ‘Tina’s Wish’ Foundation, The Claneil Foundation, The V Foundation, the Basser Center for BRCA, and the Shooting for a Cure Foundation.

### Disclaimer Statement

R.D. is an advisory board member for Repare Therapeutics. No other authors declare potential conflicts of interest.

### References

- Young, H.H. (1897) On the presence of nerves in tumors and of other structures in them as revealed by a modification of Ehrlich's method of ‘vital staining’ with methylene blue. *J. Exp. Med.* 2, 1–12
- Oertel, H. (1928) Innervation and tumour growth: a preliminary report. *Can. Med. Assoc. J.* 18, 135–139
- Servick, K. (2019) War of nerves. *Science* 365, 1071–1073
- Vermeer, P.D. (2019) Exosomal induction of tumor innervation. *Cancer Res.* 79, 3529–3535
- Ayala, G.E. *et al.* (2001) In vitro dorsal root ganglia and human prostate cell line interaction: redefining perineural invasion in prostate cancer. *Prostate* 49, 213–223
- Magnon, C. *et al.* (2013) Autonomic nerve development contributes to prostate cancer progression. *Science* 341, 1236361
- Kamiya, A. *et al.* (2019) Genetic manipulation of autonomic nerve fiber innervation and activity and its effect on breast cancer progression. *Nat. Neurosci.* 22, 1289–1305
- Zhao, C.M. *et al.* (2014) Denervation suppresses gastric tumorigenesis. *Sci. Transl. Med.* 6, 250ra115
- Renz, B.W. *et al.* (2018) Beta2 adrenergic-neurotrophin feedforward loop promotes pancreatic cancer. *Cancer Cell* 33, 75–90
- Renz, B.W. *et al.* (2018) Cholinergic signaling via muscarinic receptors directly and indirectly suppresses pancreatic tumorigenesis and cancer stemness. *Cancer Discov.* 8, 1458–1473
- Saloman, J.L. *et al.* (2016) Ablation of sensory neurons in a genetic model of pancreatic ductal adenocarcinoma slows initiation and progression of cancer. *Proc. Natl. Acad. Sci. U. S. A.* 113, 3078–3083

### Outstanding Questions

Do tumors hijack locoregional nerves?

How abundant are neural progenitor-derived tumor infiltrating nerves?

Are plastic/cancer stem cells capable of transdifferentiating and forming mature nerves?

What are the neurotrophic contents of cancer-derived sEVs?

Do nerves form functional synapses with cancer cells in non-glioma cancers?

How do nerves primarily change cancer intrinsic and extrinsic programs to support tumorigenesis?

Can nerve origins, neuroattractants, tumor-neural synapses, and effects of innervation be mitigated therapeutically?

12. Madeo, M. *et al.* (2018) Cancer exosomes induce tumor innervation. *Nat. Commun.* 9, 4284
13. Lucido, C.T. *et al.* (2019) Innervation of cervical carcinoma is mediated by cancer-derived exosomes. *Gynecol. Oncol.* 154, 228–235
14. Amit, M. *et al.* (2020) Loss of p53 drives neuron reprogramming in head and neck cancer. *Nature* 578, 449–454
15. Peterson, S.C. *et al.* (2015) Basal cell carcinoma preferentially arises from stem cells within hair follicle and mechanosensory niches. *Cell Stem Cell* 16, 400–412
16. Kovacs, A. *et al.* (2020) Tumor-infiltrating nerves create an electro-physiologically active microenvironment and contribute to treatment resistance. *bioRxiv* Published online April 25, 2020. <https://doi.org/10.1101/2020.04.24.058594>
17. Mauffrey, P. *et al.* (2019) Progenitors from the central nervous system drive neurogenesis in cancer. *Nature* 569, 672–678
18. Ottone, C. and Parrinello, S. (2015) Multifaceted control of adult SVZ neurogenesis by the vascular niche. *Cell Cycle* 14, 2222–2225
19. Eriksson, P.S. *et al.* (1998) Neurogenesis in the adult human hippocampus. *Nat. Med.* 4, 1313–1317
20. Alcantara Llaguno, S. *et al.* (2009) Malignant astrocytomas originate from neural stem/progenitor cells in a somatic tumor suppressor mouse model. *Cancer Cell* 15, 45–56
21. Curtis, M.A. *et al.* (2007) Human neuroblasts migrate to the olfactory bulb via a lateral ventricular extension. *Science* 315, 1243–1249
22. Ernst, A. *et al.* (2014) Neurogenesis in the striatum of the adult human brain. *Cell* 156, 1072–1083
23. Sanai, N. *et al.* (2011) Corridors of migrating neurons in the human brain and their decline during infancy. *Nature* 478, 382–386
24. Lytle, N.K. *et al.* (2018) Stem cell fate in cancer growth, progression and therapy resistance. *Nat. Rev. Cancer* 18, 669–680
25. Lu, R. *et al.* (2017) Neurons generated from carcinoma stem cells support cancer progression. *Signal Transduct. Target Ther.* 2, 16036
26. Zhang, D. *et al.* (2016) Stem cell and neurogenic gene-expression profiles link prostate basal cells to aggressive prostate cancer. *Nat. Commun.* 7, 10798
27. Meacham, C.E. and Morrison, S.J. (2013) Tumour heterogeneity and cancer cell plasticity. *Nature* 501, 328–337
28. Zhang, Z. *et al.* (2017) Similarity in gene-regulatory networks suggests that cancer cells share characteristics of embryonic neural cells. *J. Biol. Chem.* 292, 12842–12859
29. Yin, L. *et al.* (2020) Cancer cell's neuroendocrine feature can be acquired through cell-cell fusion during cancer-neural stem cell interaction. *Sci. Rep.* 10, 1216
30. Boilly, B. *et al.* (2017) Nerve dependence: from regeneration to cancer. *Cancer Cell* 31, 342–354
31. Pundavela, J. *et al.* (2014) ProNGF correlates with Gleason score and is a potential driver of nerve infiltration in prostate cancer. *Am. J. Pathol.* 184, 3156–3162
32. Dobrenis, K. *et al.* (2015) Granulocyte colony-stimulating factor off-target effect on nerve outgrowth promotes prostate cancer development. *Int. J. Cancer* 136, 982–988
33. Pundavela, J. *et al.* (2015) Nerve fibers infiltrate the tumor microenvironment and are associated with nerve growth factor production and lymph node invasion in breast cancer. *Mol. Oncol.* 9, 1626–1635
34. Ayala, G.E. *et al.* (2008) Cancer-related axonogenesis and neurogenesis in prostate cancer. *Clin. Cancer Res.* 14, 7593–7603
35. Jurcak, N.R. *et al.* (2019) Axon guidance molecules promote perineural invasion and metastasis of orthotopic pancreatic tumors in mice. *Gastroenterology* 157, 838–850
36. Keller, S. *et al.* (2009) Systemic presence and tumor-growth promoting effect of ovarian carcinoma released exosomes. *Cancer Lett.* 278, 73–81
37. Dai, H. *et al.* (2007) Enhanced survival in perineural invasion of pancreatic cancer: an in vitro approach. *Hum. Pathol.* 38, 299–307
38. Ayala, G.E. *et al.* (2004) Growth and survival mechanisms associated with perineural invasion in prostate cancer. *Cancer Res.* 64, 6082–6090
39. Javid, H. *et al.* (2020) The role of substance P/neurokinin 1 receptor in the pathogenesis of esophageal squamous cell carcinoma through constitutively active PI3K/Akt/NF- $\kappa$ B signal transduction pathways. *Mol. Biol. Rep.* 47, 2253–2263
40. Meldolesi, J. (2018) Neurotrophin Trk receptors: new targets for cancer therapy. *Rev. Physiol. Biochem. Pharmacol.* 174, 67–79
41. Venkatesh, H.S. *et al.* (2019) Electrical and synaptic integration of glioma into neural circuits. *Nature* 573, 539–545
42. Venkataramani, V. *et al.* (2019) Glutamatergic synaptic input to glioma cells drives brain tumour progression. *Nature* 573, 532–538
43. Ulloa, L. *et al.* (2017) Nerve stimulation: immunomodulation and control of inflammation. *Trends Mol. Med.* 23, 1103–1120
44. Steinman, L. (2004) Elaborate interactions between the immune and nervous systems. *Nat. Immunol.* 5, 575–581
45. Borovikova, L.V. *et al.* (2000) Vagus nerve stimulation attenuates the systemic inflammatory response to endotoxin. *Nature* 405, 458–462
46. Jezequel, P. *et al.* (2019) Identification of three subtypes of triple-negative breast cancer with potential therapeutic implications. *Breast Cancer Res.* 21, 65
47. Zahalka, A.H. *et al.* (2017) Adrenergic nerves activate an angiometabolic switch in prostate cancer. *Science* 358, 321–326
48. Zhao, Q. *et al.* (2014) The clinicopathological significance of neurogenesis in breast cancer. *BMC Cancer* 14, 484
49. Austin, M. *et al.* (2017) Breast cancer induced nociceptor aberrant growth and collateral sensory axonal branching. *Oncotarget* 8, 76606–76621
50. Albo, D. *et al.* (2011) Neurogenesis in colorectal cancer is a marker of aggressive tumor behavior and poor outcomes. *Cancer* 117, 4834–4845
51. Liebl, F. *et al.* (2013) The severity of neural invasion is associated with shortened survival in colon cancer. *Clin. Cancer Res.* 19, 50–61
52. Dubeykovskaya, Z. *et al.* (2016) Neural innervation stimulates splenic TFF2 to arrest myeloid cell expansion and cancer. *Nat. Commun.* 7, 10517
53. Hayakawa, Y. *et al.* (2017) Nerve growth factor promotes gastric tumorigenesis through aberrant cholinergic signaling. *Cancer Cell* 31, 21–34
54. Venkatesh, H.S. *et al.* (2015) Neuronal activity promotes glioma growth through neuropilin-3 secretion. *Cell* 161, 803–816
55. Venkatesh, H.S. *et al.* (2017) Targeting neuronal activity-regulated neuropilin-3 dependency in high-grade glioma. *Nature* 549, 533–537
56. Allen, J.K. *et al.* (2018) Sustained adrenergic signaling promotes intratumoral innervation through BDNF induction. *Cancer Res.* 78, 3233–3242
57. Rowe, C.W. *et al.* (2020) Innervation of papillary thyroid cancer and its association with extra-thyroidal invasion. *Sci. Rep.* 10, 1539

## REVIEW ARTICLE OPEN



# Research progress on morphology and mechanism of programmed cell death

Yao Chen<sup>1,6</sup>, Xiaohua Li<sup>2,6</sup>, Minfeng Yang<sup>3,4</sup> and Song-Bai Liu<sup>1,5</sup>

© The Author(s) 2024

Programmed cell death (PCD) is a basic process of life that is closely related to the growth, development, aging and disease of organisms and is one of the hotspots of life science research today. PCD is a kind of genetic control, autonomous and orderly important cell death that involves the activation, expression, and regulation of a series of genes. In recent years, with the deepening of research in this field, new mechanisms of multiple PCD pathways have been revealed. This article reviews and summarizes the multiple PCD pathways that have been discovered, analyses and compares the morphological characteristics and biomarkers of different types of PCD, and briefly discusses the role of various types of PCD in the diagnosis and treatment of different diseases, especially malignant tumors.

*Cell Death and Disease* (2024)15:327; <https://doi.org/10.1038/s41419-024-06712-8>

## FACTS

1. Different types of PCD have unique signaling pathways and specific morphological characteristics.
2. Clarifying the specific biomarkers and regulatory mechanisms of various PCD pathways and their relationships with specific diseases is vital for further targeted treatment.
3. Further studies on the molecular mechanism and mutual cross-links of various types of PCD and the identification of specific drug targets are the main directions in the future.

## OPEN QUESTIONS

1. How many forms of PCD exist, and how can they be classified more scientifically?
2. Which PCD pathways are preferentially activated, and how can different PCD pathways be coordinated for the treatment of specific diseases?
3. How can the potential off-target effects of PCD induction be effectively prevented and controlled for disease therapy?

## INTRODUCTION

Cell death is a common biological phenomenon that plays an important role in the growth and development of the body and is

related to the occurrence and development of a variety of diseases. Cell death includes accidental cell death (ACD) and programmed cell death (PCD). ACD is a kind of passive catastrophic cell death caused by extreme physical (such as high pressure, high temperature, high osmotic pressure, etc.), chemical, or mechanical damage, in which necrosis is the main type of ACD. Necrosis in ACD is an unregulated, passive form of cell death characterized by cell swelling, membrane rupture, organelle collapse, and the release of cell contents, often leading to an inflammatory response in the body. PCD is an active form of cell death that occurs under physiological conditions controlled by genes, and it is an important regulatory mechanism for the body to stabilize the internal environment and balance the number of cells in the physiological and pathological environment. It depends on a special molecular mechanism and can be regulated (i.e., delayed or accelerated) through drug or gene intervention. It is worth noting that some forms of PCD discovered in recent years, such as necroptosis and pyroptosis, also show characteristics similar to those of necrosis in ACD, including membrane rupture and inflammatory response; however, unlike necrosis in ACD, these processes are regulated and can be affected by inhibitors, activators, protein expression levels, etc.

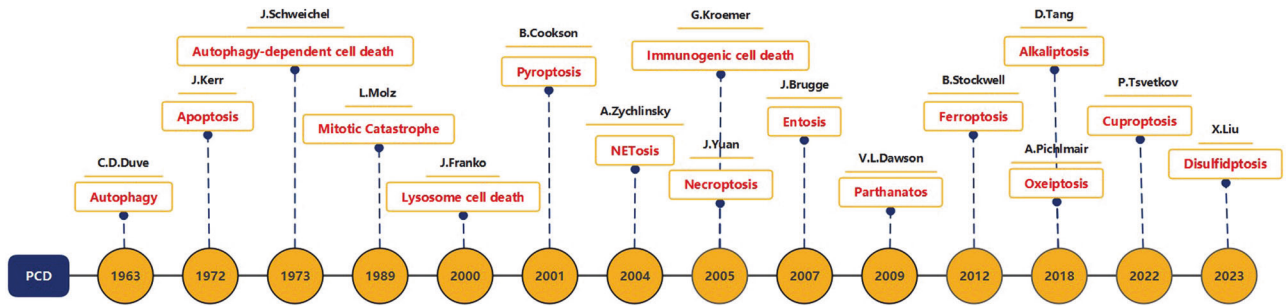
PCDs present visible morphological changes, Schweichel et al. classified PCD into three distinct morphological types based on morphology combined with the mechanism of cell death [1, 2]: (1) type I cell death or apoptosis; (2) type II cell death or autophagy; and (3) type III cell death or necrosis. Moreover, each type of PCD involves a specific molecular signaling pathway and regulatory system [3]. These regulated forms of cell death are intrinsically related to human embryonic development, homeostasis

<sup>1</sup>Suzhou Key Laboratory of Medical Biotechnology, Suzhou Vocational Health College, Suzhou, China. <sup>2</sup>Department of Thyroid and Breast Surgery, Wuzhong People's Hospital of Suzhou City, Suzhou, China. <sup>3</sup>School of Public Health, Nantong University, Nantong 226019, China. <sup>4</sup>Department of Health Technology and Informatics, The Hong Kong Polytechnic University, Kowloon, Hong Kong SAR, China. <sup>5</sup>State Key Laboratory of Radiation Medicine and Protection, Soochow University, Suzhou 215123, China. <sup>6</sup>These authors contributed equally: Yao Chen, Xiaohua Li. ✉email: minfeng.yang@connect.polyu.hk; liusongbai@126.com  
Edited by Boris Zhivotovsky

Received: 18 December 2023 Revised: 17 April 2024 Accepted: 29 April 2024

Published online: 10 May 2024





**Fig. 1** Timeline of PCD discovery. The discoverer and time of various types of PCD.

maintenance, and disease pathology, indicating broad prospects for clinical application in this field. With the continuous development of this field, new signaling pathways that regulate PCD are still being described.

In this article, we reviewed and analyzed the definitions, morphological characteristics, molecular mechanisms, and biomarkers of the main PCD types identified to date, which include apoptosis, necroptosis, pyroptosis, ferroptosis, autophagy-dependent cell death, mitotic catastrophe, immunogenic cell death, entosis, parthanatos, ferroptosis, disulfidptosis, NETosis, lysosome-dependent cell death, alkaliptosis, and oxeiptosis (Fig. 1), and briefly described the role and prospects of these major types of PCD in related diseases.

### MORPHOLOGICAL CLASSIFICATION OF THE PCDS

In 1990, Clarke et al. supplemented Schweichel and Merker's classification of cell death by dividing PCD into three types [4]. Type I refers to apoptosis (condensation, fragmentation, or phagocytosis). It is characterized by nuclear condensation and pyknosis, cell membrane coiling and blistering, cell size reduction, ribosome dissociation from polysomes and rough endoplasmic reticulum, and elimination of dead cells by heterophagolysosomes (heterophagy, engulfed by phagocytes after death). Type II is autophagic degeneration. It is characterized by the production of inwards bubbles in the cell membrane (endocytosis), the production of abundant autophagic vacuoles in the cytoplasm, the general expansion of the endoplasmic reticulum, mitochondria and Golgi apparatus, less obvious nuclear pyknosis than type I, and the elimination of dead cells by autolysosomes (autophagy). Type III is nonlysosomal vesicular degradation characterized by shrinkage and rounding or fragmentation of the cell membrane, edema, and dissolution or fragmentation of the nucleus, expansion of the endoplasmic reticulum, mitochondria, and Golgi apparatus, absence of early karyknosis, and dead cells that are not eliminated by lysosomes (cell dissolution *in situ*). Although this morphological classification is still widely used, it is mainly based on the phenotypes of the three pathways of apoptosis, autophagy, and necrosis, and these three types of characteristics are not fully representative of the other pathways in the currently known PCD type. Therefore, we optimized and supplemented the classical morphological classification, classified the known PCD pathways according to the characteristic morphology of the dead cells, how they were eliminated, whether they were accompanied by an inflammatory reaction, etc., and summarized the unique characteristics of various PCD pathways, thus providing a theoretical basis for the identification of various PCDS (Table 1).

### MOLECULAR MECHANISMS OF THE PCDS

#### Apoptosis

Apoptosis is a programmed and active death process that occurs in cells under the control of specific genes or pathways [5]. The morphological characteristics of apoptosis are as follows:

disappearance of cell junctions, reduction in volume, condensation of nuclear chromatin, nuclear lysis, cytoplasmic contraction, dilation of the endoplasmic reticulum and cell membrane blistering, and finally, division into apoptotic bodies by the cell membrane (Fig. 2A). Apoptotic bodies contain a variety of different fragments of organelles and chromatin with intact structures. The morphology of the mitochondria remained unchanged. Apoptosis is accompanied by the entire process of individual growth and development, showing the precise control of the type and number of cells.

The typical process of apoptosis involves a series of hydrolysis, caspase activation, and signal transduction processes; thus, this process is also called the caspase-dependent apoptosis pathway. Due to the different sources of apoptosis signals, classical apoptosis is divided into intrinsic apoptosis (mitochondrial apoptosis pathway) and extrinsic apoptosis (death receptor pathway) [6]. Intrinsic apoptosis is a form of PCD that initiates apoptosis due to disturbance of the intracellular microenvironment. Intrinsic apoptosis can be induced by growth factor deficiency, DNA damage, endoplasmic reticulum pressure, excess reactive oxygen species (ROS), replication pressure, microbundle changes, and mitotic defects [7]. The most important step in intrinsic apoptosis is mitochondrial outer membrane permeabilization (MOMP), which is mainly controlled by the BCL2 protein family. Among the inner and outer membranes of mitochondria, there are two main proapoptotic factors: cytochrome C (Cyt-C) and apoptosis-inducing factor (AIF). Apoptosis signals cause Cyt-C to be released from mitochondria to the cytoplasm, after which it binds to apoptotic protease activating factor 1 (APAF1), initiates a series of cascade reactions, and finally activates deoxyribonucleases and hydrolyzes nucleic acid and cytoskeleton proteins, leading to apoptosis; thus, the intrinsic pathway is also called the mitochondrial apoptosis pathway. Extrinsic apoptosis is a form of PCD caused by disturbance of the extracellular microenvironment. The extrinsic pathway induces apoptosis by activating specific death receptors on the cell surface through extrinsic death signals; thus, this pathway is also called the death receptor pathway [8]. At present, eight types of death receptors have been found, among which Fas (also known as CD95 or Apo-1) and tumor necrosis factor receptor (TNFR) are the most important. These death receptors contain a death domain of approximately 80 amino acids, which is necessary to mediate apoptosis [9, 10]. The apoptotic process of the death receptor pathway needs to be mediated by caspase8 activation [11]. Stimulation of the Fas receptor leads to its binding to the Fas ligand and Fas-associated death domain (FADD). FADD binds to pro-caspase8 to form a death-inducing signaling complex (DISC). Then, the DISC recruits pro-caspase8 to induce its activation through its own splicing, initiating the downstream caspase pathway and leading to apoptosis [12]. Both the intrinsic and extrinsic apoptosis pathways eventually undergo a cascade amplification process involving the irreversible limited hydrolysis of substrates mediated by caspase family members, which act on substrates and lead to apoptosis [13] (Fig. 3A). Caspases can be divided into two functional

**Table 1.** Morphological types of PCD and their specific characteristics.

Morphological types	PCD pathway	Specific morphological characteristics
Type I cell death (Apoptosis-like PCD) Cell shrinkage, Chromatin condensation, Eliminated by heterophagolysosome, Without inflammation.	Apoptosis	Cell contraction, Chromatin condensation, Small DNA fragments, Apoptotic body.
	Mitotic catastrophe	Chromatin Condensation, Multiple micronucleated giant cells. (most commonly induces Apoptosis)
	Immunogenic cell death	Morphological characteristics like Apoptosis.
	Oxeiptosis	Morphological characteristics like Apoptosis.
Type II cell death (Autophagy-like PCD) Cytoplasmic vacuolization, Formation of autophagic vesicles, Eliminated by autolysosomes.	Autophagy-dependent cell death	Inwards bubble of membrane, Autophagosomes form and aggregate.
	Entosis	Cell-in-cell structure.
	Lysosome-dependent cell death	Lysosome rupture.
Type III cell death (Necrosis-like PCD) Cells and organelles edema, Cell membrane destruction, Chromatin does not condense, Without the participation of lysosomes, With inflammation.	Necroptosis	Cells and organelles edema, Cell membrane disruption, Chromatin does not condense, Cell content release with inflammatory response.
	Pyroptosis	Morphological characteristics like Necroptosis.
	Ferroptosis	Mitochondrial atrophy, Mitochondrial membrane thickens, Mitochondrial membrane disruption, Mitochondrial ridge reduction.
	Parthanatos	Cell contraction, Chromatin Condensation, Large DNA fragments, Cell membrane disruption.
	Cuproptosis	Mitochondrial shrinkage, Mitochondrial membrane disruption.
	Disulfidptosis	Cell shrinkage, F-actin contraction and detachment from the plasma membrane.
	NETosis	Cell membrane rupture, Release network structure.
	Alkaliptosis	Morphological characteristics like Necroptosis.

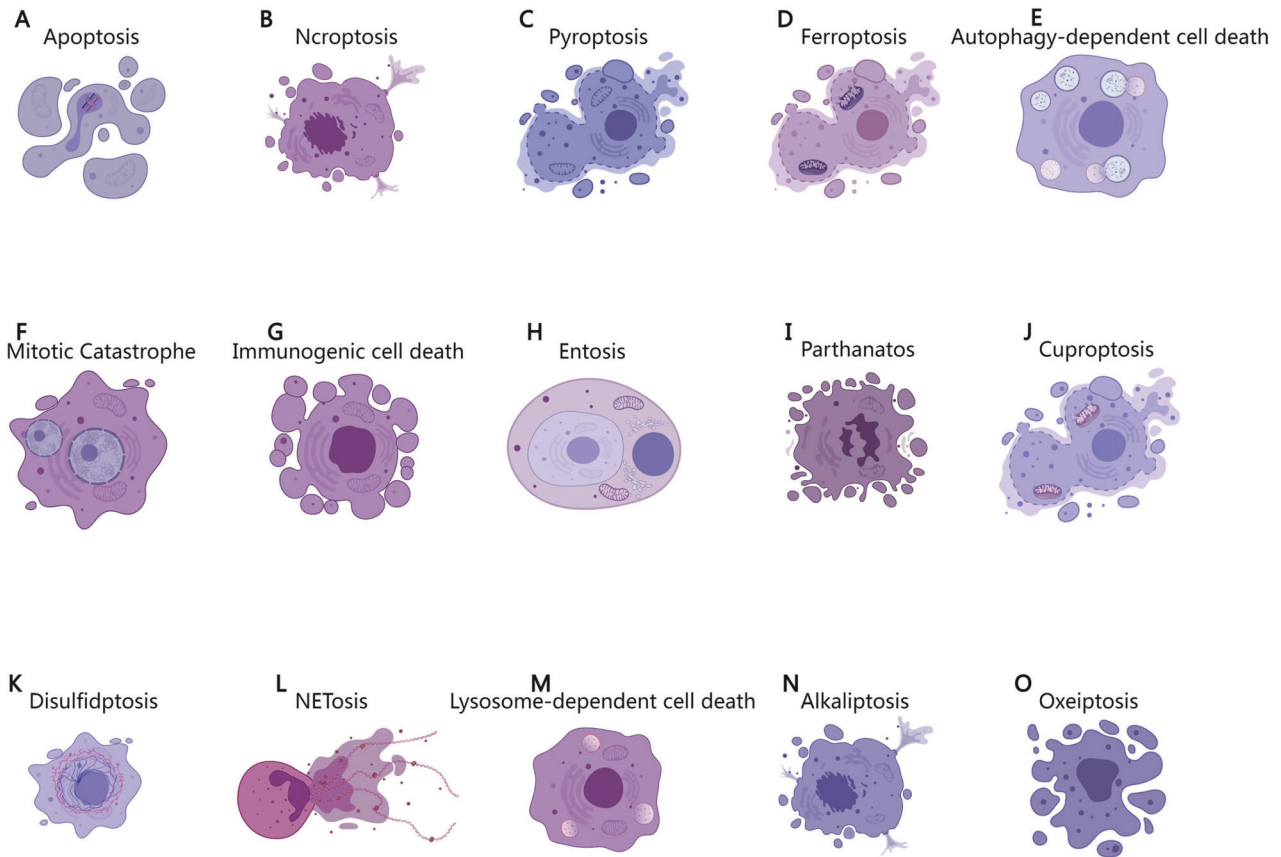
categories: inflammatory caspases, which include mainly caspase-1/4/5/11, which are involved in the processing of cytokines and the regulation of inflammation. The other type is proapoptotic caspases, which mainly include caspase-2/3/6/7/8/9/10, which are cascades involved in the process of apoptosis. The morphological characteristics of apoptosis are caused by proapoptotic caspases cutting intracellular proteins and inducing DNA cutting [14]. Caspases related to apoptosis can be divided into initial caspases (caspase 2/8/9/10) and effector caspases (caspase-3/6/7). These kinds of apoptosis regulatory factors need to be activated by the cleavage of pro-caspase. Downstream of the apoptosis pathway, its key biomarker is cleaved and activated caspase 3, making the cell irreversible to apoptosis. In addition, phosphatidylserine (PS), which is normally localized in the inner leaflet of the phospholipid bilayer of the cell membrane, is flipped to the outer leaflet when apoptosis occurs. Therefore, caspase3/8 cleavage and PS evagination can be used as the gold standards for detecting apoptosis. The study of apoptosis can not only be used for the early detection of cancer and improve the survival rate of cancer patients but can also be treated accurately and quickly by inducing tumor cell apoptosis.

### Necroptosis

Necroptosis is a form of PCD that is controlled by a unique caspase-independent signaling pathway and has necrosis-like morphological characteristics [15]. Compared with necrosis in

ACD, necroptosis involves the same subcellular changes, such as a sharp increase in intracellular peroxide, high phosphorylation of the mitochondrial membrane, increased membrane permeability, cell swelling, and cell membrane destruction (Fig. 2B). However, necroptosis is regulated by a variety of genes and is a regular mode of cell death. In contrast to apoptosis, necroptosis does not involve the formation of apoptotic bodies, and chromatin does not agglomerate.

Necroptosis can be triggered by Toll-like receptor activation, ROS accumulation in mitochondria, tumor necrosis factor- $\alpha$  (TNF- $\alpha$ ), and viral infection. When stimulated, caspase-inhibited cells may undergo necroptosis rather than apoptosis [16]. For example, TNF- $\alpha$  induced necroptosis in mouse fibroblast cells treated with the pan-caspase inhibitor Z-VAD-fmk, and it also caused necroptosis in caspase-8-deficient leukemic Jurkat cells. Thus, necroptosis may be the mechanism by which cells die when apoptosis fails to initiate normally [17]. Different cells undergo apoptosis or necroptosis according to their environment and degree of activation. In the signaling pathway, TNF- $\alpha$  binds and activates TNFR1 to induce necroptosis. After TNF- $\alpha$  binds TNFR1, TNFR1 recruits a series of proteins to form different complexes on the cytoplasmic side. Among them, complex I include TNFR-associated death domain (TRADD), receptor-interacting protein kinase 1 (RIPK1), TNFR-associated factor 2 (TRAF2), TRAF5, cellular inhibitor of apoptosis protein 1 (cIAP1), cIAP2, and the ubiquitinase complex. At this point, if RIPK1 is polyubiquitinated, it will



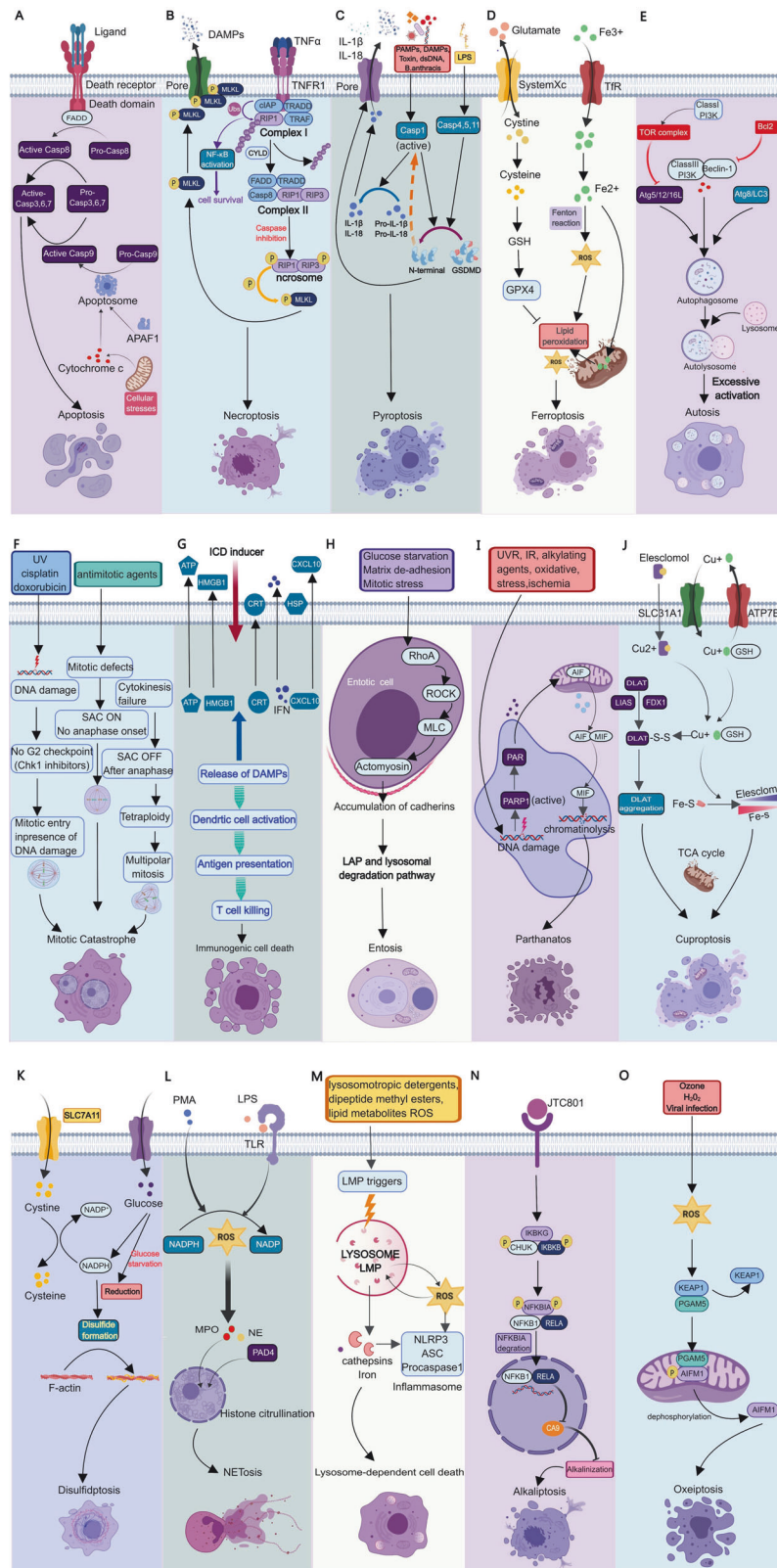
**Fig. 2 Morphological characteristics of various types of PCD.** **A** Apoptosis: cell contraction, chromatin condensation, small DNA fragments, and apoptotic body; **B** necroptosis: cell membrane disruption, cell and organelle edema, chromatin does not condense, cell content release with inflammatory response; **C** pyroptosis: morphological characteristics like necroptosis; **D** ferroptosis: mitochondrial atrophy, mitochondrial membrane thickening, mitochondrial membrane disruption, mitochondrial ridge reduction; **E** autophagy-dependent cell death: cell membrane disruption, cytoplasmic vacuolation, autophagosomes form and aggregate; **F** mitotic catastrophe: chromatin condensation, multiple micronucleated giant cells; **G** immunogenic cell death: morphological characteristics like apoptosis; **H** entosis: cell-in-cell structure; **I** parthanatos: cell contraction, chromatin condensation, large DNA fragments, cell membrane disruption; **J** cuproptosis: mitochondrial shrinkage, mitochondrial membrane disruption; **K** disulfidptosis: cell shrinkage, F-actin contraction and detachment from the plasma membrane; **L** NETosis: cell membrane rupture, release network structure; **M** lysosome-dependent cell death: lysosome rupture; **N** alkaliptosis: morphological characteristics like necroptosis; **O** oxeiptosis: morphological characteristics like apoptosis.

further form a complex to activate the signaling pathway mediated by nuclear factor- $\kappa$ B (NF- $\kappa$ B) and mitogen-activated protein kinase (MAPK) and subsequently inhibit cell death [18]. TNFR1 dissociates from complex I, and RIPK1 deubiquitinates and forms complex II with RIPK3, TRADD, and Fas-associated protein via a death domain (FADD) and caspase-8. In this complex, if caspase-8 inactivates RIPK1 and RIPK3, the cell will die by apoptosis; if caspase-8 is inhibited, RIPK1 and RIPK3 bind to each other through their respective RHIM domains to form a necroptosis complex. In this case, RIPK3 autophosphorylates, recruits, and phosphorylates MLKL, and then initiates necroptosis [19, 20]. MLKL can serve as a platform for the recruitment of  $\text{Ca}^{2+}$  and  $\text{Na}^{+}$  in the plasma membrane and can also play a role in the formation of pore complexes in the membrane [21, 22]. Therefore, the activation of MLKL by RIPK3 is a key regulatory pathway of necroptosis, and the phosphorylation of MLKL has become a biomarker of necroptosis [23] (Fig. 3B). The occurrence of necroptosis can be marked by detecting the phosphorylation of key necroptosis molecular markers, such as RIPK1, RIPK3 and MLKL, while necrostatin-1, a specific inhibitor of RIPK1, can block the occurrence of necroptosis. Necroptosis is involved in a variety of pathological processes in the body, such as bacterial and viral infections or inflammatory lesions caused by sterile lesions such as atherosclerosis. In addition, necroptosis is also considered a possible barrier against tumor formation.

### Pyroptosis

Pyroptosis is a form of PCD that is related to the innate immune response (such as pathogen invasion) and depends on Gasdermin family proteins to form plasma membrane pores [24, 25]. This process is often but not always completed by the activation of inflammatory caspases, resulting in necrosis-like morphological characteristics similar to those of necrosis in ACD, including nuclear fragmentation and dissolution; increased cell membrane permeability; swelling and lysis; and the release of cellular contents, which cause a local inflammatory response (Fig. 2C).

The molecular mechanisms of pyroptosis include classical and nonclassical pyroptosis pathways. Classical pyroptosis mainly depends on the activation of Caspase-1. When pathogens invade host cells, caspase-1 can be activated by binding to inflammatory bodies (pyroptosomes), which are a variety of protein signaling complexes, and their central scaffolds include the inflammatory bodies NLRP1, NLRP3, NAIP-NLRC4, AIM2 and pyrin [26]. After caspase-1 activation, a heterodimer with enzyme activity is formed, which specifically cleaves GSDMD in the Gasdermin domain protein family to release the N-terminal domain, inserts the cell membrane lipid bilayer, and induces oligomerization in the membrane to form pores [27], resulting in the destruction of cell membrane osmotic pressure, cell swelling and rupture, content release and inflammation, resulting in pyroptosis. Moreover, the activation of caspase-1 can also promote the activation



**Fig. 3 Signaling pathways of various types of PCD. A Apoptosis; B Necroptosis; C Pyroptosis; D Ferroptosis; E Autophagy-dependent cell death F Mitotic catastrophe; G Immunogenic cell death; H Entosis; I Parthanatos; J Cuproptosis K Disulfidptosis; L NETosis; M Lysosome-dependent cell death; N Alkaliptosis; O Oxoeptosis.**



and secretion of the inflammatory cytokines IL-18 and IL-1 $\beta$  to recruit more immune cells to the infected site and expand the inflammatory response. Nonclassical pyroptosis is premised on the activation of caspase-4,5,11. Caspases 4, 5, and 11 can be activated by direct binding to bacterial lipopolysaccharide (LPS), and activated caspases 4, 5, and 11 can also lead to pyroptosis by cleaving GSDMD, so caspase cleavage of the GSDMD substrate is a key execution event of pyroptosis [28]. In addition, upon specific binding of LPS to caspase-11, activated caspase-11 can also activate the Pannexin-1 transmembrane channel, causing the outflow of the intracellular danger signal molecule ATP and the opening of nonselective positive ion channels. Then, the outflow of intracellular K<sup>+</sup> and the influx of extracellular Na<sup>+</sup> and Ca<sup>2+</sup> will damage the integrity of the cell membrane and release the inflammatory contents in the cells, which will trigger the inflammatory response [29]. Other studies have shown that caspase-1 is also involved in the nonclassical pyroptosis pathway mediated by Caspase-11, and the activation of caspase-1 can release the proinflammatory factors IL-1 $\beta$  and IL-18 [30] (Fig. 3C). The specific biomarker of pyroptosis is the cleavage of Gasdermin family proteins, so pyroptosis can be indicated by detecting the N-terminal domain released after the cleavage of Gasdermin proteins. Pyroptosis is an important immune response of the body that is closely related to infectious diseases, cardiovascular diseases, nervous system diseases, and tumors.

### Ferroptosis

Ferroptosis is a type of PCD caused by iron-dependent oxidative damage. Ferroptosis is characterized by disordered intracellular iron ion flow and significant increases in ROS and lipid peroxide levels without the need for caspases [31–35]. Ferroptosis is completely different from other forms of death in morphology, biochemistry, and genetics. The formation of apoptotic bodies, DNA fragmentation, and activation of the Caspase family are not observed in ferroptosis, and these effects cannot be reversed by Caspase inhibitors. The morphological characteristics of ferroptosis mainly include increased mitochondrial membrane density, a significant reduction in the number of mitochondrial ridges, membrane rupture, and overall mitochondrial atrophy; the nuclei were intact and normal in size, and no condensation of chromatin was observed; the cell membrane did not bleb, but the membrane density increased; and both the structure of the membrane phospholipid bilayer and the fluidity of the cell membrane changed [36] (Fig. 2D).

A variety of substances and external conditions can trigger ferroptosis, and the main stimulus signals are related to phospholipids. Ferroptosis is caused by the accumulation of ROS on membrane lipids due to the lack of the membrane lipid repair enzyme glutathione peroxidase (GPX4), and this accumulation requires iron ions. During ferroptosis, a large amount of free Fe<sup>2+</sup> accumulates in the cell. Free Fe<sup>2+</sup> is highly oxidized and prone to the Fenton reaction with H<sub>2</sub>O<sub>2</sub>, which produces hydroxyl radicals that can cause oxidative damage to DNA, proteins, and membrane lipids, promoting lipid peroxidation, damaging the cell membrane, and leading to cell death [37]. Lipid peroxidation refers to the loss of hydrogen atoms in lipids under the action of free radicals or lipid peroxidases, which leads to the oxidation, breakage, and shortening of lipid carbon chains and the production of lipid free radicals, lipid hydroperoxide (L-OOH), reactive aldehydes (malondialdehyde, 4-hydroxynonenal) and other cytotoxic substances, resulting in cell damage from lipid oxidative degradation [38]. GSH is an important antioxidant in the human body. It not only reduces H<sub>2</sub>O<sub>2</sub> to H<sub>2</sub>O, a scavenger of free radicals, and maintains the balance of intracellular free radical content but also participates in the reduction of L-OOH as a cofactor of GPX4, repairing L-OOH in biofilms and preventing ferroptosis. The depletion of GSH leads to the inactivation of GPX4 and increases the lipid peroxidation reaction occurring on the inner side of the cell membrane

phospholipids, which leads to membrane breakage, cell disintegration, and death. Thus, GSH depletion and inhibition of GPX4 enzyme activity are necessary for cells to undergo ferroptosis (Fig. 3D). Some substances have been found to be effective in inhibiting or promoting ferroptosis. Ferroptosis activators can be roughly divided into three categories according to their targets. The first class comprises System Xc inhibitors, such as erastin and sorafenib. Erastin decreases the uptake of cysteine by cells by inhibiting System Xc, leading to the depletion of GSH, the substrate of GPX4, and further reducing GPX4 activity, leading to the accumulation of ROS and ferroptosis. The second class includes GPX4 inhibitors, such as RSL3 and FIN56. RSL3 can bind to GPX4 and inhibit its protein activity, leading to the accumulation of toxic L-OOH and triggering ferroptosis [39]. The third class includes GSH-depleting agents, such as cysteinase, BSO, and cisplatin. Cysteinase can directly degrade cysteine and block the synthesis of GSH [40]. In addition, some activators, such as ferric citrate, can trigger ferroptosis by inducing lipid peroxidation or increasing intracellular free iron ion levels. Most ferroptosis inhibitors are iron chelators or antioxidants, such as deferoxamine mesylate (DFO) and ferrostatin, which inhibit the production of lipid ROS by scavenging free iron ions, hydroxyl radicals, or lipid free radicals, thus effectively reducing or eliminating the damage caused by ferroptosis. The main causes of ferroptosis are an imbalance in cell metabolism and the accumulation of ROS, which can be confirmed by detecting the degree of ROS accumulation and ROS products in cells. Furthermore, there is some more direct evidence. For example, ferroptosis can be inhibited by several known ferroptosis inhibitors (iron chelators, antioxidants, etc.), and relatively specific morphological phenomena, such as mitochondrial atrophy, reduction or even disappearance of mitochondrial ridges, and increased membrane density, can be observed under an electron microscope to prove the occurrence of ferroptosis. In recent years, studies have shown that ferroptosis is associated with many diseases, such as Parkinson's syndrome, breast cancer, and pancreatic cancer, among which malignant tumors are most closely related. Some tumor cells are quite sensitive to ferroptosis. For example, dihydroartemisinin can induce ferroptosis in squamous cell carcinoma of the head and neck, thereby inhibiting tumor growth [41]. Therefore, ferroptosis is expected to become a new direction for disease treatment.

### Autophagy-dependent cell death

Autophagy-dependent cell death (ADCD) is a kind of PCD that must be driven by the molecular mechanism of autophagy in the process of cell death [1, 42], and its morphological feature is the observation of double-membrane phagocytic components containing vesicles, namely, autophagosomes; the expansion and fragmentation of the endoplasmic reticulum, mitochondria and other organelles; and the slight agglutination of chromatin [43, 44] (Fig. 2E).

Autophagy is a highly conserved catabolic mechanism that relies on lysosomes to degrade aging components in cells. It is often triggered by conditions such as nutrient deprivation, pathogen infection, hypoxia, or endoplasmic reticulum stress. Cells use a double-layer membrane structure to enclose pathogens, cytoplasmic proteins, or organelles to form autophagosomes or vesicles. Then, these substrates fuse with lysosomes and are degraded, thereby removing excessive or abnormal proteins, organelles, or pathogenic microorganisms. This mechanism is conducive to the maintenance of cell homeostasis and the renewal of organelles. However, excessive autophagy can lead to cell death, that is, ADCD, indicating the existence of a threshold effect of autophagy. The regulatory mechanism of autophagy is very complex and involves multiple regulatory pathways and regulators. According to current studies, autophagy is regulated mainly by the phosphatidylinositol 3-phosphate kinase-mammalian target of rapamycin (PI3K-mTOR) signal transduction pathway

upstream of autophagy-associated genes (ATG) and the Beclin1 complex [45, 46]. Type I PI3K induces the generation of phosphatidylinositol 3,4,5-triphosphate (PIP3) in response to a stimulus signal. Under the action of 3-phosphatidylinositol-dependent kinase 1 (PDK1), protein kinase B (AKT) binds to PIP3 and is activated to form phosphorylated AKT. It further acts on the downstream target protein mTOR, thereby negatively regulating autophagy [38]. The mTOR pathway is considered to be the “gatekeeper” and “monitor” of autophagy and has an inhibitory effect on autophagy [47]. ATG can also form complexes with other components to regulate autophagy. Beclin1, the first autophagy-related gene discovered, regulates the formation of autophagosomes by binding to type III PI3K to form complexes, thereby promoting the occurrence of ADCD [48]. The combination of Beclin1 and Bcl2 inhibited autophagy. Two ubiquitin-like pathways, the Atg5-Atg12-Atg16L complex, and the Atg8/LC3 system, play crucial roles in autophagosome formation (Fig. 3E). Therefore, LC3 lipidation and an increase in the LC3-II/LC3-I ratio are considered biomarkers of autophagy. Autosis is a subtype of ADCD that relies on the plasma membrane  $\text{Na}^+/\text{K}^+$ -ATPase [43]. It can be induced by nutrient deprivation or by Tat-Beclin1 and inhibited by blocking upstream  $\text{Na}^+/\text{K}^+$ -ATPase, which is a plasma pump linking ion homeostasis and ER stress [43]. ADCD may play an important role in neuronal cell death induced by neurotoxicity or hypoxia-ischemia.

### Mitotic catastrophe

Mitotic catastrophe (MC) is a regulated tumor suppressor mechanism in which the mitotic process of cells is dysregulated due to DNA damage or other causes, resulting in cell death accompanied by tetraploid or polyploid formation [49]. It was first identified in a heat-sensitive yeast mutant strain that dies as a result of abnormal chromosome segregation during cell division [50, 51]. Its main morphological features are nuclear micronization and polykeratosis (cell size becomes larger, forming giant cells containing two or more nuclei and partially condensed chromatin) [52] (Fig. 2F).

Current studies have shown that DNA damage, mitotic defects, and cytokinesis failure are the main causes of MC, which are regulated by a variety of molecules and closely related to cell cycle checkpoints and cell cycle-related kinase abnormalities. The induction of MC in cells can constitute a new target for cancer therapy. DNA damage can be caused by internal factors (such as replication errors or cell-generated ROS) or external environmental factors (such as radiation and platinum compounds), thereby disrupting the integrity and stability of the cellular genome. When the cell undergoes DNA damage and the G2 checkpoint is defective or damaged in the cell cycle, the cells with DNA damage enter the M phase of mitosis prematurely through the G2 checkpoint, resulting in chromosome segregation error and subsequent MC [53]. Therefore, an abolished or defective G2 checkpoint is essential for DNA damage-induced MC. In addition, during mitosis, the precise segregation of sister chromosomes is controlled by the spindle assembly checkpoint (SAC), also known as the mitotic checkpoint [54]. The SAC can prevent cells from entering anaphase during mitosis before the bilateral kinetochore of all sister chromosomes forms proper attachments to their respective bipolar spindle microtubules to prevent errors during chromosome segregation. The mitotic checkpoint signaling pathway is activated when spindle microtubules do not connect or misconnect kinetochores on both sides of chromosomes due to reasons such as a lack of mitotic organs. After the initiation of SAC, it can effectively inhibit the activity of the anaphase-promoting complex (APC) and prevent the continuation of mitosis [53]. However, long-term activation of the SAC can lead to mitotic arrest, resulting in mitotic defects in cells and, thus, in MCs [55]. Therefore, when the mitotic checkpoint signal is destroyed or activated for a long time, the cell can initiate anaphase or stay in

metaphase before the chromosome kinetochores have all established the correct connection with the spindle microtubule, resulting in mitotic defects, chromosome missegregation, and then aneuploidy, which further leads to the occurrence of MC. Defects in cyclins or related kinases can also lead to the inhibition of cytokinesis, resulting in polyploidy and multipolarized mitosis at anaphase of mitosis, resulting in genomic instability and stimulating MCs in the next cell cycle [53] (Fig. 3F). Related kinases include mitotic kinases such as Aurora kinases, monopolar spindle 1 (MPS1), and polo-like kinases (Plks), which play key roles in proper chromosome segregation. The exact molecular mechanism by which mitotic alterations are sensed and trigger the MC cascade is unknown, but p53 may be involved [56]. Extensive experimental evidence suggests that MC is facilitated by a signal transduction cascade dependent on caspase-2 activation, often (but not always) triggering intrinsic apoptosis regulated by the BCL2 protein family [57]. However, it has also been found that in certain cases where p53 is lacking, mitotic defects appear to drive a necrotic variant of PCD independent of caspase-2 signaling [58]. In addition, although apoptosis usually occurs when cells are in an abnormal mitotic state for a long time, there are rare cases in which MC does not lead to cell death but eventually results in senescence, a type of irreversible cell cycle arrest [59]. Thus, MCs do not always cause PCD (and can also drive cellular senescence), and the Committee on Cell Death (NCCD) recommends the use of the term mitotic death to denote PCD driven by mitotic mutations (most commonly intrinsic apoptosis) [25]. MC is biologically characterized by low expression or absence of proteins associated with G2/M phase checkpoints and mitotic spindle assembly. The detection methods mainly include optical microscopy, laser confocal microscopy, or electron microscopy to detect cells containing two or more nuclei, and flow cytometry to detect G2/M block and polyploidy. In recent years, inducing tumor cells to cross the cell cycle checkpoint with a large amount of DNA damage or errors, leading to MC and death of tumor cells, has made an important attempt at clinical treatment and targeted drug development for cancer.

### Immunogenic cell death

Immunogenic cell death (ICD) is a specific variant of PCD that is driven by stress and can induce adaptive immune responses against dead cell antigens [60, 61]. The morphological features of ICD were similar to those of apoptosis (Fig. 2G).

ICD can be caused by a variety of different stressors, including but not limited to (1) intracellular pathogens; (2) traditional chemotherapeutic drugs such as anthracyclines, DNA damage agents, and proteasome inhibitors; (3) targeted anticancer drugs such as crizotinib (a tyrosine kinase inhibitor), cetuximab (an epidermal growth factor-specific monoclonal antibody), and poly-ADP ribose polymerase (PARP) inhibitors; (4) A variety of physical therapies, including radiotherapy, external photochemotherapy, photodynamic therapy, near-infrared immunotherapy, and nanopulse stimulation [62–65]. These stressors stimulate cells to produce a series of signaling molecules called damage-associated molecular patterns (DAMPs). DAMPs released during ICD can bind to pattern recognition receptors (PRRs) on the surface of dendritic cells (DCs) and initiate a series of cytological responses that ultimately activate innate and adaptive immune responses [66]. At this point, when the target cells exhibit sufficient antigenicity, their death is executed by cytotoxic T lymphocytes (CTLs), which induce target cell apoptosis through the perforin/granzyme pathway and the death receptor pathway and trigger an adaptive immune response involving immune memory (Fig. 3G). To date, six DAMPs have been linked to the immunogenic mechanism of ICD: (1) calreticulin (CALR/CRT) [67], (2) adenosine triphosphate (ATP) [68], (3) high mobility group box protein B1 (HMGB1) [69], (4) type I interferon (IFN) [70], (5) cancer cell-derived nucleic acids [67, 71], and (6) annexin A1 (ANXA1) [72].

The DAMPs released under the above different induction conditions can be used as biomarkers for subsequent studies. ICD and its DAMPs provide a new therapeutic basis and means for tumor therapy, monitoring changes in tumor cell immunogenicity before and after chemotherapy, and combining chemotherapy and immunotherapy can improve the therapeutic effect on tumors. In the treatment of tumors, chemotherapy drugs or radiotherapy induce the death of tumor cells and upregulate the expression of certain immune signaling molecules, such as CALR, on the surface of target cells. These signaling molecules enhance the ability of DCs to recognize tumors and present antigens. After DCs are stimulated to mature, they activate tumor-specific CTLs to attack tumors and stimulate the release of interleukin (IL)-2, IL-4, and IFN- $\gamma$  to obtain more ideal antitumor therapeutic effects.

### Entosis

Entosis, also known as entotic cell death, is a form of cellular cannibalism that occurs in healthy and malignant mammalian tissues and is characterized by one cell engulfing and killing another and the formation of a “cell-in-cell” structure (CIC) [73, 74] (Fig. 2H).

Cell adhesion and cytoskeletal rearrangement pathways play important roles in the control of entosis induction and occurrence. Entosis is activated to phagocytose and kill similar cells through LC3-associated phagocytosis (LAP)- and cathepsin B (CTSB)-mediated lysosomal degradation pathways. After detachment from the extracellular matrix (ECM), epithelial cells first adhere to adjacent epithelial cells through the adhesion proteins E-cadherin/cadherin-herin 1 (CDH1) and alpha-E-catenin. Subsequently, the bacterium will invade another cell through actomyosin complex formation, the lysosomes of the invading cell will encapsulate the invasive cell, and the invasive cell will die through the lysosomal pathway [75]. There are two key factors that regulate entosis. One is the adhesion of cadherin, so calcium blockers can prevent entosis; the first adhesion protein family molecule, PCDH7, which negatively regulates entotic CIC structure formation, has also been recently discovered [76]. Another key factor is the Rho-ROCK signaling pathway, in which the expression of actomyosin complexes in the cytoskeleton is dependent on RAS homolog gene family member A (RhoA), Rho-associated protein kinase 1 (ROCK1), ROCK2 and related diaphanous factor (DIAPH1), which accumulate local activity. Therefore, blocking the Rho-ROCK signaling pathway can prevent cell invasion, thereby preventing the occurrence of entosis [73] (Fig. 3H). Studies have shown that relative intracellular Rho GTPase activity is a key factor regulating the selective clearance behavior of entosis. Malignant tumor cells with low Rho GTPase activity easily deform and engulf normal cells or benign tumor cells with high Rho GTPase activity. Therefore, in the pathological context of tumors, entosis is thought to be a mechanism of cell competition to select the best tumor cell clones by internalizing and killing failing cells [77] and to promote tumor evolution by inducing genomic instability in external cells [78]. Moreover, its role in the physiological environment has gradually been revealed. Recent studies have revealed the mechanism of mitosis surveillance in patients with entosis. This mechanism selectively promotes the penetration of aneuploid daughter cells into adjacent cells to form CIC structures by activating the p53 signaling pathway, and CIC is subsequently cleared to maintain the genomic stability of epithelial cells [79].

### Parthanatos

Parthanatos, also known as poly-ADP ribose polymerase-1 (PARP-1)-dependent cell death, is a new form of PCD in which PARP-1 is overactivated due to DNA damage [80]. It should be noted that in addition to DNA damage, oxidative stress, hypoxia, hypoglycemia, or inflammatory signals may also trigger the production of parthanatos [80]. The morphological characteristics of parthanatos are cellular atrophy, chromatin concentration, and plasma

membrane rupture. Parthanatos had some characteristics similar to apoptosis and necrosis-like morphology, but there were also obvious differences. Compared with apoptosis, parthanatos cannot form apoptotic bodies or small DNA fragments. Compared with necrosis in ACD, parthanatos is unable to induce cell and organelle swelling or cytolysis [81] (Fig. 2I).

Mechanistically, the release of apoptosis-inducing factor mitochondrion-associated 1 (AIFM1) from mitochondria-mediated by poly-ADP ribose (PAR) or calpain may be responsible for the effects of parthanatos. When PARP-1 is activated in large quantities, it can use coenzyme I (NAD) and ATP as substrates to catalyze the synthesis of PAR, and PAR induces the migration of AIFM1 in mitochondria (referring to the process of AIFM1 entering the nucleus from mitochondria), which further induces chromatin condensation and large DNA fragmentation to complete the transmission of death signals [82] (Fig. 3I). Some scholars also believe that the release of AIFM1 is dependent on calpain activation and that calpain I can promote the formation and release of mature AIF in a cell-free state [83]. Thus, parthanatos has distinct biochemical characteristics, such as rapid activation of PARP-1, early PAR accumulation, changes in mitochondrial permeability, AIFM1 migration from mitochondria to the nucleus, and intracellular NAD and ATP consumption. Parthanatos is involved in a variety of diseases of the human body and plays a key role in various nervous system diseases.

### Cuproptosis

Cuproptosis is a new form of PCD that depends on the accumulation of intracellular copper and is triggered by the direct combination of copper with the fatty acylated components of the tricarboxylic acid cycle (TCA) [84]. The morphological features of cuproptosis include mitochondrial shrinkage, cell membrane disruption, and endoplasmic reticulum and chromatin damage (Fig. 2J).

Excess copper promotes the aggregation and functional loss of lipoacylated proteins, triggers the loss of iron-sulfur cluster proteins, blocks the TCA cycle, and leads to cytotoxic stress and eventual death. The addition of Elesclomol alone did not affect cell growth, but once copper ions were added, cell growth was inhibited. Therefore, the induction of cell death by copper-binding molecules or copper ionophores mainly depends on the availability of copper. In addition, cuproptosis requires the involvement of mitochondrial respiration, and copper is not directly involved in the electron transport chain (ETC) but only plays a role in the TCA cycle. It was subsequently confirmed that the key regulatory gene of cuproptosis is FDX1, which is the upstream regulator of protein lipoacylation. FDX1 encodes a protein that is directly targeted by Elesclomol, which converts  $\text{Cu}^{2+}$  to more toxic  $\text{Cu}^{+}$  and catalyzes the lipoacylation of DLAT, DLST, and LIST in the pyruvate dehydrogenase complex.  $\text{Cu}^{+}$  further binds to the lipoacylation site of DLAT, resulting in the oligomerization of DLAT, thus resulting in copper toxicity (Fig. 3J). Loss of FDX1 leads to a complete loss of protein lipoacylation while also causing a significant decrease in cellular respiration and attenuating copper ionophore-induced cell death. The abundance of FDX1 and lipoacylated proteins is highly correlated with a variety of human tumors. Cell lines with high levels of lipoacylated proteins were shown to be more sensitive to cuproptosis. This finding suggests that copper ionophores may be a potential therapeutic approach for targeting cancer cells with such metabolic profiles. Such a mechanism could explain the pathology associated with inherited copper overload diseases and suggest novel approaches to exploit copper toxicity in the treatment of cancer.

### Disulfidptosis

Disulfidptosis is a form of PCD that differs from apoptosis, ferroptosis, etc., and has not been characterized previously.

Disulfide stress causes rapid death caused by excessive intracellular cystine accumulation in glucose-starved SLC7A11-overexpressing cells [85].

Solute carrier family 7 member 11 (SLC7A11) is a cystine/glutamate antiporter that is mainly involved in the transport of amino acids on the plasma membrane. It is also an important pathway for cancer cell survival. Studies have shown that the process of SLC7A11-mediated reduction of ingested cystine to cysteine is highly dependent on the reduced nicotinamide adenine dinucleotide phosphate (NADPH) generated by the glucose-pentose phosphate pathway [86]. Therefore, during glucose deprivation, NADPH in SLC7A11-overexpressing cells is rapidly depleted, and cystine and other disulfides abnormally accumulate, which induces disulfide stress and rapid cell death. Further studies revealed that thiol oxidants promoted cell death in cells with high SLC7A11 expression under glucose starvation conditions and led to a sharp accumulation of intracellular disulfide molecules. However, reducing agents of disulfide stress, such as dithiothreitol (DTT),  $\beta$ -mercaptoethanol (2ME), and TCEP, can completely inhibit glucose starvation-induced cell death in SLC7A11-overexpressing cells [85]. Unlike other cell death mechanisms, disulfidptosis is associated with the actin cytoskeleton. In SLC7A11-overexpressing cells, the proteins with the most significant increase in disulfide bonds during glucose deprivation were enriched mainly in biological processes or pathways related to the actin cytoskeleton and cell adhesion, which can cause abnormal disulfide bonds of actin skeleton proteins in cells and lead to subsequent F-actin contraction [85] (Fig. 3K). Thus, glucose starvation in SLC7A11-overexpressing cells induces significant changes in cell morphology, including cell shrinkage, F-actin contraction, and detachment from the plasma membrane (Fig. 2K). In addition, glucose is the starting material of glycolysis and is transported by the glucose transporter (GLUT) family through the cell membrane, so the GLUT inhibitors KL-11743 and Bay-876 can effectively inhibit glucose uptake, similar to glucose starvation. Studies have confirmed that GLUT inhibitors can induce disulfur status and cell death in cancer cells with high SLC7A11 expression, and cancer cell disulfidptosis may be a key factor in the therapeutic effect of GLUT inhibitors in the treatment of tumors with high SLC7A11 expression [86].

### Other types of PCD

NETosis, a form of PCD driven by neutrophil extracellular trap (NET) particles, is also known as NETotic cell death [87]. NETosis was initially found to be related to the extrusion of fibrous webs containing chromatin and histone proteins in neutrophils, which was later confirmed to be the extracellular network DNA protein structure released by cells in response to infection or injury [88]. The morphological characteristics of NETosis include nuclear swelling, nuclear membrane and cytoplasmic granule membrane lysis, cytoplasmic membrane rupture, chromatin contact with cytoplasmic granules, and subsequent discharge to the outside of the cell to form a grid structure (Fig. 2L). In NETosis, ROS and activated protein-arginine deiminase 4 (PAD4) are produced through the activation of NADPH oxidase, leading to chromatin densification and myeloperoxidase (MPO) and neutrophil elastase (NE) entry into the nucleus, promoting the expansion of additional chromatin and nuclear membrane destruction. After chromatin is released into the cytosol, it is decorated by cytoplasmic and granular proteins, which eventually leads to plasma membrane rupture, resulting in NET release and neutrophil death (Fig. 3L). Among them, histone citrullination and ROS production are needed. Originally identified as a means of neutrophil defense against pathogens, NETosis also occurs in aseptic inflammation, promotes thrombosis, and can mediate tissue damage. In addition, NETosis is also a type of PCD that can occur in autoimmune diseases and has significant therapeutic potential.

Lysosome-dependent cell death (LDCD) is a form of PCD mediated by lysosomal contents (including proteolytic enzymes of the cathepsin family) or iron released after enhanced lysosomal membrane permeabilization (LMP) [89, 90]. It is characterized by lysosomal disruption [91] (Fig. 2M). The accumulation of intracellular ROS or the accumulation of lipid oxides can lead to lysosomal rupture, and proteolytic enzymes in lysosomes are released into the cytoplasm, leading to the occurrence of LDCD. Cathepsin is the main executor of LDCD, and blocking the expression or activity of cathepsin can reduce the occurrence of LDCD (Fig. 3M). Lysosome-dependent cell death is associated with a variety of pathophysiological conditions, including inflammation, tissue remodeling, aging, neurodegeneration, cardiovascular disease, and intracellular pathogen response.

Alkaliptosis, a Ph-dependent novel form of PCD driven by intracellular alkalization, has recently been identified as a new strategy for the treatment of a variety of tumors, especially pancreatic cancer [92, 93]. Studies have shown that intracellular alkalization can lead to JTC801-induced alkalosis, while oxidative stress is not necessary; therefore, alkaliptosis is a form of intracellular alkali-dependent regulatory necrosis with necrosis-like morphological characteristics [93] (Fig. 2N). At the molecular level, alkaliptosis may be mediated by inhibitor of nuclear factor kappa B kinase subunit  $\beta$  (IKKB), which can induce the downregulation of nuclear factor  $\kappa$ B pathway-dependent carbonic anhydrase IX (CA9), leading to alkaliptosis (Fig. 3N). Therefore, IKKB can mediate the occurrence of alkaliptosis, while CA9 can prevent the occurrence of alkaliptosis; however, the exact mechanism involved remains unclear.

Oxeiptosis is a novel form of noninflammatory PCD induced by oxygen-free radicals and independent of caspases [94]. Oxeiptosis is morphologically characterized by apoptosis-like cell death (Fig. 2O). This process involves the interaction of a cellular ROS sensor, the antioxidant factor KEAP1, the phosphatase PGAM5, and the proapoptotic factor AIFM1 [94]. Hyperactivated KEAP1 mediates  $H_2O_2$ -induced oxeiptosis in an NFE2L2-independent manner by dephosphorylating AIFM1 at Ser116 through an interaction pathway involving Keap1-PGAM5 (Fig. 3O). Oxeiptosis is important for preventing inflammation caused by ROS or ROS-producing agents such as viral pathogens.

### CONCLUDING REMARKS

In recent years, with the in-depth study of PCD, several new cell death pathways have been identified. Among the PCDs, apoptosis mainly shows unique morphological characteristics, such as membrane integrity, cell shrinkage, and apoptotic body formation, and does not cause an inflammatory response, while necroptosis, pyroptosis, ferroptosis, and some niche cell death forms, such as NETosis, mainly show necrosis-like morphological characteristics, such as cell swelling, cell membrane rupture, organelle collapse, and an inflammatory response. In addition, it is worth noting that some of the forms of niche cell death mentioned above may not be a separate type of cell death or an independent signaling pathway but rather a side effect or phenomenon that occurs in typical cell death processes, such as apoptosis, necroptosis, pyroptosis and ferroptosis. Growing evidence highlights the "communication" between different cell death pathways [95].

Multiple modes of PCD often occur, and loss of control of single or mixed types of cell death leads to human diseases, such as cancer, hematologic disorders, autoimmune deficiency disorders, neurodegeneration, and infectious diseases. Moreover, PCD not only plays a regulatory role in maintaining the homeostasis of the body but also may play a regulatory role in unnecessary cell death. An increasing number of researchers have studied the cell death pathway as a drug target. Therefore, it is very important to clarify the mechanism and biomarkers of cell death for research on clinical drugs, especially antitumour drugs (Table 2).



**Table 2.** Biomarkers of various PCDs and their inducers and inhibitors.

PCD pathway	Biomarkers	Inducers	Inhibitors (target)
Apoptosis	Casp3,8,9 BCL-XL BCL-2 Fas	TNF + Smac-mimetic TNF + ActD TNF + CHX Trail + Smac-mimetic Fas Ligand + Smac-mimetic	Z-VAD-fmk (Pan Casp) Z-IETD-fmk (Casp8) Z-DEVD-fmk (Casp3)
Necroptosis	RIPK1 RIPK3 MLKL	TNF + Smac-mimetic + Z-VAD-fmk Trail + ActD + Z-VAD-fmk Fas Ligand + CHX + Z-VAD-fmk	Necrostatin-1 (RIPK1) Necrostatin-1s (RIPK1) GSK872 (RIPK3) Necrosulfonamide (MLKL)
Pyroptosis	Casp1,11 GSDMD IL-1 $\beta$ ,18 NLRP3	LPS + Nigericin	Ac-YVAD-cmk (Casp1) Wedelolactone (Casp11) Ac-FLTD-CMK (GSDMD) MCC950(NLRP3) Disulfiram(ALDH1)
Ferroptosis	Fe <sup>2+</sup> GSH GPX4 ROS	Erastin Glutamate Buthionine Sulfoximine (1 S,3 R)-RSL3 FIN56	Deferoxamine (Fe <sup>2+</sup> ) Deferiprone (Fe <sup>2+</sup> ) Cyclipirox (Fe <sup>2+</sup> ) Ferrostatin-1(ROS) Selenium (GPX4)
Autophagy-dependent cell death	Atg ClassIII PI3K Beclin1 LC3	Tat-Beclin1 BredeldinA C2-ceramide Rapamycin	Chloroquine (Lysosome) Concanamycin A (H <sup>+</sup> -ATPase) 3-methyladenine (ClassIII PI3K)
Mitotic Catastrophe	Casp3 P53 CyclinB1/CDK1	UV Cisplatin Microtubule targeting agents Adriamycin Radiation	NA (Not Applicable)
Immunogenic cell death	CRT ATP HMGB1 1-IFN ANXA1	Viral infection Anthracyclines Bortezomib Adriamycin Hypericin-based photodynamic therapy Radiation	ENTPD1 (ATP) NT5E (ATP)
Entosis	RhoA ROCK1,2 CDH1	CDH1	C3-toxin (RhoA) Y-27632 (ROCK) Blebbistatin (Myosin)
Parthanatos	AIFM1 MIF PARP-1	UV IR Alkylating agents Oxidative stress	BYK204165 (PARP-1) AG-14361 (PARP-1) Iniparib (PARP-1)
Cuproptosis	Cu <sup>2+</sup> FDX1 DLAT LIAS	Elesclomol + Cu <sup>2+</sup> NSC-319726 + Cu <sup>2+</sup> Disulfiram + Cu <sup>2+</sup>	Tetrathiomolybdate (Cu <sup>2+</sup> )
Disulfidptosis	SLC7A11 NADPH Cystine Disulfide	KL-11743 Bay-876	Dithiothreitol (Disulfide) $\beta$ -mercaptoethanol (Disulfide) TCEP (Disulfide)
NETosis	NETs NADPH MPO ROS	Phorbol myristate acetate IL-8 LPS	Tetrahydroisoquinolines (NETs) Cl-amidine (PAD4)
Lysosome-dependent cell death	LMP Cathepsins STAT3 ROS Fe <sup>2+</sup>	Lysosomotropic detergents Dipeptide methyl esters Sphingosine Phosphatidic acid	Leupeptin hemisulfate(Cathepsins) Deferoxamine (Fe <sup>2+</sup> ) NAC (ROS) CA-074Me (Cathepsin B)
Alkaliptosis	IKBKB CA9	JTC801	IMD0354 (IKBKB) CAY10657 ((IKBKB) N-acetyl alanine acid (pH)
Oxeiptosis	ROS PGAM5 KEAP1 AIFM1	Ozone H <sub>2</sub> O <sub>2</sub>	NAC (ROS)

There is a lack of a more rational classification of the cell death pathways that have been identified thus far. In addition, the specific mechanism of some cell death signaling pathways is not very clear, and the relationship between different types of cell death is poorly understood. For some pathways, there are still great controversies in the academic community, and further extensive and in-depth research is needed. The regulatory mechanisms of PCD on living organisms from micro- to macrosystems are the keys to the continuation of life, and a deep understanding of these mechanisms is extremely important for human development.

## REFERENCES

- Schweichel JU, Merker HJ. The morphology of various types of cell death in preliminary issues. *Teratology*. 1973;7:253–66.
- Galluzzi L, Maiuri MC, Vitale I, Zischka H, Castedo M, Zitvogel L, et al. Cell death modalities: classification and pathophysiological implications. *Cell Death Differ*. 2007;14:1237–43.
- Tang D, Kang R, Berghe TV, Vandenabeele P, Kroemer G. The molecular machinery of regulated cell death. *Cell Res*. 2019;29:347–64.
- Clarke PG. Developmental cell death: morphological diversity and multiple mechanisms. *Anat Embryol (Berl)*. 1990;181:195–213.
- Kerr JF, Wyllie AH, Currie AR. Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. *Br J Cancer*. 1972;26:239–57.
- Maiuri MC, Zalckvar E, Kimchi A, Kroemer G. Self-eating and self-killing: crosstalk between autophagy and apoptosis. *Nat Rev Mol Cell Biol*. 2007;8:741–52.
- Brumatti G, Salamanidis M, Ekert PG. Crossing paths: interactions between the cell death machinery and growth factor survival signals. *Cell Mol Life Sci*. 2010;67:1619–30.
- Tuckermann JP, Kleiman A, McPherson KG, Reichardt HM. Molecular mechanisms of glucocorticoids in the control of inflammation and lymphocyte apoptosis. *Crit Rev Clin Lab Sci*. 2005;42:71–104.
- Wilson NS, Dixit V, Ashkenazi A. Death receptor signal transducers: nodes of coordination in immune signaling networks. *Nat Immunol*. 2009;10:348–55.
- Wang L, Yang JK, Kabaleeswaran V, Rice AJ, Cruz AC, Park AY, et al. The Fas-FADD death domain complex structure reveals the basis of DISC assembly and disease mutations. *Nat Struct Mol Biol*. 2010;17:1324–9.
- Hengartner MO. The biochemistry of apoptosis. *Nature*. 2000;407:770–6.
- Gajate C, Mollinedo F. Edelfosine and perifosine induce selective apoptosis in multiple myeloma by recruitment of death receptors and downstream signaling molecules into lipid rafts. *Blood*. 2007;109:711–9.
- Datta R, Kojima H, Yoshida K, Kufe D. Caspase-3-mediated cleavage of protein kinase C theta in induction of apoptosis. *J Biol Chem*. 1997;272:20317–20.
- Earnshaw WC, Martins LM, Kaufmann SH. Mammalian caspases: structure, activation, substrates, and functions during apoptosis. *Annu Rev Biochem*. 1999;68:383–424.
- Degterev A, Huang Z, Boyce M, Li Y, Jagtap P, Mizushima N, et al. Chemical inhibitor of nonapoptotic cell death with therapeutic potential for ischemic brain injury. *Nat Chem Biol* 2005;1:112–9. [published correction appears in *Nat Chem Biol*. 2005 Sep;1(4):234].
- Kaiser WJ, Upton JW, Long AB, Livingston-Rosanoff D, Daley-Bauer LP, Hakem R, et al. RIP3 mediates the embryonic lethality of caspase-8-deficient mice. *Nature*. 2011;471:368–72.
- Mocarski ES, Upton JW, Kaiser WJ. Viral infection and the evolution of caspase 8-regulated apoptotic and necrotic death pathways. *Nat Rev Immunol*. 2011;12:79–88.
- Vanlangenakker N, Vanden Berghe T, Bogaert P, Laukens B, Zobel K, Deshayes K, et al. cIAP1 and TAK1 protect cells from TNF-induced necrosis by preventing RIP1/RIP3-dependent reactive oxygen species production. *Cell Death Differ*. 2011;18:656–65.
- Hsu H, Huang J, Shu HB, Baichwal V, Goeddel DV. TNF-dependent recruitment of the protein kinase RIP to the TNF receptor-1 signaling complex. *Immunity*. 1996;4:387–96.
- Bertrand MJ, Milutinovic S, Dickson KM, Ho WC, Boudreaux A, Durkin J, et al. cIAP1 and cIAP2 facilitate cancer cell survival by functioning as E3 ligases that promote RIP1 ubiquitination. *Mol Cell*. 2008;30:689–700.
- Dondelinger Y, Declercq W, Montessuit S, Roelandt R, Goncalves A, Bruggeman I, et al. MLKL compromises plasma membrane integrity by binding to phosphatidylinositol phosphates. *Cell Rep*. 2014;7:971–81.
- Su L, Quade B, Wang H, Sun L, Wang X, Rizo J. A plug release mechanism for membrane permeation by MLKL. *Structure*. 2014;22:1489–1500.
- Weinlich R, Oberst A, Beere HM, Green DR. Necroptosis in development, inflammation and disease. *Nat Rev Mol Cell Biol*. 2017;18:127–36.
- Brennan MA, Cookson BT. Salmonella induces macrophage death by caspase-1-dependent necrosis. *Mol Microbiol*. 2000;38:31–40.
- Galluzzi L, Vitale I, Aaronson SA, Abrams JM, Adam D, Agostinis P, et al. Molecular mechanisms of cell death: recommendations of the nomenclature committee on cell death 2018. *Cell Death Differ*. 2018;25:486–541.
- Feria MG, Taborda NA, Hernandez JC, Rugeles MT. HIV replication is associated to inflammasomes activation, IL-1 $\beta$ , IL-18 and caspase-1 expression in GALT and peripheral blood. *PLoS One*. 2018;13:e0192845.
- Kuang S, Zheng J, Yang H, Li S, Duan S, Shen Y, et al. Structure insight of GSDMD reveals the basis of GSDMD autoinhibition in cell pyroptosis. *Proc Natl Acad Sci USA*. 2017;114:10642–7.
- Lagrange B, Benaoudia S, Wallet P, Magnotti F, Provost A, Michal F, et al. Human caspase-4 detects tetra-acylated LPS and cytosolic Francisella and functions differently from murine caspase-11. *Nat Commun*. 2018;9:242.
- Yang D, He Y, Muñoz-Planillo R, Liu Q, Núñez G. Caspase-11 requires the pannexin-1 channel and the purinergic P2X7 pore to mediate pyroptosis and endotoxic shock. *Immunity*. 2015;43:923–32.
- Schmid-Burgk JL, Gaidt MM, Schmidt T, Ebert TS, Bartok E, Hornung V. Caspase-4 mediates non-canonical activation of the NLRP3 inflammasome in human myeloid cells. *Eur J Immunol* 2015;45:2911–7. [published correction appears in *Eur J Immunol*. 2021 Jun;51(6):1546].
- Dixon SJ, Lemberg KM, Lamprecht MR, Skouta R, Zaitsev EM, Gleason CE, et al. Ferroptosis: an iron-dependent form of nonapoptotic cell death. *Cell*. 2012;149:1060–72.
- Dixon SJ. Ferroptosis: bug or feature? *Immunol Rev*. 2017;277:150–7.
- Yang WS, Stockwell BR. Ferroptosis: death by lipid peroxidation. *Trends Cell Biol*. 2016;26:165–76.
- Xie Y, Hou W, Song X, Yu Y, Huang J, Sun X, et al. Ferroptosis: process and function. *Cell Death Differ*. 2016;23:369–79.
- Stockwell BR, Friedmann Angeli JP, Bayir H, Bush AI, Conrad M, Dixon SJ, et al. Ferroptosis: a regulated cell death nexus linking metabolism, redox biology, and disease. *Cell*. 2017;171:273–85.
- Li J, Cao F, Yin HL, Huang ZJ, Lin ZT, Mao N, et al. Ferroptosis: past, present and future. *Cell Death Dis*. 2020;11:88.
- Cao JY, Dixon SJ. Mechanisms of ferroptosis. *Cell Mol Life Sci*. 2016;73:2195–209.
- Ayala A, Muñoz MF, Argüelles S. Lipid peroxidation: production, metabolism, and signaling mechanisms of malondialdehyde and 4-hydroxy-2-nonenal. *Oxid Med Cell Longev*. 2014;2014:360438.
- Yang WS, Kim KJ, Gaschler MM, Patel M, Shchepinov MS, Stockwell BR. Peroxidation of polyunsaturated fatty acids by lipoxygenases drives ferroptosis. *Proc Natl Acad Sci USA*. 2016;113:E4966–E4975.
- Yu X, Long YC. Crosstalk between cystine and glutathione is critical for the regulation of amino acid signaling pathways and ferroptosis. *Sci Rep*. 2016;6:30033.
- Lin R, Zhang Z, Chen L, Zhou Y, Zou P, Feng C, et al. Dihydroartemisinin (DHA) induces ferroptosis and causes cell cycle arrest in head and neck carcinoma cells. *Cancer Lett*. 2016;381:165–75.
- Shen HM, Codogno P. Autophagic cell death: Loch Ness monster or endangered species? *Autophagy*. 2011;7:457–65.
- Liu Y, Shoji-Kawata S, Sumpster RM Jr, Wei Y, Ginet V, Zhang L, et al. Autosis is a Na<sup>+</sup>,K<sup>+</sup>-ATPase-regulated form of cell death triggered by autophagy-inducing peptides, starvation, and hypoxia-ischemia. *Proc Natl Acad Sci USA*. 2013;110:20364–71.
- Zhang H, Baehrecke EH. Eaten alive: novel insights into autophagy from multicellular model systems. *Trends Cell Biol*. 2015;25:376–87.
- Yang Z, Klionsky DJ. An overview of the molecular mechanism of autophagy. *Curr Top Microbiol Immunol*. 2009;335:1–32.
- Pattingre S, Espert L, Biard-Piechaczyk M, Codogno P. Regulation of macroautophagy by mTOR and Beclin 1 complexes. *Biochimie*. 2008;90:313–23.
- Li P, Shi J, He Q, Hu Q, Wang YY, Zhang LJ, et al. Streptococcus pneumoniae induces autophagy through the inhibition of the PI3K-I/Akt/mTOR pathway and ROS hypergeneration in A549 cells. *PLoS One*. 2015;10:e0122753.
- Maejima Y, Isobe M, Sadoshima J. Regulation of autophagy by Beclin 1 in the heart. *J Mol Cell Cardiol*. 2016;95:19–25.
- Margottin-Goguet F, Hsu JY, Loktev A, Hsieh HM, Reimann JD, Jackson PK. Prophase destruction of Emi1 by the SCF (betaTrCP/Slimb) ubiquitin ligase activates the anaphase promoting complex to allow progression beyond prometaphase. *Dev Cell*. 2003;4:813–26.
- Russell P, Nurse P. cdc25+ functions as an inducer in the mitotic control of fission yeast. *Cell*. 1986;45:145–53.
- Molz L, Booher R, Young P, Beach D. cdc2 and the regulation of mitosis: six interacting mcs genes. *Genetics*. 1989;122:773–82.
- Imreh G, Norberg HV, Imreh S, Zhivotovskiy B. Chromosomal breaks during mitotic catastrophe trigger  $\gamma$ H2AX-ATM-p53-mediated apoptosis. *J Cell Sci*. 2016;129:1950.
- Mc Gee MM. Targeting the mitotic catastrophe signaling pathway in cancer. *Mediators Inflamm*. 2015;2015:146282.

54. Chen J, Liu J. Erroneous silencing of the mitotic checkpoint by aberrant spindle pole-kinetochore coordination. *Biophys J*. 2015;109:2418–35.
55. Lara-Gonzalez P, Westhorpe FG, Taylor SS. The spindle assembly checkpoint. *Curr Biol*. 2012;22:R966–R980.
56. Vitale I, Galluzzi L, Castedo M, Kroemer G. Mitotic catastrophe: a mechanism for avoiding genomic instability. *Nat Rev Mol Cell Biol*. 2011;12:385–92.
57. Castedo M, Perfettini JL, Roumier T, Andreau K, Medema R, Kroemer G. Cell death by mitotic catastrophe: a molecular definition. *Oncogene*. 2004;23:2825–37.
58. Mansilla S, Priebe W, Portugal J. Mitotic catastrophe results in cell death by caspase-dependent and caspase-independent mechanisms. *Cell Cycle*. 2006;5:53–60.
59. Eom YW, Kim MA, Park SS, Goo MJ, Kwon HJ, Sohn S, et al. Two distinct modes of cell death induced by doxorubicin: apoptosis and cell death through mitotic catastrophe accompanied by senescence-like phenotype. *Oncogene*. 2005;24:4765–77.
60. Casares N, Pequinot MO, Tesniere A, Ghiringhelli F, Roux S, Chaput N, et al. Caspase-dependent immunogenicity of doxorubicin-induced tumor cell death. *J Exp Med*. 2005;202:1691–701.
61. Galluzzi L, Buqué A, Kepp O, Zitvogel L, Kroemer G. Immunogenic cell death in cancer and infectious disease. *Nat Rev Immunol*. 2017;17:97–111.
62. Vanpouille-Box C, Alard A, Aryankalayil MJ, Sarfraz Y, Diamond JM, Schneider RJ, et al. DNA exonuclease Trex1 regulates radiotherapy-induced tumour immunogenicity. *Nat Commun*. 2017;8:15618.
63. Galluzzi L, Bravo-San Pedro JM, Demaria S, Formenti SC, Kroemer G. Activating autophagy to potentiate immunogenic chemotherapy and radiation therapy. *Nat Rev Clin Oncol*. 2017;14:247–58.
64. Garg AD, Vandenberk L, Koks C, Verschuere T, Boon L, Van Gool SW, et al. Dendritic cell vaccines based on immunogenic cell death elicit danger signals and T cell-driven rejection of high-grade glioma. *Sci Transl Med*. 2016;8:328–327.
65. Galluzzi L, Zitvogel L, Kroemer G. Immunological mechanisms underneath the efficacy of cancer therapy. *Cancer Immunol Res*. 2016;4:895–902.
66. Garg AD, Vandenberk L, Fang S, Fasche T, Van Eygen S, Maes J, et al. Pathogen response-like recruitment and activation of neutrophils by sterile immunogenic dying cells drives neutrophil-mediated residual cell killing. *Cell Death Differ*. 2017;24:832–43.
67. Obeid M, Tesniere A, Ghiringhelli F, Fimia GM, Apetoh L, Perfettini JL, et al. Calreticulin exposure dictates the immunogenicity of cancer cell death. *Nat Med*. 2007;13:54–61.
68. Elliott MR, Chekeni FB, Trampont PC, Lazarowski ER, Kadl A, Walk SF, et al. Nucleotides released by apoptotic cells act as a find-me signal to promote phagocytic clearance. *Nature*. 2009;461:282–6.
69. Scaffidi P, Misteli T, Bianchi ME. Release of chromatin protein HMGB1 by necrotic cells triggers inflammation. *Nature*. 2002;418:191–5. [published correction appears in *Nature*. 2010 Sep 30;467(7315):622].
70. Hunger A, Medrano RF, Zanatta DB, Del Valle PR, Merkel CA, Salles TA, et al. Reestablishment of p53/Arf and interferon- $\beta$  pathways mediated by a novel adenoviral vector potentiates antiviral response and immunogenic cell death. *Cell Death Discov*. 2017;3:17017.
71. Chiba S, Baghdadi M, Akiba H, Yoshiyama H, Kinoshita I, Dosaka-Akita H, et al. Tumor-infiltrating DCs suppress nucleic acid-mediated innate immune responses through interactions between the receptor TIM-3 and the alarmin HMGB1. *Nat Immunol*. 2012;13:832–42.
72. Vacchelli E, Ma Y, Baracco EE, Sistigu A, Enot DP, Pietrocola F, et al. Chemotherapy-induced antitumor immunity requires formyl peptide receptor 1. *Science*. 2015;350:972–8.
73. Overholtzer M, Mailleux AA, Mouneimne G, Normand G, Schnitt SJ, King RW, et al. A nonapoptotic cell death process, entosis, that occurs by cell-in-cell invasion. *Cell*. 2007;131:966–79.
74. Krishna S, Overholtzer M. Mechanisms and consequences of entosis. *Cell Mol Life Sci*. 2016;73:2379–86.
75. Wang M, Ning X, Chen A, Huang H, Ni C, Zhou C, et al. Impaired formation of homotypic cell-in-cell structures in human tumor cells lacking alpha-catenin expression. *Sci Rep*. 2015;5:12223.
76. Wang C, Chen A, Ruan B, Niu Z, Su Y, Qin H, et al. PCDH7 inhibits the formation of homotypic cell-in-cell structure. *Front Cell Dev Biol*. 2020;8:329.
77. Sun Q, Luo T, Ren Y, Florey O, Shirasawa S, Sasazuki T, et al. Competition between human cells by entosis. *Cell Res*. 2014;24:1299–310.
78. Mackay HL, Moore D, Hall C, Birkbak NJ, Jamal-Hanjani M, Karim SA, et al. Genomic instability in mutant p53 cancer cells upon entotic engulfment. *Nat Commun*. 2018;9:3070. [published correction appears in *Nat Commun*. 2018 Aug 28;9(1):3540].
79. Liang J, Niu Z, Zhang B, Yu X, Zheng Y, Wang C, et al. p53-dependent elimination of aneuploid mitotic offspring by entosis. *Cell Death Differ*. 2021;28:799–813.
80. David KK, Andrabi SA, Dawson TM, Dawson VL. Parthanatos, a messenger of death. *Front Biosci (Landmark Ed)*. 2009;3:1116–28.
81. Yu SW, Wang H, Poitras MF, Coombs C, Bowers WJ, Federoff HJ, et al. Mediation of poly(ADP-ribose) polymerase-1-dependent cell death by apoptosis-inducing factor. *Science*. 2002;297:259–63.
82. Wang Y, Kim NS, Haince JF, et al. Poly(ADP-ribose) (PAR) binding to apoptosis-inducing factor is critical for PAR polymerase-1-dependent cell death (parthanatos). *Sci Signal*. 2011;4:ra20.
83. Polster BM, Basañez G, Etkebarria A, Hardwick JM, Nicholls DG. Calpain I induces cleavage and release of apoptosis-inducing factor from isolated mitochondria. *J Biol Chem*. 2005;280:6447–54.
84. Tsvetkov P, Coy S, Petrova B, Dreishpoon M, Verma A, Abdusamad M, et al. Copper induces cell death by targeting lipoylated TCA cycle proteins. *Science*. 2022;375:1254–61. [published correction appears in *Science*. 2022 Apr 22;376(6591): eabq4855].
85. Liu X, Nie L, Zhang Y, Yan Y, Wang C, Colic M, et al. Actin cytoskeleton vulnerability to disulfide stress mediates disulfidoptosis. *Nat Cell Biol*. 2023;25:404–14.
86. Liu X, Olszewski K, Zhang Y, Lim EW, Shi J, Zhang X, et al. Cystine transporter regulation of pentose phosphate pathway dependency and disulfide stress exposes a targetable metabolic vulnerability in cancer. *Nat Cell Biol*. 2020;22:476–86.
87. Brinkmann V, Reichard U, Goosmann C, Fauler B, Uhlemann Y, Weiss DS, et al. Neutrophil extracellular traps kill bacteria. *Science*. 2004;303:1532–5.
88. Fuchs TA, Abed U, Goosmann C, Hurwitz R, Schulze I, Wahn V, et al. Novel cell death program leads to neutrophil extracellular traps. *J Cell Biol*. 2007;176:231–41.
89. Franko J, Pomfy M, Prosbová T. Apoptosis and cell death (mechanisms, pharmacology and promise for the future). *Acta Med (Hradec Kralove)*. 2000;43:63–8.
90. Aits S, Jäättelä M. Lysosomal cell death at a glance. *J Cell Sci*. 2013;126:1905–12.
91. Gómez-Sintes R, Ledesma MD, Boya P. Lysosomal cell death mechanisms in aging. *Ageing Res Rev*. 2016;32:150–68.
92. Song X, Zhu S, Xie Y, Liu J, Sun L, Zeng D, et al. JTC801 induces pH-dependent death specifically in cancer cells and slows growth of tumors in mice. *Gastroenterology*. 2018;154:1480–93.
93. Liu J, Kuang F, Kang R, Tang D. Alkaliptosis: a new weapon for cancer therapy. *Cancer Gene Ther*. 2020;27:267–9.
94. Holze C, Michaudel C, Mackowiak C, Haas DA, Benda C, Hubel P, et al. Oxeiptosis, a ROS-induced caspase-independent apoptosis-like cell-death pathway. *Nat Immunol*. 2018;19:130–40.
95. Chen X, Kang R, Kroemer G, Tang D. Broadening horizons: the role of ferroptosis in cancer. *Nat Rev Clin Oncol*. 2021;18:280–96.

## AUTHOR CONTRIBUTIONS

YC, XL, and S-BL drafted the original manuscript. MY designed and created the figures. S-BL and MY structured and revised the manuscript. All authors have read and agreed to the final version of the manuscript.

## FUNDING

This work was supported by the Jiangsu Higher Education Institution Innovative Research Team for Science and Technology (2021), Program of Jiangsu Vocational College Engineering Technology Research Center (2023), Key Technology Program of Suzhou People's Livelihood Technology Projects (Grant No. SKY2021029), Open Project of Jiangsu Biobank of Clinical Resources (TC2021B009), Project of State Key Laboratory of Radiation Medicine and Protection, Soochow University (No. GZK12023013), Programs of the Suzhou Vocational Health College (Grant no. szwzy 202210, SZWZYTD202201), and Qing-Lan Project of Jiangsu Province in China (2021, 2022).

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Correspondence** and requests for materials should be addressed to Minfeng Yang or Song-Bai Liu.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024



## Review

## Pioneer factors: roles and their regulation in development

Amandine Barral<sup>1</sup> and Kenneth S. Zaret<sup>1,\*</sup> 

**Pioneer factors are a subclass of transcription factors that can bind and initiate opening of silent chromatin regions. Pioneer factors subsequently regulate lineage-specific genes and enhancers and, thus, activate the zygotic genome after fertilization, guide cell fate transitions during development, and promote various forms of human cancers. As such, pioneer factors are useful in directed cell reprogramming. In this review, we define the structural and functional characteristics of pioneer factors, how they bind and initiate opening of closed chromatin regions, and the consequences for chromatin dynamics and gene expression during cell differentiation. We also discuss emerging mechanisms that modulate pioneer factors during development.**

**Introduction: pioneer factor definition and recent discoveries**

Most DNA in the nucleus is wrapped into **nucleosomes** (see [Glossary](#)) and assembled into chromatin fibers, making DNA inaccessible to many **transcription factors** and other gene regulatory proteins ([Figure 1A](#)). By virtue of their ability to bind DNA on a nucleosome, pioneer transcription factors can target genes and enhancers in compacted chromatin. Active genes and enhancers typically harbor an open chromatin conformation and exist in **euchromatin**. Active promoters and enhancers can include nucleosome-depleted regions, where DNA is accessible to transcription factors and protein machineries [1]. They are often flanked by nucleosomes harboring the histone modifications H3K27Ac, H3K4me3, and/or H3K4me1 ([Figure 1B](#)). However, euchromatin regions represent a small fraction of the nuclear genome; the largest fraction being nucleosome dense and folded into compacted and silent chromatin. We now appreciate that silent chromatin can be in distinct molecular forms [2,3], including **naive chromatin**, which is enriched for linker histone H1 isoforms that compact chromatin [2], H3K9me3-marked heterochromatin, which is bound by the chromatin compacting protein Heterochromatin protein 1 (HP1), and H3K27me3-marked heterochromatin, which is bound by the chromatin-compacting Polycomb repressive complex 1 (PRC1) and Polycomb repressive complex 2 (PRC2) ([Figure 1B](#)). These three silent chromatin states (naive, H3K9me3-marked, and H3K27me3-marked) repress lineage-specific genes, enhancers, DNA regulatory elements, and repeat elements [3–5]. During development, new spatiotemporal, lineage-specific gene expression networks are required to induce cell differentiation, necessitating a transition from silent chromatin states to an open chromatin conformation at promoters and enhancers. The transition is often initiated by **pioneer factors**, which by binding nucleosomal DNA in silent chromatin, are among the master regulators of cell fate changes [6–12] ([Figure 1A](#)). Herein we primarily focus on the latest studies in the field.

Pioneer factors interact with silent chromatin regions and enable their remodeling into an active chromatin conformation ([Figure 1A](#)) or further compaction and silencing, as reviewed recently [13–15]. The opening of chromatin by pioneer factors can lead to activation of lineage-specific genes and enhancers [6,16,17] and guide cell differentiation [7,9,10,16,18,19]. Indeed, the

**Highlights**

Pioneer factors can be master regulators of development. They are responsible for zygotic genome activation after fertilization and drive cell fate transitions during development.

Pioneer factors bind nucleosomal DNA containing their DNA-binding motifs. The binding is mediated by the DNA-binding domain of the pioneer factor and via establishment of histone–pioneer factor interactions.

Pioneer factors initiate local chromatin decompaction by partially unwrapping DNA from nucleosomes and reducing local nucleosome interactions.

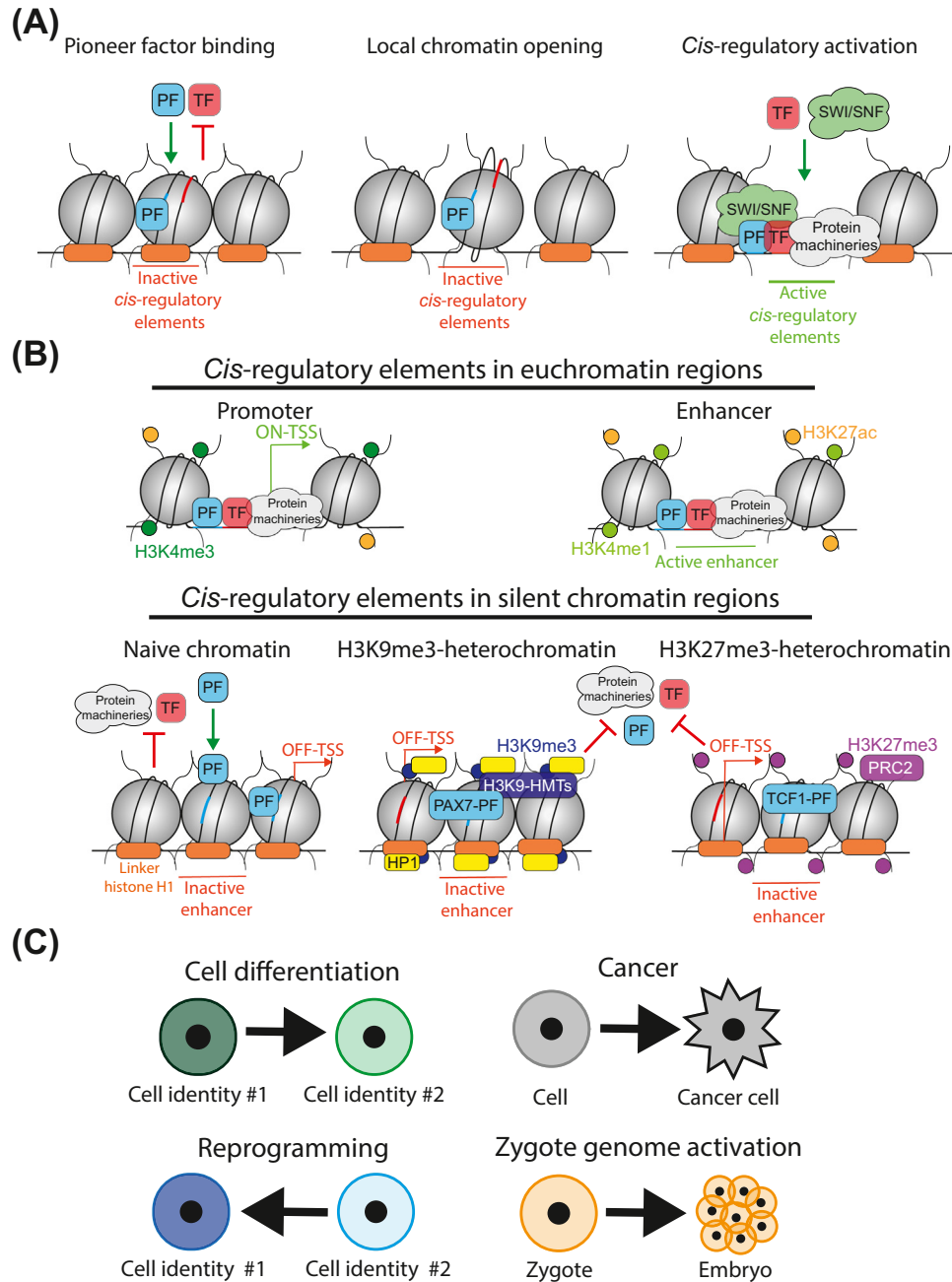
Pioneer factors recruit nucleosome-remodeling complexes and chromatin-modifying enzymes to promote silent chromatin reorganization, deposition of active chromatin modifications, and, thus, activation of underlying promoter and enhancer DNA sequences.

Pioneer factors can vary with regard to their ability to target diverse types of silent chromatin during cell differentiation.

<sup>1</sup>Institute for Regenerative Medicine and Department of Cell and Developmental Biology, Perelman School of Medicine, University of Pennsylvania, 3400 Civic Boulevard, Philadelphia, PA 19104, USA

\*Correspondence: [zaret@pennmedicine.upenn.edu](mailto:zaret@pennmedicine.upenn.edu) (K.S. Zaret).





**Figure 1. Pioneer factors (PFs) bind silent chromatin and drive cell differentiation.** (A) Left: PFs bind closed chromatin regions. Middle: PFs induce local DNA accessibility to other transcription factors (TFs) and protein complexes. Right: PFs induce chromatin reorganization and, thus, DNA sequence activation. Protein machineries include histone-modifying complexes, the RNA polymerase transcription complex, and additional protein complexes regulating *cis*-regulatory elements. (B) Chromatin state of *cis*-regulatory elements in euchromatin regions (top) and silent chromatin regions (bottom). Unlike canonical TFs, PFs are able to bind nucleosomal DNA in silent chromatin regions. (C) PFs guide cell differentiation; thus, their dysregulation might lead to cancer or cell reprogramming. They also drive zygotic gene activation. Abbreviations: H3K9-HMTs, Histone H3 Lysine 9-histone methyltransferases; PAX7, Paired box gene 7; PRC2, Polycomb repressive complex 2; SWI/SNF, Switch/sucrose nonfermenting; TCF1, T cell factor 1; TSS, transcription start site.

### Glossary

**Euchromatin:** active chromatin conformation of genome where genes are expressed and *cis*-regulatory elements are active. Active promoters and enhancers reside in euchromatin and typically have central regions that are depleted for nucleosomes, keeping the DNA accessible to transcription factors and gene regulatory protein complexes. They are also flanked by nucleosomes enriched for specific histone modifications, such as H3K27ac, H3K4me3, and H3K4me1.

**H3K27me3-heterochromatin:** silent chromatin enriched in H3K27me3 and H2AK119ub mediated by Polycomb repressive complex 2 (PRC2) and Polycomb repressive complex 1 (PRC1) complexes, respectively. H3K27me3-heterochromatin represses lineage-specific genes to ensure cell differentiation.

**H3K9me3-heterochromatin:** silent chromatin enriched in H3K9me3 and Heterochromatin protein 1 (HP1). HP1 induces H3K9me3-heterochromatin compaction. H3K9me3-heterochromatin is established by the histone methyltransferases (HMTs) Suppressor of variegation 3-9 homologs-1/2 (SUV39h-1/2) and SET domain bifurcated 1 (SETDB1) on DNA repeat sequences and lineage-specific genes and enhancers.

**Naive chromatin:** chromatin harboring a closed conformation and that is transcriptionally silent. It is enriched in linker histone H1 but does not harbor, or has low enrichment of, specific histone modifications.

**Nucleosome:** the nucleosome core particle is the basic repeating unit of chromatin. A nucleosome comprise ~150 base pairs of DNA sequence wrapped around an octamer of the four core histones: H2A, H2B, H3 and H4. Nucleosomes limit DNA sequence accessibility.

**Pioneer factor:** a subclass of transcription factor able to bind a sequence motif on free DNA and when its motif DNA is exposed on a nucleosome, although the latter may be limited by the position of the motif on the nucleosome; thus, pioneer factors are distinguished by their ability to target nucleosomal DNA. Recent studies revealed that pioneer factors interact with histones within the nucleosome core. Pioneer factors induce local nucleosomal accessibility in chromatin to

Trends in Genetics

ectopic expression of pioneer factors can initiate cell fate transitions [8,17,20–24], whereas pioneer factor-depleted cells often fail to differentiate [6,7,9,11,19,22,24–30]. Furthermore, pioneer factor dysregulation can cause aberrant cell phenotypes, such as cancer [31,32] or cell reprogramming [33–39], as recently reviewed elsewhere [40] (Figure 1C). Indeed, chromosome translocations can form oncogenic-fusion pioneer factors, such as Paired box gene 3-Forkhead box protein O1 (PAX3-FOXO1) in rhabdomyosarcoma cancer cells, leading to cancer [41] by altering enhancer repertory usage and gene expression profile [42–44].

The fundamental role of pioneer factors in development has been affirmed by their ability to drive **zygotic gene activation** [6,11,25,26,30] (Figure 1C). First characterized in *Drosophila*, the pioneer factor Zelda initiates zygotic gene transcription in the early embryo [45,46]. In mouse and human, Double homeobox (DUX) family factors, which can target closed chromatin [47], have been identified as being required for zygotic genome activation, since their zygotic depletion leads to defective genome zygotic activation and impairs embryonic development [48]. In zebrafish, the POU domain, class 5, transcription factor 3 (Pou5f3), and the SRY-box transcription factor 19b (Sox19b) act together as pioneer factors with homeobox protein NANOG (Nanog) to activate zygotic gene transcription. Triple zygotic mutants are deficient in zygotic gene expression and fail to induce embryonic development [26]. Upon fertilization, the zygotic genome harbors closed and inactive chromatin. Pioneer factors, such as Nuclear receptor subfamily 5 group A (Nr5a2) in murine zygotic cells [6], target closed chromatin and initiate its opening, which results in zygotic gene expression during early developmental stage [45,48].

While the distinction between canonical transcription factors and pioneer factors can be considered to be a continuum [49], pioneer factors are defined by specific properties that distinguish them from other transcription factors [50–52]. Whereas canonical transcription factors have weak affinity for closed chromatin regions and their binding on DNA target site is impeded by nucleosomes [51–53], pioneer factors present a higher affinity for closed chromatin [50,51,53–56] and weak dissociation from nucleosomal DNA [51,53,57]. Therefore, pioneer factors are defined by their ability to bind their target DNA site on nucleosomes *in vitro* [25,36,52,53,58–72] and *in vivo* [33,73–75]. Moreover, recent biophysical and structural studies highlight details by which pioneer factors can differentially interact with nucleosome core particles [51,59,63,65,67,72,76], whereas canonical transcription factors do it poorly or undetectably [51]. Such structural studies are relatively new; thus, we can anticipate that additional mechanisms of nucleosome interactions will be revealed as more structural studies are performed [77]. Pioneer factors are also characterized by their ability, upon binding *in vitro*, to perturb the structure of an underlying nucleosome in an ATPase-independent manner [51,58,59,63,64,67,71,73], although full opening of chromatin sites targeted by the factors *in vivo* involves ATP-dependent remodelers [32,34,78,79]. In this review, we describe the latest studies of characteristics of pioneer factors and how they regulate cell fate, and then present mechanisms of regulation of pioneer factors during development.

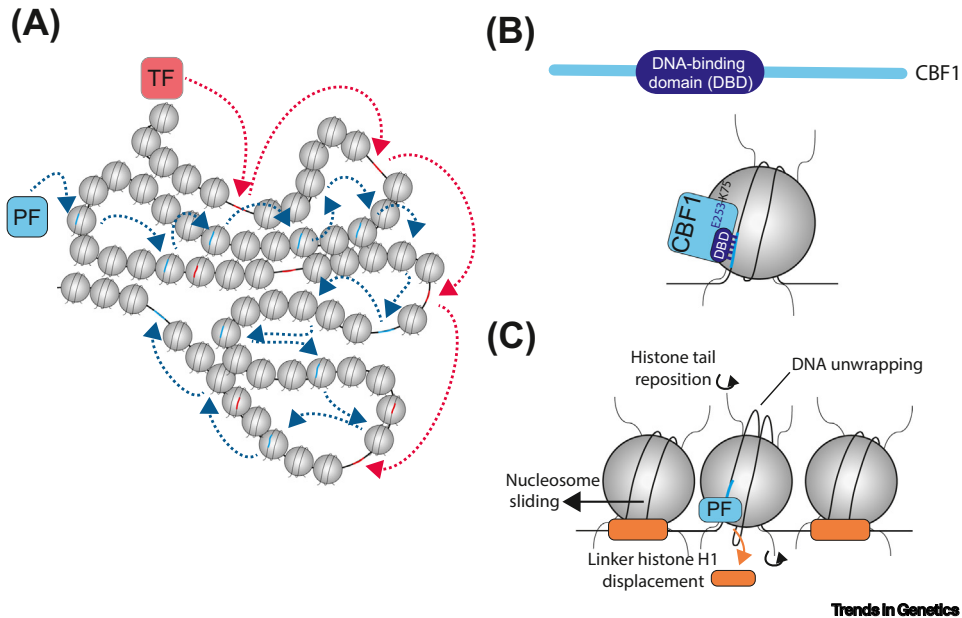
### Pioneer factors scan and bind locally closed chromatin

Investigations into the underlying mechanisms of action reveal that the act of ‘pioneering’ occurs through multiple steps: chromatin scanning, nucleosomal DNA targeting, and local chromatin reorganization. First, pioneer factors need to identify their specific DNA-binding sites by scanning the chromatin fiber. While they can freely diffuse in the nucleus, they also exhibit many associations with nucleosomal chromatin to reach their DNA-binding site [80,81] (Figure 2A). Single molecule-tracking microscopy assays showed that pioneer factors have a short residence time on closed chromatin, while non-pioneers appear to avoid it altogether [50,54–56,80]. Nonspecific electrostatic interactions with the phosphodiester backbone, as it is exposed on the nucleosome surface, allow the pioneer factors to scan closed chromatin loci across the nucleus [50,55,74].

promote binding of other transcription factors or protein complexes that can either open the chromatin further or lead to its stronger compaction.

**Transcription factor:** protein able to bind a specific DNA motif on free DNA, through its DNA-binding domain to regulate gene expression.

**Zygotic gene activation:** first transcriptional expression of embryonic genes after egg fertilization.



**Figure 2. Mechanism of action of pioneer factors (PFs) to induce silent chromatin opening.** (A) Unlike canonical transcription factors (TFs), PFs can transiently interact with silent chromatin to identify their binding sites. (B) PFs interact with nucleosomal DNA containing its DNA motif through its DNA-binding domain (DBD) and protein domains outside the DBD. The non-DBD can establish contacts with histones in the nucleosome core particles, as described by Donovan *et al.* [51]. (C) PF binding induces local silent chromatin opening by inducing local DNA accessibility and local chromatin decompaction.

While their on-off behavior and local proximity to DNA mass would allow them to scan nucleosomal DNA via 1D diffusion [80], a recent biophysical study of the GAGA pioneer factor (GAF) indicated that nucleosome scanning is more of a 3D process [56]. Detailed single molecule-tracking studies comparing the Forkhead box A1 (FOXA1) and SRY-box transcription factor (SOX2) pioneer factors for chromatin scanning showed that FOXA1 diffuses slowly in closed chromatin domains, whereas SOX2 diffuses more rapidly in such domains; yet, in ectopic expression experiments, both factors similarly target closed chromatin target sites with low-turnover nucleosomes [50]. By contrast, comparably expressed Hepatocyte nuclear factor-4-alpha (HNF4A), a transcription factor that appears not to be a pioneer factor, scans closed chromatin poorly and, when it does so, targets sites with high-turnover nucleosomes [50]. Observing differences in chromatin scanning among the few pioneer factors assessed in this fashion suggests that there are a variety of mechanisms by which such factors scan chromatin.

While scanning chromatin, pioneer factors can transiently interact with target motifs on nucleosomal DNA [25,51] (Figure 2B). The DNA-binding domain (DBD) itself is sufficient to bind nucleosomes, as observed for FOXA1 [82] and SOX2 [58], but remains to be confirmed for additional pioneer factors. DBD deletion of Octamer-binding transcription factor 4 (OCT4) abolishes binding [69], confirming its fundamental role in nucleosomal DNA binding. DBDs of pioneer factors typically fold into an  $\alpha$ -helix or short helical twist structure, allowing nucleosome invasion and DNA anchoring [60]. Pioneer factors preferentially bind their target DNA motif when it localizes at nucleosome entry/exit site or at DNA groove exposed on the nucleosome surface [32,52,59,65,67,71,72,83].

Mutations of pioneer factors, outside the DBD, can selectively impact nucleosome binding [10,84–86] and the ability to drive cell fate [76]. A striking example is provided by the deletion of four amino acids within the linker region between two DBDs of the OCT4, which ablates nucleosome binding and the reprogramming function of OCT4, without affecting how the factor binds



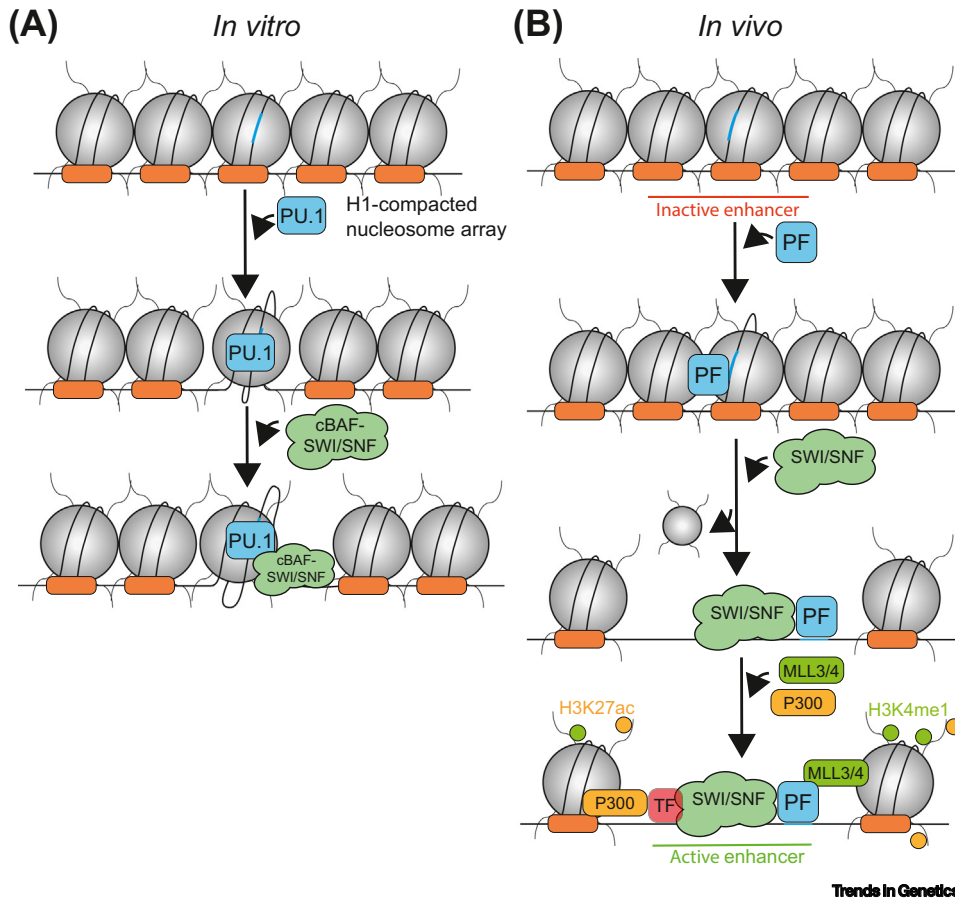
free DNA [69]. Recent cross-linking studies showed that non-DBDs can be in close proximity to the nucleosome surface and establish pioneer factor–nucleosome interactions [51,59,67,69,72,76]. Different pioneer factors use different forms of interaction with the nucleosome. A pioneer factor residue can interact with specific residues of the histone core or N-terminal extensions (tails) [59,76], such as for Cbf1, where its E253 residue interacts with K75 residue of H2A histone [51] (Figure 2B). Pioneer factors can also establish electrostatic interactions with the nucleosome surface [67]. OCT4 has a small acidic patch with which it can interact with positively charged lysine and arginine residues of histones [67,87]. The various emerging mechanisms illustrate how we are only beginning to understand how pioneer factors interact with nucleosomes.

### Pioneer factors can open locally closed chromatin

After binding to closed chromatin, pioneer factors can induce local chromatin decompaction. Recent chromatin immunoprecipitation (ChIP)-nexus experiments on *Drosophila* embryos demonstrated that the Zelda pioneer factor initiates and drives chromatin opening, while other transcription factors tested only increase chromatin accessibility that was already initiated by pioneer factors [88]. How does this initiate? Pioneer factor binding to nucleosomes often leads to DNA accessibility, making the nucleosomal DNA sensitive to nucleases [58,60,65,71,89] or accessible to other proteins [58,89–91]. Incubating FOXA1 [92], PU.1 [61], or OCT4 [71] with a linker histone (H1)-compacted nucleosome array induced DNaseI cleavage underlying its binding site, in the absence of an ATP-dependent nucleosome remodeler. Pioneer factors elicit the initial step of chromatin opening by inducing DNA distortions [51,74,87] and limited DNA unwrapping of the nucleosome [51,63,65,72,87] (Figure 2C). The binding of OCT4 on nucleosome unwraps 25 base pairs of nucleosomal DNA [71]. Pioneer factors alone can also induce nucleosome sliding [51,74,87] and histone tail reposition [58]. Upon its binding, OCT4 and H4/H2B histone tails are in close proximity, leading to conformation changes and relocalization of histone tails [67,71]. This could decrease internucleosome interactions and nucleosome stacking, mechanistically explaining local chromatin opening by pioneer factors [74] (Figure 2C). Finally, due to their ability to bind DNA at nucleosome entry/exit site, or near the nucleosome dyad axis, pioneer factors can displace and compete with histone linker H1 [59,71,89] (Figure 2C). Indeed, the Leafy (LFY) pioneer factor shows structural similarity with H1 [73]. We could assume that this competition mechanism could be shared by several pioneer factors, since many bind the nucleosome entry/exit site [6,32,51,59,65,67,69,71].

### Pioneer factors enable ATP-dependent chromatin reorganization and *cis*-regulatory element activation

After local chromatin decompaction, additional protein machineries are required to promote complete opening of closed chromatin regions. Foundational *in vitro* studies using chromatin assembled with *Drosophila* embryo extracts showed that nucleosome disruption by the pioneer factor GAF depends on ATP hydrolysis via the Nucleosome-remodeling factor (NURF) complex [93], indicating that complete chromatin remodeling is an energy-dependent pathway. In a recent *in vitro* study using fully purified, H1-compacted nucleosome arrays, purified PU.1 pioneer factor and the purified canonical BRG1/BRM-associated factor (cBAF) Switch/sucrose nonfermenting (SWI/SNF) nucleosome-remodeling complex showed that PU.1 alone could initiate DNA hypersensitivity underlying its binding site, which was enabled by an unstructured domain of PU.1. The initially open domain was expanded by the cBAF complex only in the presence of ATP [61] (Figure 3A). In addition, incubation of the cBAF complex with a PU.1 mutant deficient in its recruitment of cBAF, but which still harbors the PU.1 unstructured domain and a robust local chromatin opening activity, did not allow cBAF action [61]. These results argue against a model whereby the cBAF remodeler scans chromatin for a partially open site and argues for a model in which the



**Figure 3. Pioneer factors (PFs) drive closed chromatin reorganization and gene expression activation.** (A) *In vitro*, PU.1 is able to recruit the canonical BRG1/BRM-associated factor (cBAF) Switch/sucrose nonfermenting (SWI/SNF) nucleosome-remodeling complex on H1-compacted nucleosome arrays to expand local DNA accessibility, as described by Frederick *et al.* [61]. (B) *In vivo*, a PF leads to chromatin reorganization and, thus, enhancer activation. A PF recruits SWI/SNF nucleosome-remodeling complexes to promote nucleosome eviction and chromatin opening. PFs also promote deposition of active chromatin on enhancer by recruiting Mixed-lineage leukemia 3/4 (MLL3/4) and p300 enzymes. Abbreviation: TF, transcription factor.

pioneer factor must elicit local chromatin opening and simultaneously recruit or stabilize the remodeler at the site.

During cell differentiation, pioneer factors directly recruit SWI/SNF nucleosome-remodeling complexes by interacting with different subunits of SWI/SNF complexes [94,95], such as Brahma-related gene 1 (BRG1) [16,34,36,78,79,96] or BAF [9]. Thus, pioneer factors guide SWI/SNF-dependent remodeling complexes to specific silent loci, which are important for cell differentiation. After their recruitment on chromatin, SWI/SNF nucleosome-remodeling complexes evict nucleosomes [34,78] and promote chromatin rearrangement, leading to an open chromatin conformation [9,36,78,94,96,97] (Figure 3B). While pioneer factors recruit and stabilize SWI/SNF complexes on chromatin [34,36,61,78,94,98], the open domains created by SWI/SNF complexes can subsequently stabilize pioneer factors on chromatin [34,78,79]. Indeed, inhibition of BRG1 led to a reduction in PU.1 binding [79]. These recent results suggest a positive feedback loop between pioneer factors and SWI/SNF nucleosome-remodeling complexes that can maintain an open chromatin conformation on *cis*-regulatory elements.

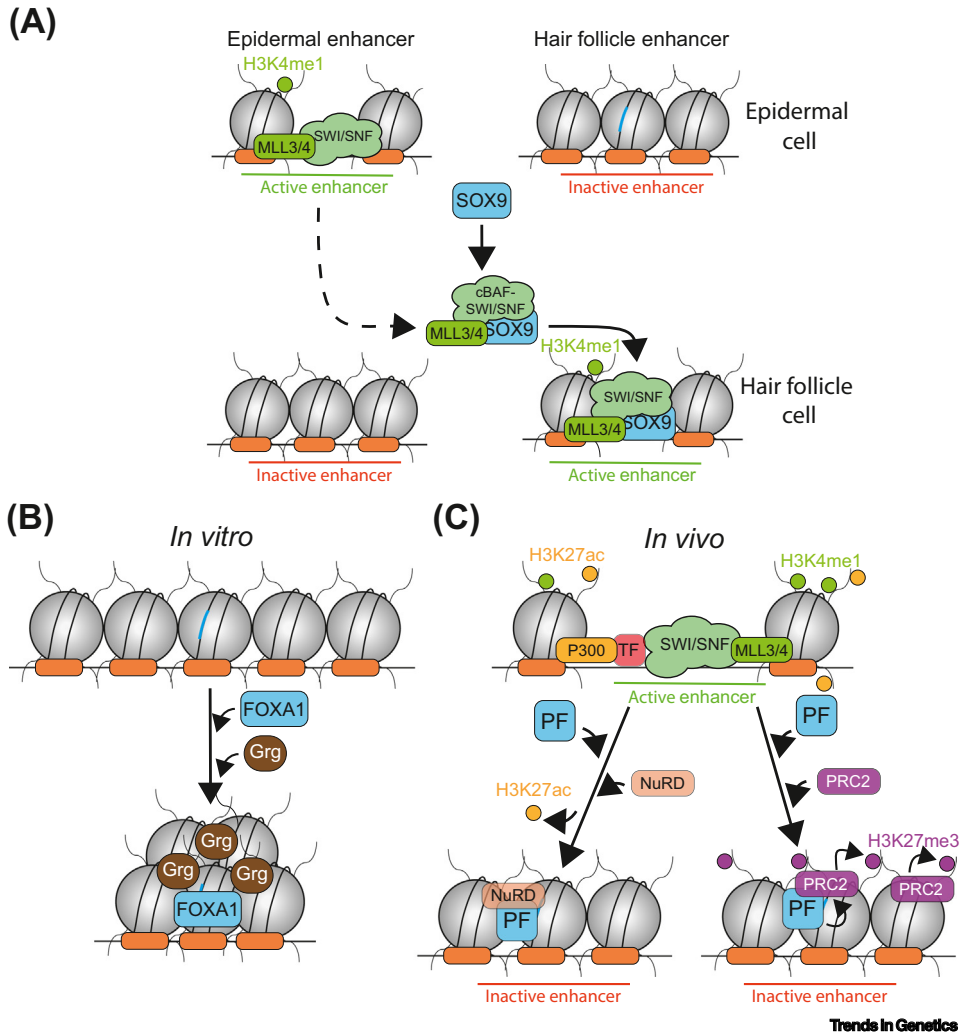
After initial chromatin opening, additional transcription factors and other protein complexes bind the newly accessible region [97,99]. Pioneer factor binding correlates with a reduction in cytosine methylation of 5-methylCpG residues [86,100–103]; CpG methylation (5mC) normally represses enhancers [104]. Indeed pioneer factors can recruit Ten-eleven translocation (TET) enzymes to enhancers [101,105], leading to conversion of 5mC into 5-hydroxymethylcytosine (5hmC) [105,106]. Pioneer factors also establish an active chromatin state on enhancers by recruiting Mixed-lineage leukemia-3/4 (MLL3/4) [107,108] and p300 [16,26,47,109] enzymes, which deposit H3K4me1 and H3K27ac on enhancers, respectively (Figure 3B). FOXA1 can recruit TET enzymes and reduce 5mC on enhancers [106,110], while also recruiting MLL3 and promoting H3K4me1 deposition [108]; both of which increase enhancer activity and gene expression. Finally, pioneer factors mediate enhancer–promoter proximity [33,94,111] and, thus, stimulate target lineage-specific gene expression. T cell factor 1 (TCF1) regulates CCCTC-binding factor (CTCF) on enhancers [112]. A recent mechanism based on protein aggregation has also been identified [113–115]. Indeed, OCT4 [114] and Nanog [115] can form protein aggregates to favor enhancer–promoter communications and gene activation.

### Pioneer factors can also result in further closing of chromatin and gene silencing

To direct cell differentiation, pioneer factors also elicit silencing of other lineage-specific genes [19,21,23,33,95,100]. They can indirectly inactivate gene expression by inducing passive enhancer inactivation. That is, pioneer factor binding on chromatin leads to transcription factor and protein machinery relocalization on the enhancer repertoire, away from active enhancers [17,116,117]. For example, induction of SRY-box transcription factor 9 (SOX9), a pioneer factor driving the hair follicle fate, in epidermal stem cells indirectly silences epidermal-related genes by interacting with SWI/SNF remodeling complexes and MLL4/3, and recruiting them away from active epidermal enhancers [95] (Figure 4A). Pioneer factors can also directly inactivate genes by establishing silent chromatin on enhancers and promoters, which can include chromatin closure. For example, when it recruits the Groucho homolog (Grg) transcriptional repressor to naive chromatin, FOXA1 can induce local chromatin compaction and impairs binding of other transcription factors [118] (Figure 4B). Pioneer factors can also drive chromatin closure by recruiting histone deacetylases, such as the Nucleosome-remodeling and deacetylase (NuRD) complex, which causes erasure of the H3K27ac active histone mark and loss of chromatin accessibility on enhancers [119,120] (Figure 4C). A recent study showed that the GAF pioneer factor can elicit H3K9me3 deposition on DNA repeats and transcriptional silencing during *Drosophila* zygotic genome activation [121]. Pioneer factors interact with diverse repressive complexes [33,120]. Some pioneer factors, such as PU.1, can recruit and stimulate the PRC2 complex, inducing H3K27me3-marked heterochromatin deposition on lineage-specific enhancers and promoters [23,33,100,119] (Figure 4C). To summarize, pioneering occurs when the transcription factor targets a naive domain of chromatin and, via partners, elicits a new functional capacity to the domain: either open and active or closed and further silenced.

### Pioneer factors maintain cell identity

During mitosis, chromatin is massively condensed into mitotic chromosomes, many DNA-binding proteins are evicted from chromosomes, and transcription is diminished, but not completely silent [122]. After mitosis exit, cells reactivate lineage-specific gene expression to restore their identity. Live imaging, using fluorescent protein-fused pioneer factors, showed that at least a subset of pioneer factors, such as GATA binding protein 2 (GATA2), FOXA1, PAX3, and SOX2, stay associated with mitotic chromosomes [123–125]. Their mitotic chromosome-binding ability is related to their affinity for closed chromatin, since it depends on their DBDs and their ability to establish nonspecific electrostatic interactions with closed chromatin [123–127]. SOX2 degradation



**Figure 4. Pioneer factors (PFs) induce gene silencing.** (A) PFs induce passive enhancer inactivation by relocating protein complexes on chromatin. As described by Yang *et al.* [95], SRY-box transcription factor 9 (SOX9) recruits Mixed-lineage leukemia 3/4 (MLL3/4) and Switch/sucrose nonfermenting (SWI/SNF) nucleosome-remodeling complex away from active epidermal enhancers during a hair follicle cell fate transition. (B) *In vitro*, Forkhead box protein A1 (FOXA1) increases local compaction of closed chromatin by recruiting Groucho homolog (Grg) transcriptional corepressor, as described by Sekiya *et al.* [118]. (C) *In vivo*, PFs directly induce enhancer inactivation. They erase active chromatin and induce chromatin closure on enhancers by recruiting nucleosome remodeling and deacetylase (NuRD) complex. PFs also establish H3K27me3-marked heterochromatin by recruiting and stimulating the Polycomb repressive complex 2 (PRC2) complex. Abbreviation: TF, transcription factor.

specifically in mitosis leads to a loss of pluripotency in mouse embryonic stem cells (ECs) [127]. Thus, pioneer factors have an important role in the maintenance of cell identity.

### Modulating pioneer factor activity

During tissue specification, pioneer factors have dynamic binding and opening activity on silent chromatin to progressively guide lineage-specific gene expression [7,10,17,19–21,24,100,117,128,129]. Pioneer factors do not generate ‘peaks’ in binding assays at all of their potential DNA-binding sites in a given cell [20,79,83]. Yet, examining subthreshold peaks reveals extensive binding at alternative cell sites, called ‘sampling’ sites [20,50]. Moreover, DNA motif enrichment analyses showed that pioneer



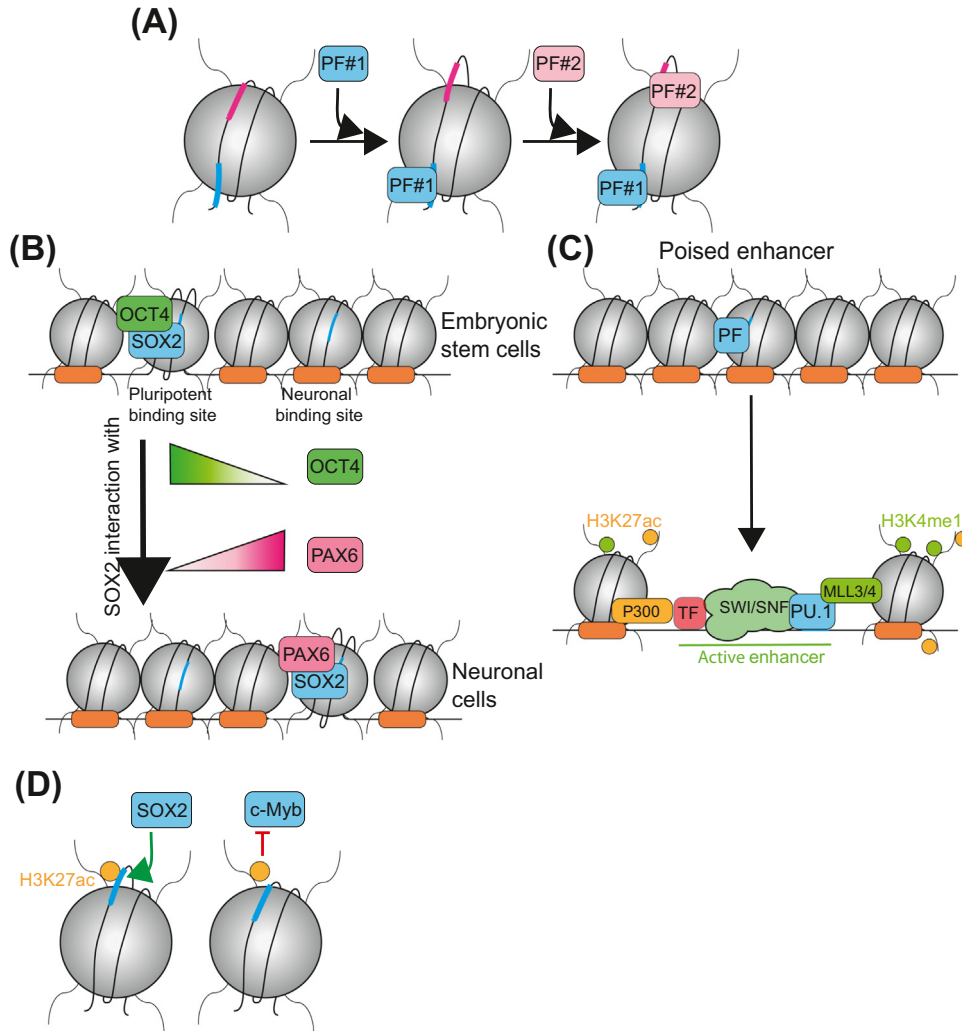
factors can bind chromatin regions that are weakly or not enriched for their DNA-binding motif [9,21,130]. In addition, pioneer factors bind closed chromatin regions without necessarily inducing DNA accessibility [36,90,107]. These observations indicate that pioneer factors are modulated for their stable binding events and chromatin-opening activity.

What are ways in which pioneer factors can be modulated during chromatin binding? Pioneer factor binding on nucleosomal DNA depends on nucleotide sequence and the position and orientation of its target DNA motif on the nucleosome surface [52,64,83,87,131,132]. A target DNA motif too close or facing the nucleosome surface could hinder factor binding due to steric clashes [64,72,87,131]. A recent *in vitro* study using a basic helix–loop–helix (bHLH) pioneer factor-binding motif tiled at one base-pair intervals across the DNA template showed that bHLH factors preferentially bind nucleosomes when the binding motif is exposed at the  $+/-7$  to  $+/-5$  superhelical region on the nucleosome surface [72], which is in agreement with an *in vitro* binding selection study for bHLH proteins [52]. To limit steric clashes, pioneer factors are able to bind partial or degenerate DNA motifs [65,75], due to flexibility of their DBD [68]. Given the emerging structural information showing that pioneer factors interact with the core histones within a bound nucleosome, binding to only a partial motif may be compensated for by histone interactions that stabilize binding.

During development, pioneer factors contribute to binding of other pioneer and/or transcription factors [88], leading to their cooperative binding on nucleosomal DNA [24,38,65,131]. This may be elicited by a first-bound pioneer factor to distort nucleosomal DNA leading to DNA motifs being accessible for other pioneer [58,64,72,87,131] or nonpioneer [107,133] factor binding (Figure 5A). An increase in pioneer factor concentration correlates with an increase in nucleosome binding [25,51,80] and with binding of new genomic regions during development [10,26,134]. More recently, pioneer factors have been identified as being able to form aggregates on chromatin [81,115,135–139]. The mechanism may ensure a local increase in pioneer factor concentration and their binding events [81]. It may also promote cooperation between pioneer factors, since several pioneer factors may colocalize in the same aggregate [136]. However, how pioneer factors form aggregates is poorly understood, since deletion of their unstructured domains, known to promote protein aggregate formation, does not always destabilize pioneer factor aggregates [135,136].

Pioneer factor function on silent chromatin can be regulated by the partner of the pioneer factor. For example, Tripartite motif containing 24 protein (Trim24), a cofactor of p53, prevents p53 opening activity on closed chromatin regions [140]. In addition, during cell differentiation, the loss of certain partners of pioneer factors and acquisition of new partners can lead to relocalization of pioneer factor genomic occupancy and modulation of chromatin opening activity [129,141]. Proteomic analysis of SOX2 in ESCs versus neuronal cells showed cell-specific interactors [128,142]. In ESCs, SOX2 interacts with OCT4. During ESC differentiation into neural progenitors, SOX2 loses its interaction with OCT4 and progressively gains interactions with a new partner: Paired box 6 (PAX6) [143] (Figure 5B). Thus, SOX2 targets and regulates different sets of lineage-specific genes during cell differentiation [143].

A ‘poised’ state of enhancers has been identified during cell differentiation, whereby the enhancers are bound by a pioneer factor but still harbor a closed conformation [19,107,144–146] (Figure 5C). They become open and, thus, activated, later during development [7]; such enabling of developmental competence characterized the initial pioneer factor description for FOXA1 in undifferentiated mouse endoderm [147]. Interestingly, in cells, ectopically expressed pioneer factor can bind chromatin within 0.5–24 h, but overt chromatin remodeling and opening may be delayed



## Trends in Genetics

**Figure 5. Modulation of pioneer factor (PF) binding and opening activity.** (A) Cooperative binding. A first PF binds and distorts nucleosomal DNA, leading to exposure of a second DNA motif and, thus, binding of a second factor. (B) PF binding and opening activity depend on PF cofactors. As described by Zhang *et al.* [143], SRY-box transcription factor 2 (SOX2) interacts with Octamer-binding transcription factor 4 (OCT4) in embryonic stem cells, then loses its interaction with OCT4 to gain an interaction with Paired box gene 6 (PAX6) during neural fate transition, leading to genome-wide SOX2 relocalization. (C) Poised PFs bind closed chromatin regions but do not induce their local opening. Closed chromatin opening may occur later during development. (D) Histone modifications might favor or impede PF binding/local opening activity. Abbreviations: MLL3/4, Mixed-lineage leukemia 3/4; c-Myb, Myb proto-oncogene protein; SWI/SNF, Switch/sucrose nonfermenting; TF, transcription factor.

by 1–5 days [90]. The delay could be due to the time required for subsequent recruitment of a nucleosome remodeler [61,129], a requirement for additional developmental signals [24,146], or loss of binding of a local inhibitor, which then allows chromatin opening [140].

Histone post-translational modifications (PTMs) might favor [67,148] or impede [21,149] pioneer factor binding or local opening action by inducing new steric clashes [64], altering electrostatic histone surface charges [149], or de/stabilizing pioneer factor–histone interactions or DNA–histone interactions [67]. Cryo-electron microscopy structural studies using H3K27-acetylated nucleosomes showed that acetylation induces DNA sliding, increases DNA-binding site exposition, and, thus,

partially modulates pioneer factor binding [67] (Figure 5D). Moreover, peptide microarray assays showed that some pioneer factors are unable to interact with specific modified histone tails [21,149]. For example, H3K27ac histone modification prevents Myb proto-oncogene protein (c-Myb) binding on the histone tail and, thus, poorly colocalizes on chromatin [149] (Figure 5D). This suggests that H3K27ac might protect active enhancers from aberrant c-Myb binding or ensures c-Myb detachment and recycling from chromatin after enhancer activation. Thus, some histone modifications, and the protein complexes that bind them, might repulse pioneer factors from chromatin to protect *cis*-regulatory elements from their action at specific stages of development.

Given all of the parameters that could affect pioneer factor binding, it may be no surprise that pioneer factors harbor diverse affinities and mechanism of actions on the three states of silent chromatin: naive chromatin, H3K9me3-marked, and H3K27me3-marked heterochromatin. First, pioneer factors appear to preferentially target naive chromatin regions [38], since most pioneer factor-bound regions have a closed conformation but contain low or no H3K9me3 and H3K27me3-repressive histone marks [17,20,107] (Figure 1B). Moreover, **H3K9me3-heterochromatin** had been characterized as a barrier to reprogramming [38,150]. Yet, recent epigenomic studies found that certain pioneer factors, such as Paired box 7 (PAX7) or TCF1, target regions covered by H3K27me3 [22–24,140] or H3K9me3 marks [21,22,90,140] (Figure 1B) and then induce underlying DNA sequence activation [22–24,140]. However, mechanisms of engagement of pioneer factors with these heterochromatin regions are poorly understood. During pituitary lineage differentiation, PAX7 has a weaker enrichment on heterochromatin regions than on naive chromatin regions. Moreover, PAX7 induces activation of enhancers covered by H3K9me2-heterochromatin, but not by H3K9me3-heterochromatin [90], suggesting different mechanisms of regulation between the diverse types of silent chromatin. It may be that H3K9me3- and **H3K27me3-heterochromatin** have to be destabilized by pioneer factor-independent mechanisms to facilitate pioneer factor binding and action on a more naive-like chromatin structure during cell differentiation.

Finally, recent studies revealed consequential PTMs on the pioneer factors themselves. The modifications might modulate their scanning, binding, or local opening activities [60,151–156]. For example, acetylation of PAX7 modulates its binding on chromatin and, thus, gene expression profile [154], while its methylation is required for its mitotic chromosome-binding activity [156]. Identification of PTMs on pioneer factors, the associated modifying enzymes, and their consequences for pioneer factor function is a key point for future work. Such studies will allow a better understanding of the roles of pioneer factors during development and diseases, such as cancer, where pioneer factors might be mutated on potential post-translational-modified residues or where modifying enzymes might be dysregulated or mutated.

### Concluding remarks

The nucleosome-targeting feature of pioneer factors enables them to drive development by targeting naive and heterochromatic domains to elicit chromatin opening or further closing and promote subsequent cell fate transitions. RNA-binding domains have been recently identified on pioneer factors [157], which impact DNA binding; however, their role in scanning, binding, and local opening of silent chromatin is unknown. In general, non-DBD segments, including apparently unstructured domains that are critical for local chromatin opening by pioneer factors [61], need to be characterized to understand their role in pioneer factor function (see [Outstanding questions](#)). Recent studies identified additional pioneer factor co-factors and PTMs, but underlying mechanisms need to be better investigated to fully understand the dynamic role of pioneer factors in terms of mechanisms and roles in cell differentiation. Finally, pioneer factors are able to target diverse types of silent chromatin [21–24,90]. However, it is unclear how pioneer factors recognize and reorganize H3K9me3- or H3K27me3-marked-heterochromatin. Addressing these

### Outstanding questions

What are different modes of interaction with nucleosomes for pioneer factors? What are mechanistic consequences of pioneer factor–nucleosome interactions on nucleosome-wrapped DNA and local chromatin decompaction? How do histone variants and histone modifications modulate pioneer factor–nucleosome interactions and chromatin opening by pioneer factors?

What is the role of non-DNA-binding domains, especially intrinsically disordered regions, in pioneer factor scanning, binding, and opening of chromatin? How do post-translational modifications of pioneer factors modulate the activities?

What are the developmental signals and their molecular mechanisms that activate enhancers that are ‘poised’ by pioneer factors?

How do pioneer factors distinguish diverse types of silent chromatin? How do pioneer factors recognize, destabilize, and reorganize H3K9me3- and H3K27me3-heterochromatin regions?

questions is key to fully understanding pioneer factor function on silent chromatin during development and in disease (see Outstanding questions).

### Acknowledgments

A.B. was supported by a postdoctoral EMBO fellowship (ALTF 266-2022). Work on pioneer factors in the lab of K.S.Z. is supported by NIH R01GM36477.

### Declaration of interests

The authors declare no conflicts of interest.

### References

- Thurman, R.E. *et al.* (2012) The accessible chromatin landscape of the human genome. *Nature* 489, 75–82
- Kundaje, A. *et al.* (2015) Integrative analysis of 111 reference human epigenomes. *Nature* 518, 317–330
- Filion, G.J. *et al.* (2010) Systematic protein location mapping reveals five principal chromatin types in *Drosophila* cells. *Cell* 143, 212–224
- Boyer, L.A. *et al.* (2006) Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* 441, 349–353
- Nicetto, D. *et al.* (2019) H3K9me3-heterochromatin loss at protein-coding genes enables developmental lineage specification. *Science* 363, 294–297
- Gassler, J. *et al.* (2022) Zygotic genome activation by the totipotency pioneer factor Nr5a2. *Science* 378, 1305–1315
- Lee, K. *et al.* (2019) FOXA2 is required for enhancer priming during pancreatic differentiation. *Cell Rep.* 28, 382–393
- Magli, A. *et al.* (2019) Time-dependent Pax3-mediated chromatin remodeling and cooperation with Six4 and Tead2 specify the skeletal myogenic lineage in developing mesoderm. *PLoS Biol.* 17, e3000153
- Päun, O. *et al.* (2023) Pioneer factor ASCL1 cooperates with the mSWI/SNF complex at distal regulatory elements to regulate human neural differentiation. *Genes Dev.* 37, 218–242
- Ungerback, J. *et al.* (2018) Pioneering, chromatin remodeling, and epigenetic constraint in early T-cell gene regulation by SPI1 (PU.1). *Genome Res.* 28, 1508–1519
- Veil, M. *et al.* (2019) Pou5f3, SoxB1, and Nanog remodel chromatin on high nucleosome affinity regions at zygotic genome activation. *Genome Res.* 29, 383–395
- Larson, E.D. *et al.* (2021) Pioneering the developmental frontier. *Mol. Cell* 81, 1640–1650
- Zaret, K.S. (2020) Pioneer transcription factors initiating gene network changes. *Annu. Rev. Genet.* 54, 367–385
- Balsalobre, A. and Drouin, J. (2022) Pioneer factors as master regulators of the epigenome and cell fate. *Nat. Rev. Mol. Cell Biol.* 23, 449–464
- Mayran, A. and Drouin, J. (2018) Pioneer transcription factors shape the epigenetic landscape. *J. Biol. Chem.* 293, 13795–13804
- Heslop, J.A. *et al.* (2021) GATA6 defines endoderm fate by controlling chromatin accessibility during differentiation of human-induced pluripotent stem cells. *Cell Rep.* 35, 109145
- Thompson, J.J. *et al.* (2022) Extensive co-binding and rapid redistribution of NANOG and GATA6 during emergence of divergent lineages. *Nat. Commun.* 13, 4257
- Spruce, C. *et al.* (2020) HELLS and PRDM9 form a pioneer complex to open chromatin at meiotic recombination hot spots. *Genes Dev.* 34, 398–412
- Cernilogar, F.M. *et al.* (2019) Pre-marked chromatin and transcription factor co-binding shape the pioneering activity of Foxa2. *Nucleic Acids Res.* 47, 9069–9086
- Donaghey, J. *et al.* (2018) Genetic determinants and epigenetic effects of pioneer-factor occupancy. *Nat. Genet.* 50, 250–258
- Fuglerud, B.M. *et al.* (2022) SOX9 reprograms endothelial cells by altering the chromatin landscape. *Nucleic Acids Res.* 50, 8547–8565
- Johnson, J.L. *et al.* (2018) Lineage-determining transcription factor TCF-1 initiates the epigenetic identity of T cells. *Immunity* 48, 243–257
- Pataskar, A. *et al.* (2016) NeuroD1 reprograms chromatin and transcription factor landscapes to induce the neuronal program. *EMBO J.* 35, 24–45
- van Oevelen, C. *et al.* (2015) C/EBP $\alpha$  activates pre-existing and de novo macrophage enhancers during induced pre-B cell transdifferentiation and myelopoiesis. *Stem Cell Rep.* 5, 232–247
- Duan, J. *et al.* (2021) CLAMP and Zelda function together to promote *Drosophila* zygotic genome activation. *eLife* 10, e69937
- Miao, L. *et al.* (2022) The landscape of pioneer factor activity reveals the mechanisms of chromatin reprogramming and genome activation. *Mol. Cell* 82, 986–1002
- Pályi, M. *et al.* (2020) Chromatin accessibility established by Pou5f3, Sox19b and Nanog primes genes for activity during zebrafish genome activation. *PLoS Genet.* 16, e1008546
- Rothstein, M. and Simoes-Costa, M. (2020) Heterodimerization of TFAP2 pioneer factors drives epigenomic remodeling during neural crest specification. *Genome Res.* 30, 35–48
- Delás, M.J. *et al.* (2023) Developmental cell fate choice in neural tube progenitors employs two distinct cis-regulatory strategies. *Dev. Cell* 58, 3–17
- Gaskill, M.M. *et al.* (2021) GAF is essential for zygotic genome activation and chromatin accessibility in the early *Drosophila* embryo. *eLife* 10, e66668
- Parolia, A. *et al.* (2019) Distinct structural classes of activating FOXA1 alterations in advanced prostate cancer. *Nature* 571, 413–418
- Takaku, M. *et al.* (2020) Cancer-specific mutation of GATA3 disrupts the transcriptional regulatory network governed by estrogen receptor alpha, FOXA1 and GATA3. *Nucleic Acids Res.* 48, 4756–4768
- Di Giammartino, D.C. *et al.* (2019) KLF4 is involved in the organization and regulation of pluripotency-associated three-dimensional enhancer networks. *Nat. Cell Biol.* 21, 1179–1190
- Gong, W. *et al.* (2022) ETV2 functions as a pioneer factor to regulate and reprogram the endothelial lineage. *Nat. Cell Biol.* 24, 672–684
- Horisawa, K. *et al.* (2020) The dynamics of transcriptional activation by hepatic reprogramming factors. *Mol. Cell* 79, 660–676
- Takaku, M. *et al.* (2016) GATA3-dependent cellular reprogramming requires activation-domain dependent recruitment of a chromatin remodeler. *Genome Biol.* 17, 36
- Wapinski, O.L. *et al.* (2017) Rapid chromatin switch in the direct reprogramming of fibroblasts to neurons. *Cell Rep.* 20, 3236–3247
- Soufi, A. *et al.* (2012) Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell* 151, 994–1004
- Chen, K. *et al.* (2020) Heterochromatin loosening by the Oct4 linker region facilitates Klf4 binding and iPSC reprogramming. *EMBO J.* 39, e99165
- Sunkel, B.D. and Stanton, B.Z. (2021) Pioneer factors in development and cancer. *iScience* 24, 103132
- Pandey, P.R. *et al.* (2017) PAX3-FOXO1 is essential for tumour initiation and maintenance but not recurrence in a human myoblast model of rhabdomyosarcoma. *J. Pathol.* 241, 626–637



42. Gryder, B.E. *et al.* (2017) PAX3-FOXO1 establishes myogenic super enhancers and confers BET bromodomain vulnerability. *Cancer Discov.* 7, 884–899
43. Loupe, J.M. *et al.* (2016) Comparative transcriptomic analysis reveals the oncogenic fusion protein PAX3-FOXO1 globally alters mRNA and miRNA to enhance myoblast invasion. *Oncogenesis* 5, e246
44. Sunkel, B.D. *et al.* (2021) Evidence of pioneer factor activity of an oncogenic fusion transcription factor. *iScience* 24, 102867
45. Harrison, M.M. *et al.* (2011) Zelda binding in the early *Drosophila melanogaster* embryo marks regions subsequently activated at the maternal-to-zygotic transition. *PLoS Genet.* 7, e1002266
46. Liang, H.L. *et al.* (2008) The zinc-finger protein Zelda is a key activator of the early zygotic genome in *Drosophila*. *Nature* 456, 400–403
47. Choi, S.H. *et al.* (2016) DUX4 recruits p300/CBP through its C-terminus and induces global H3K27 acetylation changes. *Nucleic Acids Res.* 44, 5161–5173
48. De Iaco, A. *et al.* (2017) DUX-family transcription factors regulate zygotic genome activation in placental mammals. *Nat. Genet.* 49, 941–945
49. Hansen, J.L. *et al.* (2022) A test of the pioneer factor hypothesis using ectopic liver gene activation. *eLife* 11, e73358
50. Lerner, J. *et al.* (2023) Different chromatin-scanning modes lead to targeting of compacted chromatin by pioneer factors FOXA1 and SOX2. *Cell Rep.* 42, 112748
51. Donovan, B.T. *et al.* (2023) Basic helix-loop-helix pioneer factors interact with the histone octamer to invade nucleosomes and generate nucleosome-depleted regions. *Mol. Cell* 83, 1251–1263
52. Zhu, F. *et al.* (2018) The interaction landscape between transcription factors and the nucleosome. *Nature* 562, 76–81
53. Donovan, B.T. *et al.* (2019) Dissociation rate compensation mechanism for budding yeast pioneer transcription factors. *eLife* 8, e43008
54. Garcia, D.A. *et al.* (2021) An intrinsically disordered region-mediated confinement state contributes to the dynamics and function of transcription factors. *Mol. Cell* 81, 1484–1498
55. Lerner, J. *et al.* (2020) Two-parameter mobility assessments discriminate diverse regulatory factor behaviors in chromatin. *Mol. Cell* 79, 677–688
56. Tang, X. *et al.* (2022) Kinetic principles underlying pioneer function of GAGA transcription factor in live cells. *Nat. Struct. Mol. Biol.* 29, 665–676
57. Cirillo, L.A. and Zaret, K.S. (1999) An early developmental transcription factor complex that is more stable on nucleosome core particles than on free DNA. *Mol. Cell* 4, 961–969
58. Dodonova, S.O. *et al.* (2020) Nucleosome-bound SOX2 and SOX11 structures elucidate pioneer factor function. *Nature* 580, 669–672
59. Echigoya, K. *et al.* (2020) Nucleosome binding by the pioneer transcription factor OCT4. *Sci. Rep.* 10, 11832
60. Fernandez Garcia, M. *et al.* (2019) Structural features of transcription factors associating with nucleosome binding. *Mol. Cell* 75, 921–932
61. Frederick, M.A. *et al.* (2023) A pioneer factor locally opens compacted chromatin to enable targeted ATP-dependent nucleosome remodeling. *Nat. Struct. Mol. Biol.* 30, 31–37
62. Huertas, J. *et al.* (2020) Nucleosomal DNA dynamics mediate Oct4 pioneer factor binding. *Biophys. J.* 118, 2280–2296
63. MacCarthy, C.M. *et al.* (2022) OCT4 interprets and enhances nucleosome flexibility. *Nucleic Acids Res.* 50, 10311–10327
64. Malaga Gadea, F.C. and Nikolova, E.N. (2023) Structural plasticity of pioneer factor Sox2 and DNA bendability modulate nucleosome engagement and Sox2-Oct4 synergism. *J. Mol. Biol.* 435, 167916
65. Michael, A.K. *et al.* (2020) Mechanisms of OCT4-SOX2 motif readout on nucleosomes. *Science* 368, 1460–1465
66. Pluta, R. *et al.* (2022) Molecular basis for DNA recognition by the maternal pioneer transcription factor FoxH1. *Nat. Commun.* 13, 7279
67. Sinha, K.K. *et al.* (2023) Histone modifications regulate pioneer transcription factor cooperativity. *Nature* 619, 378–384
68. Soufi, A. *et al.* (2015) Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell* 161, 555–568
69. Roberts, G.A. *et al.* (2021) Dissecting OCT4 defines the role of nucleosome binding in pluripotency. *Nat. Cell Biol.* 23, 834–845
70. Ozden, B. *et al.* (2023) Molecular mechanism of nucleosome recognition by the pioneer transcription factor Sox. *J. Chem. Inf. Model.* 63, 3839–3853
71. Guan, R. *et al.* (2023) Structural mechanism of LIN28B nucleosome targeting by OCT4. *Mol. Cell* 83, 1970–1982
72. Michael, A.K. *et al.* (2023) Cooperation between bHLH transcription factors and histones for DNA access. *Nature* 619, 385–393
73. Jin, R. *et al.* (2021) LEAFY is a pioneer transcription factor and licenses cell reprogramming to floral fate. *Nat. Commun.* 12, 626
74. Mivelaz, M. *et al.* (2020) Chromatin fiber invasion and nucleosome displacement by the Rap1 transcription factor. *Mol. Cell* 77, 488–500
75. Meers, M.P. *et al.* (2019) Pioneer factor-nucleosome binding events during differentiation are motif encoded. *Mol. Cell* 75, 562–575
76. Iwafuchi, M. *et al.* (2020) Gene network transitions in embryos depend upon interactions between a pioneer transcription factor and core histones. *Nat. Genet.* 52, 418–427
77. Luzete-Monteiro, E. and Zaret, K.S. (2022) Structures and consequences of pioneer factor binding to nucleosomes. *Curr. Opin. Struct. Biol.* 75, 102425
78. King, H.W. and Klose, R.J. (2017) The pioneer factor OCT4 requires the chromatin remodeler BRG1 to support gene regulatory element function in mouse embryonic stem cells. *eLife* 6, e22631
79. Minderjahn, J. *et al.* (2020) Mechanisms governing the pioneering and redistribution capabilities of the non-classical pioneer PU.1. *Nat. Commun.* 11, 402
80. Chen, J. *et al.* (2014) Single-molecule dynamics of enhanceosome assembly in embryonic stem cells. *Cell* 156, 1274–1285
81. Mir, M. *et al.* (2018) Dynamic multifactor hubs interact transiently with sites of active transcription in *Drosophila* embryos. *eLife* 7, e40497
82. Cirillo, L.A. *et al.* (1998) Binding of the winged-helix transcription factor HNF3 to a linker histone site on the nucleosome. *EMBO J.* 17, 244–254
83. Barozzi, I. *et al.* (2014) Coregulation of transcription factor binding and nucleosome occupancy through DNA features of mammalian enhancers. *Mol. Cell* 54, 844–857
84. Kumar, D.K. *et al.* (2023) Complementary strategies for directing in vivo transcription factor binding through DNA binding domains and intrinsically disordered regions. *Mol. Cell* 83, 1462–1473
85. Pelletier, A. *et al.* (2021) Pax7 pioneer factor action requires both paired and homeo DNA binding domains. *Nucleic Acids Res.* 49, 7424–7436
86. Boller, S. *et al.* (2016) Pioneering activity of the C-terminal domain of EBF1 shapes the chromatin landscape for B cell programming. *Immunity* 44, 527–541
87. Tan, C. and Takada, S. (2020) Nucleosome allostery in pioneer transcription factor binding. *Proc. Natl. Acad. Sci. U. S. A.* 117, 20586–20596
88. Brennan, K.J. *et al.* (2023) Chromatin accessibility in the *Drosophila* embryo is determined by transcription factor pioneering and enhancer activation. *Dev. Cell* 58, 1898–1916
89. Iwafuchi-Doi, M. *et al.* (2016) The pioneer transcription factor FoxA maintains an accessible nucleosome configuration at enhancers for tissue-specific gene activation. *Mol. Cell* 62, 79–91
90. Mayran, A. *et al.* (2018) Pioneer factor Pax7 deploys a stable enhancer repertoire for specification of cell fate. *Nat. Genet.* 50, 259–269
91. Li, T. *et al.* (2022) Dynamic nucleosome landscape elicits a non-canonical GATA2 pioneer model. *Nat. Commun.* 13, 3145
92. Cirillo, L.A. *et al.* (2002) Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol. Cell* 9, 279–289
93. Tsukiyama, T. *et al.* (1994) ATP-dependent nucleosome disruption at a heat-shock promoter mediated by binding of GAGA transcription factor. *Nature* 367, 525–532
94. Wolf, B.K. *et al.* (2023) Cooperation of chromatin remodeling SWI/SNF complex and pioneer factor AP-1 shapes 3D enhancer landscapes. *Nat. Struct. Mol. Biol.* 30, 10–21

95. Yang, Y. *et al.* (2023) The pioneer factor SOX9 competes for epigenetic factors to switch stem cell fates. *Nat. Cell Biol.* 25, 1185–1195
96. Wang, Y. *et al.* (2020) A prion-like domain in transcription factor EBF1 promotes phase separation and enables B cell programming of progenitor chromatin. *Immunity* 53, 1151–1167
97. Chambers, C. *et al.* (2023) SWI/SNF blockade disrupts PU.1-directed enhancer programs in normal hematopoietic cells and acute myeloid leukemia. *Cancer Res.* 83, 983–996
98. Judd, J. *et al.* (2021) Pioneer-like factor GAF cooperates with PBAP (SWI/SNF) and NURF (ISWI) to regulate transcription. *Genes Dev.* 35, 147–156
99. Xiong, L. *et al.* (2022) Oct4 differentially regulates chromatin opening and enhancer transcription in pluripotent stem cells. *eLife* 11, e71533
100. Li, R. *et al.* (2018) Dynamic EBF1 occupancy directs sequential epigenetic and transcriptional events in B-cell programming. *Genes Dev.* 32, 96–111
101. Suzuki, T. *et al.* (2017) RUNX1 regulates site specificity of DNA demethylation by recruitment of DNA demethylation machineries in hematopoietic cells. *Blood Adv.* 1, 1699–1711
102. Tan, D.S. *et al.* (2023) The homeodomain of Oct4 is a dimeric binder of methylated CpG elements. *Nucleic Acids Res.* 51, 1120–1138
103. Vanzan, L. *et al.* (2021) High throughput screening identifies SOX2 as a super pioneer factor that inhibits DNA methylation maintenance at its binding sites. *Nat. Commun.* 12, 3337
104. Wiench, M. *et al.* (2011) DNA methylation status predicts cell type-specific enhancer activity. *EMBO J.* 30, 3028–3039
105. Yang, Y.A. *et al.* (2016) FOXA1 potentiates lineage-specific enhancer activation through modulating TET1 expression and function. *Nucleic Acids Res.* 44, 8153–8164
106. Reizel, Y. *et al.* (2021) FoxA-dependent demethylation of DNA initiates epigenetic memory of cellular identity. *Dev. Cell* 56, 602–612
107. Mayran, A. *et al.* (2019) Pioneer and nonpioneer factor cooperation drives lineage specific chromatin opening. *Nat. Commun.* 10, 3807
108. Jozwik, K.M. *et al.* (2016) FOXA1 directs H3K4 monomethylation at enhancers via recruitment of the methyltransferase MLL3. *Cell Rep.* 17, 2715–2723
109. Lee, J.W. *et al.* (2019) RUNX3 regulates cell cycle-dependent chromatin dynamics by functioning as a pioneer factor of the restriction-point. *Nat. Commun.* 10, 1897
110. Lemma, R.B. *et al.* (2022) Pioneer transcription factors are associated with the modulation of DNA methylation patterns across cancers. *Epigenetics Chromatin* 15, 13
111. Magli, A. *et al.* (2019) Pax3 cooperates with Ldb1 to direct local chromosome architecture during myogenic lineage specification. *Nat. Commun.* 10, 2316
112. Wang, W. *et al.* (2022) TCF-1 promotes chromatin interactions across topologically associating domains in T cell progenitors. *Nat. Immunol.* 23, 1052–1062
113. Christou-Kent, M. *et al.* (2023) CEBPA phase separation links transcriptional activity and 3D chromatin hubs. *Cell Rep.* 42, 112897
114. Boija, A. *et al.* (2018) Transcription factors activate genes through the phase-separation capacity of their activation domains. *Cell* 175, 1842–1855
115. Kuznetsova, K. *et al.* (2023) Nanog organizes transcription bodies. *Curr. Biol.* 33, 164–173
116. Hosokawa, H. *et al.* (2018) Transcription factor PU.1 represses and activates gene expression in early T cells by redirecting partner transcription factor binding. *Immunity* 48, 1119–1134
117. Chronis, C. *et al.* (2017) Cooperative binding of transcription factors orchestrates reprogramming. *Cell* 168, 442–459
118. Sekiya, T. and Zaret, K.S. (2007) Repression by Groucho/TLE/Grg proteins: genomic site recruitment generates compacted chromatin in vitro and impairs activator binding in vivo. *Mol. Cell* 28, 291–303
119. Gregoricchio, S. *et al.* (2022) HDAC1 and PRC2 mediate combinatorial control in SPI1/PU.1-dependent gene repression in murine erythroleukaemia. *Nucleic Acids Res.* 50, 7938–7958
120. Reed-Inderbitzin, E. *et al.* (2006) RUNX1 associates with histone deacetylases and SUV39H1 to repress transcription. *Oncogene* 25, 5777–5786
121. Gaskill, M.M. *et al.* (2023) Localization of the *Drosophila* pioneer factor GAF to subnuclear foci is driven by DNA binding and required to silence satellite repeat expression. *Dev. Cell* 58, 1610–1624
122. Ito, K. and Zaret, K.S. (2022) Maintaining transcriptional specificity through mitosis. *Annu. Rev. Genomics Hum. Genet.* 23, 53–71
123. Raccaud, M. *et al.* (2019) Mitotic chromosome binding predicts transcription factor properties in interphase. *Nat. Commun.* 10, 487
124. Silvério-Alves, R. *et al.* (2023) GATA2 mitotic bookmarking is required for definitive haematopoiesis. *Nat. Commun.* 14, 4645
125. Caravaca, J.M. *et al.* (2013) Bookmarking by specific and non-specific binding of FoxA1 pioneer factor to mitotic chromosomes. *Genes Dev.* 27, 251–260
126. Festuccia, N. *et al.* (2016) Mitotic binding of Esrrb marks key regulatory regions of the pluripotency network. *Nat. Cell Biol.* 18, 1139–1148
127. Deluz, C. *et al.* (2016) A role for mitotic bookmarking of SOX2 in pluripotency and differentiation. *Genes Dev.* 30, 2538–2550
128. Hovland, A.S. *et al.* (2022) Pluripotency factors are repurposed to shape the epigenomic landscape of neural crest cells. *Dev. Cell* 57, 2257–2272
129. Vierbuchen, T. *et al.* (2017) AP-1 transcription factors and the BAF complex mediate signal-dependent enhancer selection. *Mol. Cell* 68, 1067–1082
130. Larson, E.D. *et al.* (2021) Cell-type-specific chromatin occupancy by the pioneer factor Zelda drives key developmental transitions in *Drosophila*. *Nat. Commun.* 12, 7153
131. Li, S. *et al.* (2019) Nonreciprocal and conditional cooperativity directs the pioneer activity of pluripotency transcription factors. *Cell Rep.* 28, 2689–2703.e4
132. Tanaka, H. *et al.* (2020) Interaction of the pioneer transcription factor GATA3 with nucleosomes. *Nat. Commun.* 11, 4136
133. Liu, Z. and Kraus, W.L. (2017) Catalytic-independent functions of PARP-1 determine Sox2 pioneer activity at intractable genomic loci. *Mol. Cell* 65, 589–603
134. Sharma, A. *et al.* (2020) GATA6 mutations in hiPSCs inform mechanisms for maldevelopment of the heart, pancreas, and diaphragm. *eLife* 9, e53278
135. Nguyen, T. *et al.* (2022) Chromatin sequesters pioneer transcription factor Sox2 from exerting force on DNA. *Nat. Commun.* 13, 3988
136. Sharma, R. *et al.* (2021) Liquid condensation of reprogramming factor KLF4 with DNA provides a mechanism for chromatin organization. *Nat. Commun.* 12, 5579
137. Kang, H. *et al.* (2018) Sequence-dependent DNA condensation as a driving force of DNA phase separation. *Nucleic Acids Res.* 46, 9401–9413
138. Pownall, M.E. *et al.* (2023) Chromatin expansion microscopy reveals nanoscale organization of transcription and chromatin. *Science* 381, 92–100
139. Dufourt, J. *et al.* (2018) Temporal control of gene expression by the pioneer factor Zelda through transient interactions in hubs. *Nat. Commun.* 9, 5194
140. Isbel, L. *et al.* (2023) Readout of histone methylation by Trim24 locally restricts chromatin opening by p53. *Nat. Struct. Mol. Biol.* 30, 948–957
141. Bullerwell, C.E. *et al.* (2021) EBF1 drives hallmark B cell gene expression by enabling the interaction of PAX5 with the MLL H3K4 methyltransferase complex. *Sci. Rep.* 11, 1537
142. Bunina, D. *et al.* (2020) Genomic rewiring of SOX2 chromatin interaction network during differentiation of ESCs to postmitotic neurons. *Cell Syst.* 10, 480–494
143. Zhang, S. *et al.* (2019) OCT4 and PAX6 determine the dual function of SOX2 in human ESCs as a key pluripotent or neural factor. *Stem Cell Res. Ther.* 10, 122
144. Jacobs, J. *et al.* (2018) The transcription factor Grainy head primes epithelial enhancers for spatiotemporal activation by displacing nucleosomes. *Nat. Genet.* 50, 1011–1020
145. Johnson, T.A. *et al.* (2018) Conventional and pioneer modes of glucocorticoid receptor interaction with enhancer chromatin in vivo. *Nucleic Acids Res.* 46, 203–214
146. Wang, A. *et al.* (2015) Epigenetic priming of enhancers predicts developmental competence of hESC-derived endodermal lineage intermediates. *Cell Stem Cell* 16, 386–399

147. Gualdi, R. *et al.* (1996) Hepatic specification of the gut endoderm in vitro: cell signaling and transcriptional control. *Genes Dev.* 10, 1670–1682
148. Zhang, S. *et al.* (2022) H3K27ac nucleosomes facilitate HMGN localization at regulatory sites to modulate chromatin binding of transcription factors. *Commun. Biol.* 5, 159
149. Fuglerud, B.M. *et al.* (2018) The pioneer factor activity of c-Myb involves recruitment of p300 and induction of histone acetylation followed by acetylation-induced chromatin dissociation. *Epigenetics Chromatin* 11, 35
150. Kaemena, D.F. *et al.* (2023) B1 SINE-binding ZFP266 impedes mouse iPSC generation through suppression of chromatin opening mediated by reprogramming factors. *Nat. Commun.* 14, 488
151. Ali, F.R. *et al.* (2020) Dephosphorylation of the proneural transcription factor ASCL1 re-engages a latent post-mitotic differentiation program in neuroblastoma. *Mol. Cancer Res.* 18, 1759–1766
152. Gao, S. *et al.* (2020) Chromatin binding of FOXA1 is promoted by LSD1-mediated demethylation in prostate cancer. *Nat. Genet.* 52, 1011–1017
153. Loffreda, A. *et al.* (2017) Live-cell p53 single-molecule binding is modulated by C-terminal acetylation and correlates with transcriptional activity. *Nat. Commun.* 8, 313
154. Sincennes, M.C. *et al.* (2021) Acetylation of PAX7 controls muscle stem cell self-renewal and differentiation potential in mice. *Nat. Commun.* 12, 3253
155. Woods, L.M. *et al.* (2022) Elevated ASCL1 activity creates de novo regulatory elements associated with neuronal differentiation. *BMC Genomics* 23, 255
156. Wu, T.F. *et al.* (2015) Loading of PAX3 to mitotic chromosomes is mediated by arginine methylation and associated with Waardenburg syndrome. *J. Biol. Chem.* 290, 20556–20564
157. Oksuz, O. *et al.* (2023) Transcription factors interact with RNA to regulate genes. *Mol. Cell* 607, 176–184

# Mechanisms and consequences of sex differences in immune responses

Shannon E. Dunn<sup>1,2</sup>, Whitney A. Perry<sup>3</sup> & Sabra L. Klein<sup>4</sup>✉

## Abstract

Biological sex differences refer to differences between males and females caused by the sex chromosome complement (that is, XY or XX), reproductive tissues (that is, the presence of testes or ovaries), and concentrations of sex steroids (that is, testosterone or oestrogens and progesterone). Although these sex differences are binary for most human individuals and mice, transgender individuals receiving hormone therapy, individuals with genetic syndromes (for example, Klinefelter and Turner syndromes) and people with disorders of sexual development reflect the diversity in sex-based biology. The broad distribution of sex steroid hormone receptors across diverse cell types and the differential expression of X-linked and autosomal genes means that sex is a biological variable that can affect the function of all physiological systems, including the immune system. Sex differences in immune cell function and immune responses to foreign and self antigens affect the development and outcome of diverse diseases and immune responses.

## Sections

Introduction

Causes of sex differences in immune cells

Sex differences in immune responses

Consequences of sex differences

Conclusions

<sup>1</sup>Department of Immunology, University of Toronto, Toronto, Ontario, Canada. <sup>2</sup>Women's College Research Institute, Women's College Hospital, Toronto, Ontario, Canada. <sup>3</sup>Division of Geographic Medicine and Infectious Diseases, Tufts Medical Center, Boston, MA, USA. <sup>4</sup>W. Harry Feinstone Department of Molecular Microbiology and Immunology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA. ✉e-mail: [sklein2@jhu.edu](mailto:sklein2@jhu.edu)



## Key points

- Biological sex refers to the differences between males and females caused by a differential sex chromosome complement (most commonly XX or XY in mammals), whereas gender refers to socially constructed norms that determine the roles, relationships and positional power of men and women across their lifetime.
- Current evidence supports the notion that sex chromosomes and gonadal hormones modulate the number and functions of immune cells.
- There are well characterized sex differences in the innate and adaptive immune response; there is strong evidence that type I and type II interferon signalling and humoral responses are greater in females than in males across diverse species.
- Sex differences in both innate and adaptive immunity contribute to the increased prevalence of autoimmunity in females and increase the propensity of females to reject their organs post-transplantation.
- Sex differences research is uncovering novel therapeutic pathways that could be targeted to improve disease outcomes in all sexes.

## Introduction

Biological sex refers to the differences between males and females caused by a differential sex chromosome complement (most commonly XX or XY in mammals). These differences in sex chromosomes specify the reproductive tissues that develop (that is, ovaries or testes), and the concentrations of sex steroids (that is, oestrogens and progesterone (P4) or androgens). Sex differences are typically binary in mammals with males being XY and females being XX; however, variations in biological sex exist. For example, intersex individuals are born with reproductive characteristics of both males and females; individuals with Turner syndrome have an XO chromosome complement where the 'O' indicates an absent or non-functional X chromosome, and those with Klinefelter syndrome have an XXY chromosome complement. Although variation in biological sex is rare, these conditions deserve empirical consideration in the context of susceptibility to immune-related diseases, including cancers and autoimmune diseases. Whereas sex is a purely biological assignment, gender refers to the socially constructed norms that determine the roles, relationships and positional power of individuals across their lifetime<sup>1</sup>. Gender is not a binary construct, but can be used to define men and women<sup>1</sup>. Of note, sex and gender are not mutually exclusive and can intersect to affect disease outcomes<sup>1</sup>. For example, XX females may vary along a spectrum for a defined immune parameter because of differences in gender roles and behaviours that affect immunity. For instance, gender differences in the experience of pregnancy, occupational exposures, or stress can differentially affect immunity<sup>1</sup>. Although our understanding of the influence of gender on immunity is in its infancy, much progress has been made in our understanding of how biological sex influences the immune system.

Sex differences in immune cell function can occur as a result of sex differences in the expression of X-encoded or Y-encoded genes in immune cells, or differences in the expression of autosomal genes in immune cells as a result of steroid receptor signalling and epigenetic modifications<sup>2</sup>. Sex differences in immune cell function have been

observed across various innate and adaptive immune cell types, both in the resting state and in the context of disease, including allergy, autoimmunity, infection, inflammatory diseases, cancer and organ rejection post-transplantation<sup>2–5</sup>. Sex differences in immunity not only affect disease pathogenesis, but also the response of individuals to treatments including immunomodulatory therapies, vaccines and immune checkpoint inhibitors<sup>6</sup>. Improved understanding of the mechanisms that underlie sex differences in immunity, the breadth of their impact, and how these differences can be used to inform treatment optimization and testing of therapies in human clinical trials will require the concerted efforts of basic scientists, clinicians and epidemiologists to develop a consensus about sex differences in immunity. Here, we highlight our current understanding of sex differences in immunity. We also describe the key mechanisms that mediate sex differences in immune responses and illustrate the functional relevance of such differences for immune-related diseases based on data from preclinical models and clinical studies.

## Causes of sex differences in immune cells

### Genetic differences

The XX complement provides female mammals with two copies of the X chromosome, one derived from the male parent (Xp) and one from the female (Xm). Biological females are a mosaic of Xp and Xm because one X chromosome is randomly inactivated in every cell of a female to compensate for gene dosage effects. X chromosome mosaicism provides many biological advantages to female mammals. For example, mutations in one X chromosome only affect half of the cells in a female, but all cells in a male. The XX chromosome complement also results in allelic diversity that can affect responses to immune challenges<sup>7,8</sup>. For example, people with Klinefelter syndrome have immune responses that are more like that of biological females, at least in the context of autoimmune diseases<sup>9</sup>. Females with Turner syndrome have lower lymphocyte counts and reduced antibody production than XX females<sup>10</sup>. Because the sex chromosome complement also specifies gonad development, the differences in immunity among individuals with XXY and XO genotypes may also relate to effects on sex hormone concentrations<sup>9,11</sup>.

Genetic differences also specify immunological differences between the sexes. A gene expression analysis of immune cells from human males and females revealed differences in the expression of 1,875 transcripts that encoded 1,553 unique proteins and 196 long non-coding RNAs; 7% of these sex-based differentially expressed transcripts were encoded on sex chromosomes, with the remaining encoded on autosomes<sup>12</sup>. The X chromosome contains many immune-related genes, including Toll-like receptor 7 (*TLR7*), *IRAK1*, *FOXP3* and *CD40L*, as well as many microRNAs (miRNAs) and histone-modifying enzymes (such as *KDM6A*) that are important for the regulation of genes involved in immune responses<sup>7,13</sup>. Several of these genes show evidence of escape from X-inactivation in mice and humans. For example, B cells from female humans express higher levels of the X-escapee gene *TLR7* than those from male humans. This greater expression of *TLR7* in human female B cells is associated with enhanced antibody responses, predisposing females to some autoimmune disorders<sup>14</sup>, but also contributes to a stronger immune response following influenza vaccination in mice<sup>15</sup>. Conversely, deleterious mutations in *TLR7*, which are only observed among males, are associated with increased susceptibility to viral infections, including SARS-CoV-2 infection<sup>16</sup>. Another escapee of X-inactivation that can have broad effects on gene expression is the lysine demethylase *KDM6A* (*Kdm6a* in mice), which encodes the

histone-modifying enzyme UTX<sup>13</sup>. *KDM6A* is one of the top genes that is differentially expressed between male and female human T cells<sup>13</sup> and natural killer (NK) cells<sup>17</sup>. Loss-of-function studies in mice have shown that UTX regulates the expression of genes involved in both T helper 1 (T<sub>H</sub>1) and T<sub>H</sub>2 pathways<sup>13</sup> as well as NK cell effector function<sup>17</sup> (discussed later). In addition, a disproportionately high number of miRNAs (~10% of those in the human genome) are encoded on the X chromosome<sup>18</sup>, and some of these contribute to sex differences in immunity<sup>19</sup>.

The contribution of sex chromosome complement to immunity in mice has been studied using the four core genotype (FCG) model wherein the sex chromosome complement is disconnected from gonadal sex. This model is achieved by deleting the testes-determining gene *Sry* from the Y chromosome (ChrY) of mice and transgenically (Tg) expressing it on chromosome 3. This approach enables Mendelian inheritance of *Sry*, producing the four core genotypes: gonadal XX females, gonadal XY<sup>*Sry*-/-</sup> females, gonadal XY males, and gonadal XX<sup>*Sry*-Tg</sup> males<sup>20</sup>. Gene expression analysis of activated T cells from XY<sup>*Sry*-/-</sup> and XX mice demonstrated that some X-encoded genes are actually expressed at higher levels in XY<sup>*Sry*-/-</sup> than in XX T cells. This is due to reduced methylation of genes on the X chromosome compared with the Xp chromosome, whereby mice that only inherit one ChrX have less methylation at X-encoded gene loci than mice that inherit two ChrX<sup>21</sup>.

Only biological males can be affected by ChrY polymorphisms. ChrY consomic mice – that is, mice that are genetically identical with the exception of the ChrY – have been used to study how the ChrY regulates immune function in the context of autoimmunity and viral infection<sup>22</sup>. Genetic variation in ChrY affects the severity of experimental autoimmune encephalomyelitis (EAE), experimental myocarditis and influenza A virus infection (IAV)<sup>22,23</sup>. Differences in immune phenotype between ChrY consomic strains correlate with copy number variation in Y-encoded multicopy genes, whereas differences in the development of lung inflammation and survival of different male ChrY consomic mice after IAV infection is associated with differential expression of small RNAs in the lung<sup>23</sup>. Another important process that affects the XY complement is loss of the ChrY that can occur in human leukocytes during ageing, which leads to the dysregulation of autosomal genes in leukocytes of males<sup>24</sup>.

Thus, sex chromosome complement, the activity of gene regulatory elements on sex chromosomes, and the expression of sex chromosomal genes contribute to observed sex differences in gene expression in immune cells at steady state. Moreover, these factors can affect immune responses to viral infections and can influence the development of autoimmunity.

## Sex hormonal differences

The sex chromosome complement also determines gonadal specification and the associated concentrations of sex steroids. Studies of castrated or ovariectomized rodents have clearly established that gonadal hormones are the major regulators of sex differences in immune responses<sup>25,26</sup>. Receptors for oestrogens, P4 and androgens are expressed in innate and adaptive immune cells, and signalling through these receptors influences both the development and function of immune cell populations. The effects of sex steroids on immune function are dependent on the concentration of the specific hormone and the cellular expression of the cognate hormone receptor (HR).

Canonical sex steroid signalling involves the binding of the hormone to its nuclear HR in the cytoplasm, receptor dimerization and the nuclear translocation of the liganded HR. The liganded HR then modulates the expression of target genes through direct binding to

hormone response elements, or through recruitment of transcriptional activators or co-repressor proteins to target genes<sup>27</sup>. Alternatively, ligand-bound HRs can indirectly modulate transcription of target genes by associating with and inhibiting the activity of other transcription factors such as NFκB or AP1 (ref. 27). Unliganded HRs can also modulate target gene expression through the formation of complexes with other proteins at gene promoter regions<sup>28,29</sup>. Beyond these genomic actions of hormones, sex steroids can influence cell function by binding to membrane-tethered nuclear HRs or other membrane-localized HRs, which triggers rapid changes in gene expression through intracellular signalling cascades involving calcium mobilization, generation of cyclic AMP, modulation of ion currents or activation of protein kinases<sup>29</sup>.

17β-Oestradiol (E2) circulates at greater levels in females and has a role in the proliferation and differentiation of many cell types, including immune cells<sup>30</sup>. E2 affects immune function by signalling through the nuclear HRs oestrogen receptor-α (ERα) (encoded by *ESR1*) and ERβ (encoded by *ESR2*), and through the G protein-coupled oestrogen receptor 1 (GPER1)<sup>30,31</sup>. ERα, ERβ and GPER1 are expressed in haematopoietic stem cells and in mature immune cells, albeit the expression level of HRs varies according to the cell type examined<sup>30–32</sup>. A comparison of *ESR1* and *ESR2* expression in human peripheral blood mononuclear cell populations illustrates that expression of both receptors is greater in lymphocytes than in NK cells or monocytes, and that CD4<sup>+</sup> T cells express greater levels of *ESR1*, whereas B cells express greater levels of *ESR2* (ref. 32). Despite the higher circulating levels of E2 in females, *ESR1* and *ESR2* are not differentially expressed in T or B cells between the sexes<sup>32</sup>. The expression of *ESR1* and *ESR2* also does not fluctuate during the menstrual cycle in females<sup>33</sup>. *ESR1* and *ESR2* expression is relatively stable in CD4<sup>+</sup> T cells and does not change in response to stimulation<sup>32</sup>.

Oestrogens can have both enhancing and suppressive effects on immune responses depending on the concentration of the ligand, the receptor and cell type<sup>2</sup>. E2 can transcriptionally regulate genes that have oestrogen response elements (EREs), including activation of *IFNG*<sup>34</sup> and repression of *IL6* (ref. 35) and *TNF*<sup>28</sup>. The tolerogenic effects of E2 are mediated through ERα and GPER1 (refs. 36,37). For example, E2 acting through ERα expands FoxP3<sup>+</sup> regulatory T (T<sub>reg</sub>) cells, induces expression of PD1 on T<sub>reg</sub> cells, induces thymus involution and promotes thymocyte apoptosis<sup>36,37</sup>.

Another steroid that can circulate at higher levels in females is P4, which acts as a ligand for two nuclear P4 receptors (nPR) – nPRA and nPRB – that are encoded by a single gene, *PGR*<sup>38</sup>, as well as membrane PR<sup>38</sup>. Natural and synthetic forms of P4 also signal through other steroid receptors including the androgen, glucocorticoid and mineralocorticoid receptors<sup>39,40</sup>. PRs are expressed in many human immune cells, including NK cells, macrophages, dendritic cells (DCs) and T cells. Serum P4 levels correlate with the percentage of pDCs in human peripheral blood positive for interferon (IFN)-α<sup>41</sup> and the induction of antiviral genes<sup>42</sup>. P4 promotes a tolerogenic state, including the production of tissue repair cytokines (including TGFβ, IL-6 and IL-22), activation CD39<sup>+</sup> T<sub>H</sub>17 cells and production of epidermal growth factor amphiregulin (AREG) by lung epithelial cells during IAV infection in mice<sup>43</sup>. Physiological concentrations of P4 have been shown to protect the ovaries from excessive inflammation during ovulation<sup>44</sup>. Treatment of bone marrow-derived DCs with P4 at physiological concentrations suppresses lipopolysaccharide-induced secretion of TNF and IL-1β and expression of CD80, and reduces the capacity of bone marrow-derived DCs to activate allogenic T cells in vitro<sup>45</sup>. In murine splenic B cells, physiological P4 concentrations decrease the mRNA expression of

activation-induced deaminase (AID), an enzyme involved in class switch recombination (CSR) and affinity maturation of antibodies<sup>46</sup>.

Androgens, including testosterone and the more potent non-aromatizable androgen, dihydrotestosterone (DHT), bind the androgen receptor (AR), which resides in the cytoplasm and translocates to the nucleus when engaged by ligand. AR is expressed in several types of immune cell, including thymocytes, peripheral T cells, B cells, neutrophils, monocytes, macrophages and type 2 lymphoid cells as well as stromal cells<sup>47,48</sup>. Collectively, androgens can exert suppressive or inhibitory influence on the activity of immune cell populations and effector immune responses, which contributes to sex-specific differences in clinical disease states, including autoimmunity and cancer<sup>47</sup>.

## Sex differences in immune responses

### Innate immunity

The innate immune system comprises innate T lymphocytes, innate lymphoid cells (ILCs), granulocytes, macrophages and DCs. These cells are the first responders during infection and sense the presence of infected cells through germline-encoded receptors (for example, TLRs) and cytokine receptors. Sex-specific differences in components of the innate immune system can affect the early response to pathogens, as well as the strength and nature of the subsequent adaptive immune response.

**Type I IFN signalling.** One of the most well-studied examples of sex differences in the innate system is the greater production of type I IFN by female plasmacytoid DCs (pDCs) than by male pDCs in humans in response to TLR7 activation<sup>41,49–52</sup>. This sex difference in the TLR7-induced type I IFN response extends to all IFN $\alpha$  isoforms and to IFN $\beta$ , and is not attenuated by inhibition of type I IFN receptor signalling, suggesting that it is specified by signalling events that occur downstream of TLR7 and upstream of the first signalling loop of type I IFN transcription<sup>50</sup>. In this regard, human female pDCs express higher levels of TLR7 (ref. 53) and IRF5 (ref. 54), which is a transcription factor that signals downstream of TLR7 to induce type I IFNs. Compared with pDCs of male mice, pDCs of female mice also express higher levels of *Unc93b1*, which can enhance TLR7 signalling by regulating the trafficking and stability of TLR7 in endosomes<sup>55</sup>.

Studies from the past few years have advanced our understanding of the genetic and hormonal mechanisms that underlie sex differences in TLR7-induced IFN $\alpha$  production by human pDCs (Fig. 1). One study that compared pre-pubertal and post-pubertal females (XX), transgender XX and XY individuals receiving hormone therapy, and females with Turner syndrome (XO) found that sex differences in TLR7-induced IFN $\alpha$  production by pDCs are already present in pre-pubertal children but are amplified after puberty<sup>53</sup>. In addition, this study showed that TLR7 expression increases with increasing XX chromosome gene dosage<sup>53</sup>. Studies conducted in immunodeficient mice transplanted with human male (XY) or female (XX) progenitor cells purified from cord blood cells further demonstrated that TLR7-induced IFN $\alpha$  expression by human pDCs is greatest when cells have an XX chromosome complement and are developed in a female-typic host environment<sup>56</sup>. A notable fraction (~30%) of female human pDCs exhibit bi-allelic X-expression of *TLR7*, which is associated with greater *TLR7* and IFN $\alpha$  expression at the single-cell level compared with that of mono-allelic female or male pDCs, suggesting that TLR7 is an X chromosome-encoded gene that contributes to the enhanced IFN $\alpha$  response by human female pDCs<sup>14,57</sup>. Nonetheless, the observation that *TLR7* is not differentially expressed

in pDCs prior to puberty<sup>53</sup> suggests that another factor accounts the greater type I IFN signalling seen in female than in male infants and young children<sup>52,53</sup>.

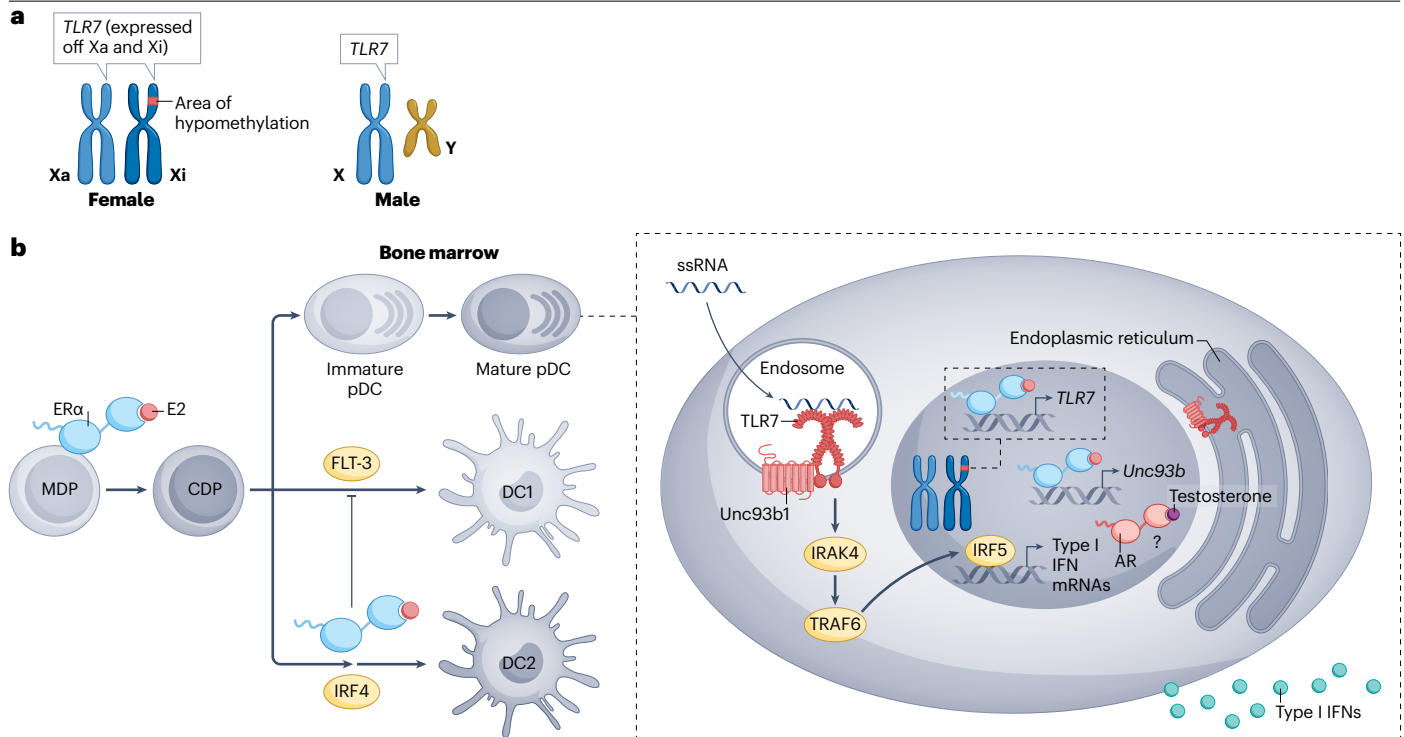
E2 also enhances TLR7-induced type I IFN signalling, as exemplified by the finding that supplementation of exogenous E2 to post-menopausal females increases the frequency of pDCs that produce IFN $\alpha$  upon TLR7 stimulation<sup>51</sup>. Furthermore, mice lacking *Esr1* exhibit impaired IFN $\alpha$  production by pDCs<sup>51</sup>. However, stimulation of pDC IFN $\alpha$  production is not observed when pDCs are treated with E2 in vitro<sup>49,51</sup>, suggesting that the effect of E2 is hard-wired during the development of pDCs in the bone marrow. The expressions of *Unc3b1* (ref. 58) and *IRF5* (ref. 54) are both induced by E2, and the expression of *IRF5* correlates with that of *ESR1* in human pDCs<sup>54</sup>, suggesting that *IRF5* may mediate the effect of E2 on pDC IFN $\alpha$  production.

Analyses of whole-blood cytokine levels in response to TLR and STING activators revealed that the TLR9-induced IFN $\alpha$  response is also higher in human female pDCs than in male pDCs<sup>59</sup>, contrasting with previous studies that found no sex differences in TLR9 signalling<sup>41,49,53</sup>. In further support of a female bias in the TLR9-induced IFN $\alpha$  response, in vivo E2 treatment potentiates IFN $\alpha$  and TNF production by human pDCs in response to stimulation with a synthetic TLR9 ligand or with serum from patients with systemic lupus erythematosus (SLE) that contains self-nucleic-acid-containing immune complexes<sup>51</sup>. In addition, pDCs from mice deficient in *Esr1* exhibit reduced IFN $\alpha$  production in response to stimulation with either TLR7 or TLR9 ligands<sup>51</sup>.

Unbiased analyses of sex differences in the transcriptomic or epigenomic landscape of human and murine immune cells have revealed a greater expression of genes in the canonical IFN stimulation pathway in females than in males, particularly in cells of the monocyte lineage<sup>12,60</sup>. For example, an analysis of sex-based differentially expressed genes (DEGs) in the DICE (Database of Immune Cell Expression, Expression quantitative trait loci (eQTLs) and Epigenomics) study identified a more prominent IFN signature in human females than in males across every immune cell examined<sup>12</sup>. Type I IFN, type II IFN, IFN regulatory factors and TLRs were predicted to be upstream regulators of these female-enriched DEGs<sup>12</sup>. The Immunological Genome Project (ImmGen) examined the transcriptome of 11 different murine cell types and identified sex-based DEGs only in macrophages, with genes in IFN-stimulated and antiviral immune pathways being particularly enriched in cells from female mice<sup>60</sup>. Together, these findings highlight that IFN transcriptional activity in immune cells is greater in females in the steady state, which may render females poised for more rapid antiviral immune responses when exposed to viruses or vaccines.

**Antigen presentation.** Sex-specific functional differences exist in the phenotype of antigen presenting cells (APCs) that extend beyond type I IFN signalling. For example, greater allorecognition of mismatched MHCII is observed in mixed lymphocyte reactions that contain female rather than male mouse spleen cells; this effect is dampened by androgen treatment of female cells<sup>25</sup>. Studies that have profiled APCs in male and female mice in the context of vaccination have not identified sex differences in antigen processing or in the expressions of MHC class II, CD80 or CD86; however, a common finding in these studies is that draining lymph node cells from female mice produce greater levels of the cytokine IL-12p40 (refs. 61,62). IL-12p40 can be secreted by DCs either as a monomer, a homodimer or as a heterodimer with IL-12p35 to form the T<sub>H</sub>1-promoting cytokine IL-12p70, or with IL-23p19 to form the T<sub>H</sub>17-promoting cytokine IL-23 (ref. 63). Macrophages or DCs isolated from female mice secrete greater levels of IL-12p40 than cells





**Fig. 1 | Mechanisms underlying sex differences in the production of type I interferon.** **a**, Females have two copies of the X chromosome (blue), one that is inherited from the female parent (Xm) and one from the male (Xp). Males have a Y chromosome (yellow) and inherit Xm. In male offspring, the X chromosome is active (Xa, light blue), whereas in females, one X chromosome (either Xm or Xp) becomes randomly inactivated and is hypermethylated (Xi, dark blue) to compensate for X-encoded gene dosage effects. Biological females are a mosaic of cells that express genes from Xp and Xm. *TLR7* is expressed on the X chromosome and is a gene that can partially escape X-inactivation on Xi; some female plasmacytoid dendritic cells (pDCs) and B cells show bi-allelic expression of *TLR7*. **b**, Various mechanisms may account for the greater production of type I interferon (IFN) in pDCs from human females in response to endosomal TLR7 stimulation. The sex difference in the IFN $\alpha$  response is encoded as pDCs develop in the bone marrow and is regulated by both X chromosome gene dosage and gonadal hormone levels. The effect of X chromosome gene dosage relates

in part to the bi-allelic expression of *TLR7*. The effect of X chromosome gene dosage is stronger in individuals with lower testosterone levels (shown by the purple circle and potential interactions with androgen receptor (AR) signalling); although the mechanisms of this regulation are unknown. Oestradiol (E2, red circle) acts through oestrogen receptor- $\alpha$  (ER $\alpha$ ) to enhance TLR7-induced type I IFN production via a number of mechanisms, including the increased expression of *TLR7* and IRF5. In mice, E2 increases the expression of Unc93b1, which helps shuttle TLR7 from the endoplasmic reticulum to the endosomes. E2-ER $\alpha$  signalling also acts on DC progenitors in the bone marrow to enhance myeloid DC generation via upregulation of IRF4; E2 also inhibits the survival of DCs differentiated with FLT3 ligand, albeit the generated DC1 cells are more mature. Collectively these mechanisms contribute to sex differences in antiviral immunity and autoimmune diseases, including systemic lupus erythematosus. CDP, common dendritic progenitor; MDP, monocyte dendritic cell progenitor; ssRNA, single-stranded RNA.

from male mice when co-cultured with activated T cells; the greater secretion by female cells is associated with higher IFN $\gamma$  secretion by T cells in co-cultures<sup>64,65</sup>. This sex difference in IL-12p40 production by macrophages is not present when macrophages are sourced from castrated male mice<sup>64</sup>, suggesting that androgens limit IL-12p40 secretion. In line with this proposal, treatment of mice with E2 enhances, whereas ovariectomy lowers, the production of IL-12p40 and IL-12p70 by spleen cells after stimulation with *Staphylococcus aureus*<sup>66</sup>. Thus, female murine DCs exhibit greater secretion of IL-12p40 and IL-12p70 in response to T cell-derived signals, which may serve to amplify the sex difference in the T<sub>H</sub>1 response (discussed below).

ER $\alpha$  signalling is also crucial in the DC differentiation pathway<sup>67</sup>. During homeostatic Flt3-ligand-driven DC development, E2 signalling through ER $\alpha$  enhances the expression of MHC class II and co-stimulatory molecules and the potential of the generated DC to produce IL-6, IL-12p40 and IL-12p70 in response to T cell-derived

stimuli<sup>67</sup>. Under conditions of inflammation, E2-ER $\alpha$  signalling induces granulocyte-macrophage colony-stimulating factor (GM-CSF)-stimulated myeloid progenitors to express IRF4 and to differentiate into inflammatory DCs<sup>67</sup>. Inflammatory DCs that are differentiated in the presence of GM-CSF and E2 are more mature than those differentiated without E2, and are more effective at priming T cells<sup>68</sup>. Thus, the differential concentrations of sex steroids modulate APC maturation and function to regulate T cell priming.

**Macrophage polarization and function.** One study found that the activation and phagocytic activity of certain macrophage and monocyte populations is greater in female mice than in male mice<sup>2,69</sup>. Characterization of immune cells in the peritoneal cavity of rats and mice revealed that macrophages and lymphocytes are more abundant in females than in males, correlating with greater chemokine expression by peritoneal cells in the females<sup>69</sup>. Peritoneal macrophages from female mice also



demonstrate greater basal expression of TLR2, TLR3, TLR4 and MyD88, enhanced phagocytosis, and better NADPH oxidase-mediated bacterial killing activities than macrophages from males, which correlates with greater resistance against group B streptococcal (GBS)-induced sepsis<sup>69</sup>. However, a more recent study that involved population-level and single-cell RNA sequencing of murine peritoneal macrophages under homeostatic conditions found no significant sex difference in TLR or MyD88 transcripts in the dominant F4/80<sup>hi</sup>MHCII<sup>+</sup>CD102<sup>+</sup> peritoneal macrophage population<sup>70</sup>. Rather, that study found that macrophages from female mice exhibit a higher expression of genes associated with lipid uptake and transport, phagocytosis (for example, *CD209b*) and recruitment of B1 cells, whereas male macrophages show a higher expression of genes associated with cell proliferation<sup>70</sup>. These differences in gene expression were attributed to the greater representation of Timd4<sup>+</sup> macrophages within the peritoneal macrophage pool in the females. Timd4<sup>+</sup> peritoneal macrophages are derived from an embryonic monocyte precursor and are the pioneer population of macrophages in the peritoneum. After sexual maturation, Timd4<sup>+</sup> macrophages are gradually replaced by bone marrow-derived Timd4<sup>-</sup>Ly6C<sup>hi</sup> monocytes, with this turnover being more rapid in the male mice<sup>70</sup>.

Macrophage polarization states have also been reported to vary between males and females in the heart and lung in the context of infection and allergy, contributing to sex differences in disease outcomes<sup>71–74</sup>. For example, coxsackievirus B3 (CVB3) infection causes more severe myocarditis in male than in female BALB/c mice<sup>75</sup>. Although age-matched male and female infected BALB/c mice show similar cardiac myocyte viral titres and numbers of infiltrating macrophages, the polarization of macrophages differs between the sexes, with greater IFN $\gamma$  and pro-inflammatory M1 polarization in males and greater IL-4 and M2 phenotypic differentiation in females<sup>72,75</sup>.

A similar sex difference in macrophage polarization was demonstrated in mice with GBS-induced pneumonia whereby male mice show a stronger early innate immune response to infection than female mice with a greater accumulation of monocytes and M1-like macrophages; this occurs despite the males and females having an equivalent bacterial burden<sup>73</sup>. That study also implicates the X-linked miRNA, miR-223-3p, in regulating sex differences in macrophage polarization<sup>73</sup>. Higher expression of miRNA 223-3p in female mice correlated positively with the proportion of M2 macrophages and negatively with the proportion of M1 macrophages<sup>73</sup>. In line with these preclinical findings, the incidence of GBS bloodstream infections is higher among older human males than among older females, and infant males have a higher incidence of GBS sepsis than infant females<sup>76,77</sup>. Sex differences in macrophage polarization were also seen in the lung in a murine model of ovalbumin-induced airway inflammation<sup>74</sup>. In contrast to male mice, which demonstrate M1 polarization, female mice develop lung inflammation characterized by the presence of T<sub>H</sub>2 effector cells, correlating with a greater predominance of M2-like (Ym1<sup>+</sup>CD206<sup>+</sup>) macrophages<sup>74</sup>.

Whether sex differences in macrophage polarization are driven by intrinsic sex differences in macrophages, differences in the developmental origin of these macrophages, or differences in the early cytokine cues provided by ILCs in the tissue is unknown. During murine CVB3 infection, sex differences in the extent of heart damage are associated with differential activities of  $\gamma\delta$ T cells<sup>75</sup>. Depletion of V $\gamma$ 4T cells, which secrete more IFN $\gamma$  (a cytokine cue for M1-like macrophages) or administration of exogenous M2 macrophages protects males from myocarditis<sup>72,75</sup>. Finally, single-cell sequencing studies of tissue macrophages have concluded that the gene profile of macrophages aligns more closely to the developmental origin of the cell than to M1 or M2

phenotype<sup>78</sup>. Future studies using single-cell approaches are needed to re-examine the phenotype of male and female macrophage populations in various tissues under steady state and disease conditions to better understand the basis of these sex differences in macrophage phenotype.

**Granulocytes.** Neutrophils comprise the greatest portion of circulating human blood cells and have a central role in the first line of innate immune defence. Studies that have profiled isolated murine neutrophils have detected sex-based differences in their gene expression. For example, The ImmGen study found that neutrophils from female mice exhibit a higher expression of IFN-stimulated genes at steady state than those from male mice<sup>60</sup>. In addition, a multiomics assessment that included transcriptomics, metabolomics and lipidomics investigations found sex differences in the genes, metabolites and lipids of primary bone marrow-derived neutrophils. Moreover, ageing has a greater effect on neutrophils from male mice than on those from female mice<sup>79</sup>. Female-biased genes include collagen-encoding genes and receptors that regulate neutrophil migration across endothelial barriers. By contrast, neutrophils from males have a greater expression of cell cycle-related genes and a gene signature consistent with more condensed or repressed chromatin. This study also found that neutrophils from male mice exhibit a gene signature suggestive of greater elastase release and increased serum neutrophil elastase at baseline and in response to lipopolysaccharide injection<sup>79</sup>. Further studies are required to better understand how transcriptional differences in neutrophils translate to phenotypic sex differences in immune function.

Studies have also implicated a critical role for the AR in the differentiation and function of neutrophils<sup>80,81</sup>. Mice lacking the AR exhibit neutropenia owing to the reduced proliferation of neutrophil precursor cells<sup>80</sup>. Although neutrophils from these AR-knockout mice exhibit normal phagocytosis and oxidative burst activities, their capacity for chemokine and cytokine production and CXCR2-mediated migration is diminished compared with that of neutrophils from wild-type mice<sup>80</sup>. Castration of male mice also decreases neutrophil number and impacts neutrophil function in the context of tumour metastasis<sup>81</sup>. In a mouse model of B16 melanoma, males exhibit a lower metastatic tumour burden in the lung than female mice, which correlates with an increased number of neutrophils and greater neutrophil-dependent cytotoxicity in the males<sup>81</sup>. Neutrophil depletion in male mice or castration increases tumour burden, whereas testosterone treatment reduces tumour burden in female mice<sup>81</sup>. Consistent with these findings, neutrophils isolated from patients with prostate cancer undergoing androgen deprivation therapy have a less mature, myeloid-derived suppressor cell-like phenotype compared with patients not undergoing treatment<sup>81</sup>. Thus, strong evidence supports a role for androgens in the control of neutrophil differentiation and function.

**Innate lymphoid cells and natural killer T cells.** ILCs sense the presence of virus-infected cells, respond to damage through cytokine receptors (such as the IL-12 or IL-33 receptor), and generate effector cytokines (for example, IFN $\gamma$  and IL-5) to direct adaptive immune responses. ILCs encompass five cell subsets, including type 1, type 2 and type 3 ILCs (ILC1–3), NK cells and lymphoid tissue inducer cells. The cytokine signatures of ILC1, ILC2 and ILC3 subsets resemble those of T<sub>H</sub>1, T<sub>H</sub>2 and T<sub>H</sub>17 cells, respectively<sup>29</sup>. Of these subtypes, sex differences have been reported for NK cells, ILC2 and ILC3 cells.

NK cells have been detected in the peripheral blood at higher frequencies in males than in females in both mouse and human studies<sup>17,82</sup>;

however, activated female NK cells secrete higher concentrations of IFN $\gamma$ , perforin and granzyme B than male NK cells<sup>17,82</sup>. One study found that sex differences in neutrophil number and function are a result of the higher expression of the X-escapee gene *Kdm6a* (which encodes UTX) in female than in male NK cells<sup>17</sup>. Normalizing the sex difference in *Kdm6a* copy number through the generation of female UTX heterozygous mice, normalizes sex differences in NK cell numbers and responses<sup>17</sup>. Decreasing UTX copy number increases NK cell numbers by increasing genes in cell fitness; this is also associated with reduced expression of genes associated with NK effector function (such as *ffng*)<sup>17</sup>. Although gonadectomy in this study had no effect on NK cell number, other studies have found that NK cell number and cytotoxic activity fluctuate with the menstrual or oestrus cycle<sup>83,84</sup>. Thus, both X chromosome complement and sex steroids affect NK cell numbers and activity.

In contrast to NK cells, sex differences in ILC2 and ILC3 cells are controlled by androgens<sup>29,85</sup>. Androgens limit the development of ILC2 cells in the bone marrow, which is associated with a steep decline in these cells in male humans and mice after puberty<sup>86,87</sup>. Studies in mice illustrate that androgens also drive sex differences in the phenotype of ILC2s. Compared with those from male mice, ILC2s from female mice demonstrate higher secretion of type 2 cytokines (for example, IL-5 and IL-13) in response to IL-2 and IL-33 and reduced expression of killer cell lectin-like receptor 1 by these cells<sup>29</sup>. In models of allergic inflammation, sex differences in ILC2 activity in the lung drive the type 2 inflammatory response, which is more notable in females than in males<sup>88,89</sup>. Androgens limit the expression of IL-33 in the lung, which is important for sustaining the proliferation and phenotype of ILC2s in this tissue, resulting in higher numbers of ILC2 cells in the lungs of females at steady state<sup>29,87,88</sup>. By suppressing IL-33 secretion, androgens also increase the stability of T<sub>reg</sub> cells, reduce eosinophil infiltration and decrease T<sub>H2</sub> cell-mediated inflammation in the lungs of males during allergic inflammation<sup>87–90</sup>. However, following immunization with myelin and complete Freund's adjuvant (CFA), male mice are more prone than female mice to T<sub>H2</sub> cell-mediated inflammation in the lymph node. This inflammatory response correlates with higher ILC2 responses in the lymph nodes of male mice than in female mice as a result of androgen-induced IL-33 production by mast cells<sup>91</sup>. Thus, the effect of androgens on ILC2-mediated inflammation is tissue-specific and context-dependent.

The effect of androgens on ILC3s has been demonstrated in a murine model of *Escherichia coli* bladder infection<sup>85</sup>. Female mice demonstrate better protection than male mice against infection; however, this sex difference is eliminated when females are treated with testosterone<sup>85</sup>. Increased protection against *E. coli* in females correlates with greater IL-17 production and greater numbers of ILC3s and  $\gamma\delta$ T cells in the bladder early during infection. These effects are eliminated by neutralizing IL-17.

NK T cells (NKT) are innate T cells that express semi-invariant TCR chains (whereby V $\alpha$ 14J $\alpha$ 18 pairs with V $\beta$ 8, V $\beta$ 7 or V $\beta$ 2 in mice, and V $\alpha$ 24J $\alpha$ 18 pairs with V $\beta$ 11 in humans) and rapidly produce cytokines in response to glycolipid ligands presented by CD1d<sup>92</sup>. The CD4<sup>+</sup>CD8<sup>-</sup>NKT cell subset is present in the peripheral blood at a higher frequency in female than in male humans<sup>93,94</sup>. Stimulation of NKT cells from human females with the synthetic ligand  $\alpha$ -galactosylceramide ( $\alpha$ -GalCer) in vitro elicits greater levels of effector cytokines (for example, TNF, IL-17, IFN $\gamma$  and IL-4) than stimulation of NKT cells from males<sup>94</sup>. Similarly, treatment of female mice with  $\alpha$ -GalCer increases, whereas ovariectomy or deletion of the ER $\alpha$  receptor decreases, serum

concentrations of IFN $\gamma$ <sup>95</sup>, suggesting that E2 promotes NKT cell function. In line with this suggestion, androgens inhibit the production of IFN $\gamma$  by murine NKT cells in the liver, which limits the capacity of male mice to clear infection with the protozoan *Entamoeba histolytica*<sup>96,97</sup>.

## Adaptive immunity

Adaptive immunity refers to both the cellular (T cell-mediated) and humoral (antibody-mediated) arms of the immune system that function to clear pathogens and establish memory to protect against future infections. Strong evidence suggests that females exhibit stronger T cell and humoral immune responses than males.

**T cell-mediated immune responses.** A number of studies have demonstrated that the expansion of T cells in response to an immune challenge is more extensive in females than in males (Fig. 2). For example, antigen-specific T cells from female, young adult mice treated with a protein antigen and CFA, proliferate more and secrete higher levels of IFN $\gamma$  upon re-exposure to the antigen *ex vivo* than cells from similarly treated male mice<sup>25,98,99</sup>. Stronger T cell responses are also observed in post-pubertal female mice than in post-pubertal male mice that are vaccinated with malarial sporozoites<sup>100</sup> or infected with HIV<sup>101</sup>. Likewise, human females develop stronger antigen-specific T cell responses to hepatitis C virus (HCV), SARS-CoV-2 and measles vaccination<sup>102–104</sup>.

Differences in the roles of sex hormones and sex chromosomes in the T cell response to vaccination have been investigated using the FCG mouse model<sup>98</sup>. Following vaccination with a model antigen and CFA, gonadal females exhibit stronger antigen-specific T cell proliferative responses and greater IFN $\gamma$  production than gonadal males regardless of the underlying sex chromosome complement<sup>98</sup>. This sex difference is blunted by castrating males, but not by ovariectomizing females<sup>98</sup>, suggesting that androgens are a major driver of the sex difference in the expansion and/or differentiation of T<sub>H1</sub> cells.

Female humans and mice have greater frequencies of circulating CD4<sup>+</sup> T cells than do males<sup>105,106</sup>. Compared with young adult males, a greater fraction of circulating CD4<sup>+</sup> T cells in young adult females have a naive or recent thymic emigrant status<sup>107–109</sup>, which may relate to the comparatively slower decline in the rate of thymic involution in females after puberty<sup>109</sup>. Sex differences in androgen levels seem to underlie the sex differences in the thymic output of T cells, at least in rodents, by limiting thymocyte development through effects on AR in thymic stromal cells<sup>110</sup>. Both castration and ovariectomy increase thymic size and cellularity, with castration having more potent effects than ovariectomy in restoring thymic function<sup>111,112</sup>.

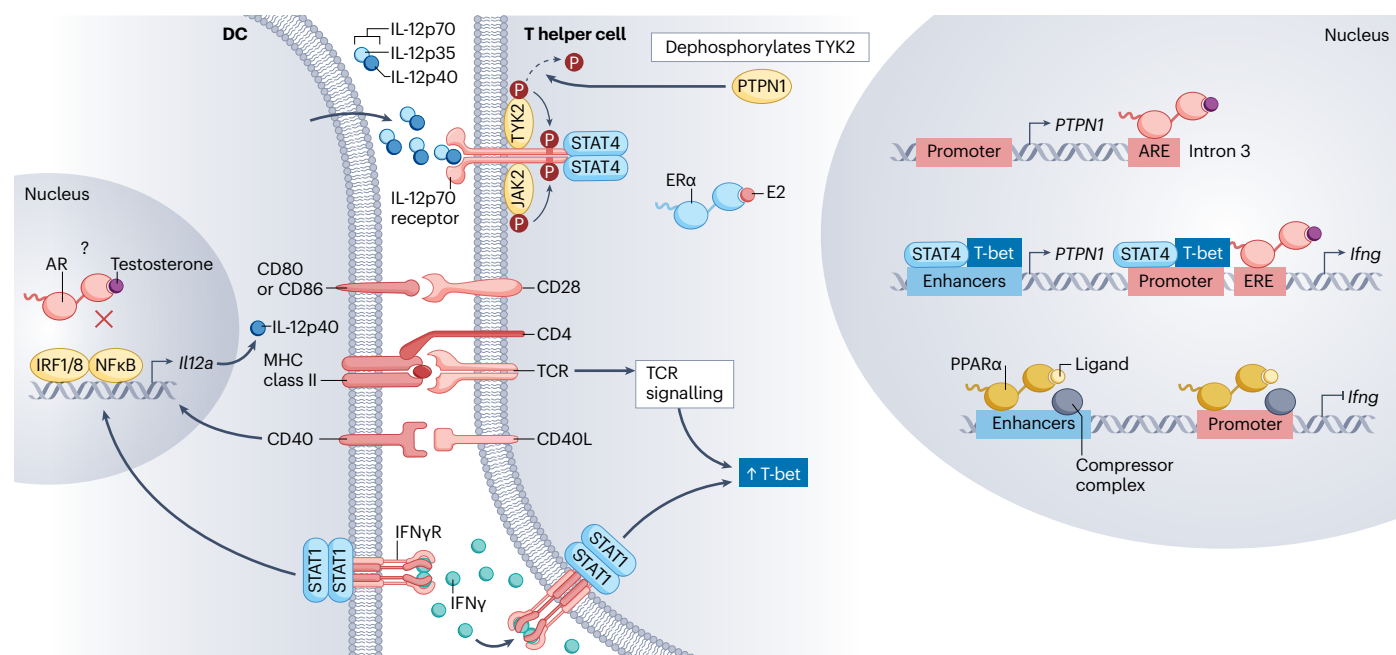
Differential expression of *Aire*, which encodes an autoimmune regulator, may contribute to sex differences in thymic function and T cell immunity. *Aire* is expressed in medullary thymic epithelial cells and induces the expression of tissue-specific antigens that mediate the negative selection of autoreactive thymocytes<sup>113</sup>. *Aire* is expressed at higher levels in thymic epithelial cells from young adult males than in those from females through its induction by androgens and repression by E2 (refs. 114,115). *Aire*-knockout mice with EAE are resistant to the protective effect of androgens<sup>114</sup>, suggesting that the effects of androgens on T cell immunity in this model of multiple sclerosis may be mediated, in part, by increasing *Aire* expression. However, the interpretation of these findings is clouded by the fact that *Aire*-knockout mice also demonstrate enhanced T<sub>H17</sub> responses and develop more severe EAE than wild-type mice<sup>114</sup>. Moreover, although it is established that thymus size and *Aire* expression are under the control of sex steroids in young adult mice, this does not seem to be the case in middle-aged mice<sup>116,117</sup>.

Thus, the role of *Aire* as a regulator of sex differences in autoimmunity remains controversial.

In the steady state, T cells in the peripheral blood and gut mucosa of human females are more activated than those from males, with female T cells exhibiting greater expression of HLA-DR, FoxP3, CD69 and Ki67 (refs. 118,119). When cultured in vitro with anti-CD3 and anti-CD28 or phytohaemagglutinin, female murine and human CD4<sup>+</sup> and CD8<sup>+</sup> T cells produce higher levels of IFN $\gamma$  than cells from males<sup>82,120</sup>. The mechanisms of the sex-dependent regulation of T cell IFN $\gamma$  production are well understood (Fig. 2). Androgens negatively regulate IFN $\gamma$  production by altering chromatin accessibility at the *Ifng* locus<sup>65,82,121</sup>, which occurs through the activation of AR on control elements in this gene<sup>121</sup> or indirectly through the androgen-induced expression of peroxisome proliferator-activated receptor- $\alpha$  (PPAR $\alpha$ ) which acts as a repressor at the IFN $\gamma$  gene locus<sup>65,82,122</sup>. Knockdown of PPAR $\alpha$  or treatment with a PPAR $\alpha$  antagonist enhances IFN $\gamma$  production by CD4<sup>+</sup> and CD8<sup>+</sup> T cells from male, but not female, mice, and also abrogates sex differences in murine survival following infection with *Listeria monocytogenes*<sup>82</sup>. Beyond androgens, available evidence suggests that E2 can enhance IFN $\gamma$  production by T cells. E2 treatment

potentiates IFN $\gamma$  production by T cells in vivo<sup>66,123</sup> and treatment of activated splenocytes in vitro with doses of E2 that mimic levels seen in the menstrual cycle (0.01–1 nM E2) enhances *Ifng* expression and promoter activity in T cells<sup>34</sup>. Sex steroids also regulate T<sub>H</sub>1 differentiation by modulating IL-12 receptor–STAT4 signalling<sup>124,125</sup>. E2 enhances murine IL-12-induced STAT4 phosphorylation<sup>124</sup>, whereas androgens reduce human IL-12-induced STAT4 phosphorylation by promoting expression of the tyrosine protein phosphatase non-receptor type 1 (ref. 126).

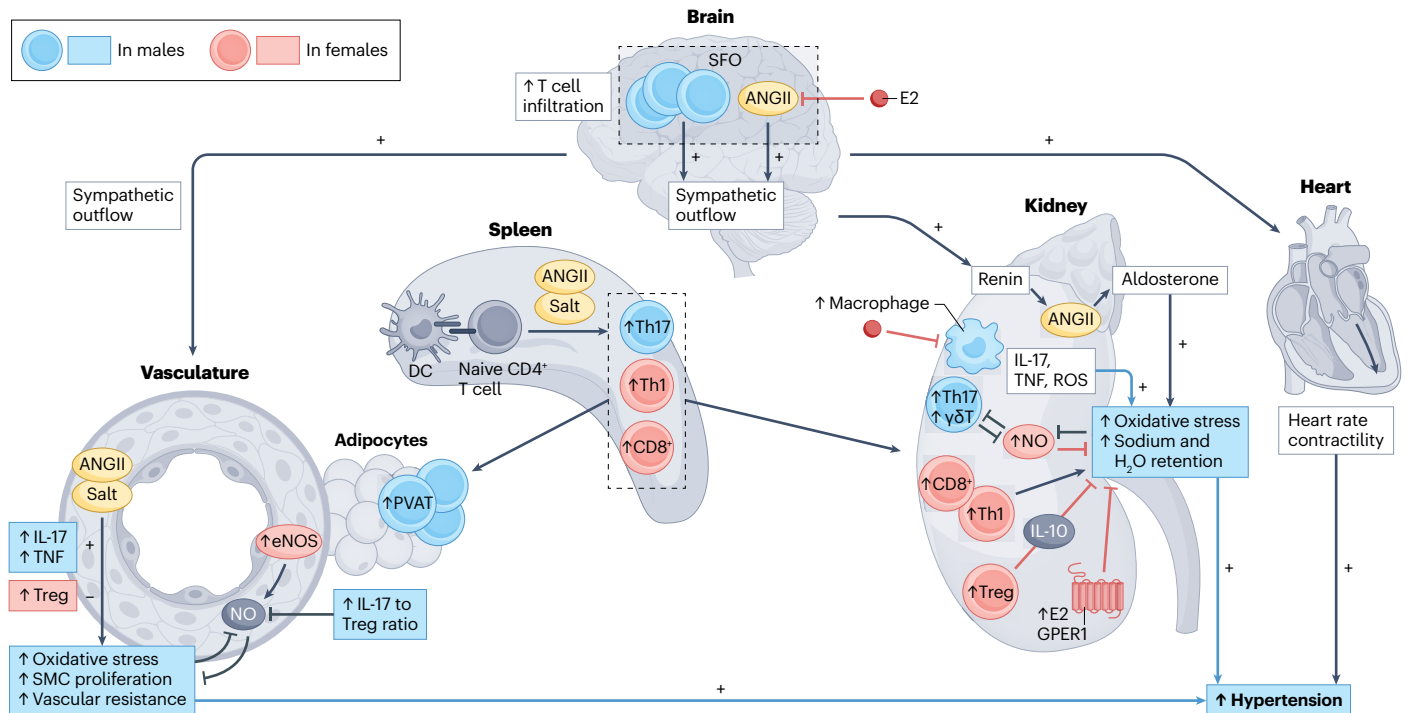
Female T<sub>H</sub> cells are more prone than male T<sub>H</sub> cells to produce IFN $\gamma$ , whereas isolated male T cells produce more IL-4 (refs. 62,127) or IL-17A<sup>65,120</sup> upon anti-CD3 and anti-CD28 stimulation. De-repression of IFN $\gamma$  in male T cells by castration is associated with reduced IL-4 or IL-17 secretion by T cells<sup>65,127</sup>. Sex differences in the expression of the nuclear receptor, PPAR- $\gamma$  may contribute to sex bias in the production of IL-17 by T cells<sup>65,122</sup>. PPAR- $\gamma$  expression is repressed by androgens<sup>65</sup> and induced by E2 (ref. 122) in mice. Studies using T cell-specific knockout mice have illustrated that endogenous PPAR- $\gamma$  activity limits IL-17 production in female mice but not in male mice<sup>65</sup>. The male bias towards T<sub>H</sub>2 and T<sub>H</sub>17 is seen in some contexts in vivo (for example, in the kidney during hypertension, as discussed below), but not in others (for example,



**Fig. 2 | Sex differences in the production of IFN $\gamma$  by T cells.** T cells derived from female humans and mice intrinsically produce greater levels of IFN $\gamma$  than those from conspecific males, which is beneficial in the clearance of tumours and bacterial infections, but can be overcome by the activities of type 2 lymphoid cells in the context of allergic inflammation in the lung. This mechanism seems to be regulated by the balance in sex hormone levels. When naive CD4<sup>+</sup> T cells encounter antigen in the context of co-stimulation, signalling pathways get turned on (for example, p38, ERK and NF $\kappa$ B) that lead to early *Ifng* and *Tbx21* (encodes T-bet) expression. The early production of IFN $\gamma$  by T cells feeds back in an autocrine fashion through the IFN $\gamma$  receptor (IFN $\gamma$ R) and STAT1 to amplify T-bet expression through transcriptional mechanisms (not shown). T-bet induces expression of IL-12R  $\beta$ -chain 2 which sensitizes T cells to IL-12p70 and increases STAT4 activity. STAT4, T-bet (shown) and other transcription factors such as NF $\kappa$ B and AP1 (not shown) bind to the IFN $\gamma$  promoter and enhancer regions to induce *Ifng* expression. Androgens acting through the androgen receptor (AR) limit *Ifng*

expression by increasing expression of the nuclear receptor PPAR $\alpha$  in T cells, which tethers a co-repressor complex containing histone modifiers at this locus, limiting STAT4 and T-bet-dependent transcription. Androgens also induce the expression of tyrosine phosphatase non-receptor type 1 (PTPN1) in human T cells by binding of the AR to an intronic element in *PTPN1*. PTPN1 is a phosphatase that dephosphorylates TYK2 downstream of IL-12 signalling to limit STAT4 activity. By contrast, oestradiol (E2) enhances IL-12 receptor–STAT4 signalling through an unknown mechanism and binds to an oestrogen response element (ERE) in the IFN $\gamma$  promoter to induce *Ifng* expression. In antigen-presenting cells, such as dendritic cells, androgens limit the transcription of *Il12a* downstream of T cell-derived signals, including CD40 crosslinking and IFN $\gamma$  receptor signalling, although the exact mechanism is not understood. *Il12a* encodes IL-12p40, which dimerizes with IL-12p35 to make the pro-T<sub>H</sub>1 cytokine IL-12p70. ARE, androgen response element; DC, dendritic cell; ER $\alpha$ , oestrogen receptor- $\alpha$ ; PPAR $\alpha$ , peroxisome proliferator-activated receptor- $\alpha$ .





**Fig. 3 | Sex differences in the balance of IL-17-producing T cells and T regulatory cells regulate sex differences in experimental hypertension.**

Blood pressure (BP) is the product of blood volume (which is regulated by sodium and water secretion and retention by the kidney), cardiac output and vascular resistance and is positively regulated by sympathetic outflow to these organs. The sympathetic nervous system induces the production of renin from the kidney, which together with angiotensin-converting enzyme promotes the production of angiotensin II (ANGII). ANGII acts through the ANGII type 1 receptor in the renal cortex to induce the production of aldosterone, which promotes the reabsorption of sodium. ANGII also promotes vascular constriction and smooth muscle cell hypertrophy and increases heart rate and sympathetic outflow. T helper 17 (T<sub>H</sub>17) inflammatory responses, which predominate in males, enhance these mechanisms, leading to hypertension. ANGII and salt both promote the differentiation of naive CD4<sup>+</sup> T cells into T<sub>H</sub>17 cells, particularly in males. Moreover, pro-inflammatory, IL-17-expressing T<sub>H</sub>17 and  $\gamma\delta$ T cells infiltrate

the kidney and the peripheral vasculature adipose tissue (PVAT) to a greater extent in males. IL-17 and TNF promote inflammation and oxidative stress, leading to sodium and water retention in the kidney, smooth muscle cell (SMC) proliferation and inflammation in the vasculature, which increases vascular resistance. T<sub>H</sub>17 cells also inhibit the production of nitric oxide (NO) – a potent vasodilator and inhibitor of T cell proliferation. T cell infiltration is also greater in the subformal organ (SFO) of the brain in male mice during ANGII-induced hypertension, which may increase sympathetic outflow to the kidney and heart. Although pro-inflammatory T<sub>H</sub>1 and CD8<sup>+</sup> T cells are more abundant in females and can promote hypertension, the activities of these cells are tempered by the increased anti-inflammatory activities of T regulatory (T<sub>reg</sub>) cells and IL-10 in the kidney and vasculature. Estrogens also prevent hypertension by countering the functions of macrophages and microglia in the kidney and brain, respectively, and by activating GPER1, which can reduce oxidative stress in the kidney and promote sodium excretion.

in allergic inflammation in the lung<sup>128</sup> or bladder infection<sup>85</sup>) in which greater cytokine production by innate lymphocytes early in immune responses may overcome these intrinsic T cell differences (Fig. 3).

T follicular helper (T<sub>FH</sub>) cells are a subset of T<sub>H</sub> cells that are defined by high BCL6 and CXCR5 expression and promote germinal centre reactions. Although the numbers of T<sub>FH</sub> cells do not reportedly differ between the sexes at steady state<sup>122</sup>, the abundance of these cells increases to a greater extent in females than in males in rodent models of SLE and rheumatoid arthritis<sup>129–131</sup>. For example, female rats with collagen II and CFA-induced arthritis develop more severe arthritis than males, associated with a twofold higher antibody response and a fourfold higher number of T<sub>FH</sub> cells<sup>129</sup>. Similarly, in a murine model of SLE that is driven by allogeneic recognition of host MHC class II (H-2<sup>b/d</sup>) by donor H-2<sup>d</sup>-restricted T<sub>H</sub> cells, increased numbers of T<sub>FH</sub> cells in the lymph nodes is associated with more severe disease in females<sup>130</sup>. Increased T<sub>FH</sub> cell expansion and the development of SLE-like disease is also seen in female mice with DC-specific knockout of *Blimp1* – a key

negative regulator of Bcl6 and T<sub>FH</sub> differentiation. In that study, the sex difference in T<sub>FH</sub> cell expansion corresponded to a higher expression of IL-6 in female than in male DCs<sup>131</sup>. In sum, female T<sub>FH</sub> cells expand more than male T<sub>FH</sub> cells in the context of autoimmune responses; however, the underlying mechanisms of these sex differences require further exploration.

FoxP3<sup>+</sup> T<sub>reg</sub> cells are crucial for maintaining immune tolerance. The proportion of T<sub>reg</sub> cells in the blood is greater in young adult male humans than in young adult female humans<sup>132,133</sup>. Moreover, T<sub>reg</sub> cell frequencies in the blood are higher in cisgender males than in females even prior to puberty, with frequencies changing after puberty or sex hormone treatment in transgender individuals<sup>132</sup>. Transgender males (that is, individuals who have an XX chromosome complement receiving testosterone treatment) exhibit T<sub>reg</sub> cell frequencies similar to those in cisgender males, and transgender females (that is, individuals who have an XY chromosome complement receiving oestrogen treatment) have T<sub>reg</sub> cell frequencies similar to those in cisgender females<sup>132</sup>. These data



suggest that  $T_{reg}$  cell frequencies in the blood change with the levels of sex steroids. By contrast, the suppressive function of  $T_{reg}$  cells (being greater in males than in females) tracks with the sex of individuals at birth, and does not change with hormone therapy in transgender individuals<sup>132</sup>. In contrast to findings in young adults, sex differences were not observed in a study that investigated the frequency of blood  $T_{reg}$  cells in male and female human donors having a larger age range (18–75 years)<sup>108</sup>. It is possible that the large age range of participants in that study masked sex differences, since  $T_{reg}$  cell frequencies decline with age<sup>134</sup> (Box 1).

Whether  $T_{reg}$  cells are differentially abundant between male and female mice is controversial. Sex differences in  $T_{reg}$  cells have not generally been reported<sup>90,135,136</sup>, with the exception of inbred SJL mice,

in which they are reported to be greater in males<sup>137</sup>. However, as reported for humans,  $T_{reg}$  cells from male mice are more stable and have greater suppressive activity than those from female mice. Specifically,  $T_{reg}$  cells from wild-type males suppress T effector cell proliferation in vitro to a greater extent than  $T_{reg}$  cells from female or male AR-knockout mice<sup>90</sup>. In addition,  $T_{reg}$  cells from male SJL mice are reported to express higher levels of immunosuppressive molecules CTLA4 and IL-10 than those from female mice<sup>137</sup>.

$CD8^+$  T cells from female mice and humans expand more extensively than those of males and are more likely to acquire a short-lived T effector profile characterized by high expression of IFN $\gamma$ , TNF and granzyme B<sup>15,100,102,138</sup> and increased expression of markers of T cell memory<sup>15</sup>.  $CD8^+$  T cells from human females express more perforin per cell than those from males<sup>139</sup>. Studies in mice infected with *L. monocytogenes* or vaccinia virus have demonstrated that the effect of female sex in increasing the expansion of pathogen-specific  $CD8^+$  T cells is cell-extrinsic, meaning that sex differences are not present in the absence of sex differences in APC-derived signals. By contrast, sex differences in  $CD8^+$  T cell effector function are cell-intrinsic, with female  $CD8^+$  T cells demonstrating higher expression of IFN $\gamma$ , TNF and granzyme B and acquisition of a short-lived effector fate in response to an immune challenge as compared with male  $CD8^+$  T cells, which are instead skewed towards a memory precursor cell phenotype<sup>138</sup>. These sex differences in  $CD8^+$  T cell effector function are an important factor that contributes to sex differences in tumour progression in murine models<sup>121,140,141</sup>. Thus, female T cells expand to a greater extent in response to activating signals and are intrinsically geared to produce higher levels of IFN $\gamma$  and cytotoxic mediators than those of males; this difference is largely regulated by sex steroids.

**B cells and humoral immunity.** Adult females of diverse species exhibit stronger antibody responses than males in response to an immune challenge<sup>2</sup>. For example, in response to vaccination against simian immunodeficiency virus (SIV), female primates exhibit greater local IgA antibody titres, numbers of memory B cells, and numbers of plasma cells compared to males<sup>142,143</sup>. When SIV-vaccinated primates are challenged with SIV, females are better able to clear the infection compared to males<sup>142</sup>. Similarly, following vaccination with either whole inactivated IAV, trivalent inactivated vaccine, or quadrivalent inactivated vaccine, adult female mice generate greater influenza-specific antibody responses, including live virus neutralizing responses, than male mice<sup>15,144,145</sup>. Following vaccination with inactivated IAV, adult female mice produce greater virus-specific, class-switched total IgG and IgG2c antibodies against both the virus in the vaccine and variants that contain mutations in the haemagglutinin domain of the virus<sup>146</sup>. Antibodies from IAV-vaccinated female mice also exhibit broader reactivity, recognizing more unique, linear haemagglutinin epitopes than antibodies from males. Antibodies derived from vaccinated females better protect naive males and females than antibodies from males, and this protection is associated with higher antibody specificity and avidity to the virus<sup>15</sup>. After inactivated IAV vaccination, females also have higher numbers of germinal centre B cells containing superior somatic hypermutation (SHM) frequencies than those of males<sup>146</sup>. Deletion of the gene that encodes AID (*Aicda*), which mediates SHM in B cells, eliminates the sex-based differences in humoral immunity and protection against live virus challenge<sup>146</sup>. The expression of the X-linked gene, *Tlr7*, in B cells is also higher in IAV-vaccinated female mice than in IAV-vaccinated male mice, as a consequence of reduced DNA methylation in the *Tlr7* promoter region in female B cells. Sex-based

## Box 1

### Effect of ageing on immune cells in males versus females

Our appreciation of sex differences in the ageing of the immune system has increased remarkably with the emergence of ‘omics’ analyses. Older age is associated with a decline in the frequency of naive T cells and T regulatory cells, a reduction in the de novo production of antibodies and development of increased inflammatory immune responses; these changes were, in the past, incorrectly assumed to have the same kinetics and magnitude between older males and older females. However, several studies performed over the human life course have now demonstrated that indices of declining immunity, including genetic and epigenetic modifications in peripheral blood mononuclear cells (PBMCs), occurs earlier and with a greater magnitude in males than in females<sup>255</sup>. Among both mice and humans, ‘inflammageing’ is associated with an age-associated increase in the activity of inflammatory pathways. Single-cell transcriptional analyses have revealed that dendritic cells and monocytes of aged human men express significantly higher levels of inflammatory genes than those of aged women<sup>118</sup>. Multiomics analyses of primary bone marrow-derived neutrophils from young and aged mice did not reveal any sex difference on the impact of ageing on the abundance of bone marrow neutrophils; however, transcriptional changes associated with ageing were greater in neutrophils from male mice than in those from female mice<sup>79</sup>. By contrast, TLR7 induced IFN $\alpha$  secretion to a greater extent in whole human blood from females than in blood from males across diverse adult ages (19–97 years)<sup>59</sup>. ATAC sequencing of human PBMCs identified more regions of open chromatin at gene loci such as *JAK3* and *STAT5B* that are important for T cell signalling in aged females compared to aged males<sup>255</sup>. By contrast, single-cell RNA sequencing of PBMCs from young and aged individuals identified a striking upregulation of genes involved in TCR and MAPK signalling pathways, including *JUNB*, *NFKBIA*, *DUSP* and *CXCR4*, especially in male T cells with ageing, suggesting that male T cells acquire a greater inflammatory phenotype with age than female T cells<sup>118</sup>. Thus, collectively, across diverse species and immune cells, ageing seems to have a greater effect on cellular activity in males than in females, which is likely to impact sex and age-related differences in disease susceptibility.

differences in the neutralizing antibody response, CSR, and antibody avidity is eliminated in *Tlr7*-knockout mice<sup>15</sup>.

Human clinical trials that were conducted to identify public health strategies for managing influenza vaccine shortages revealed that adult females (aged 18–49 years) consistently generated haemagglutinin inhibition (HAI) antibody titres that were twice that of males in response to trivalent inactivated influenza vaccine (TIV), regardless of whether a full dose or half dose was administered<sup>147</sup>. In a trial in adults with a broader age range (20–89 years) females again generated greater neutralizing antibody titres following TIV compared than males<sup>148</sup>. In response to the H1N1 vaccine, adult females (aged 18–45 years) developed greater IL-6 and neutralizing antibody responses compared than males; this sex difference was diminished in individuals aged  $\geq 65$  years<sup>149</sup>. The attenuated sex differences in the older individuals were a consequence of reproductive senescence in the females, since females that had the highest circulating levels of E2 also had the greatest antibody responses to the vaccine<sup>149</sup>. In mice, gonadectomy reduced, whereas exogenous E2 treatment increased, IAV vaccine-induced immunity in females of reproductive age, providing additional support for a role of E2 in mediating sex differences in vaccine-induced immunity<sup>149,150</sup>. Mechanistically, E2 enhances SHM and CSR in B cells via the upregulation of AID, which contains an ERE in its promoter region<sup>126</sup>. Several additional genes associated with CSR and SHM, including *TLR7* and *NF $\kappa$ B* also have confirmed EREs in their promoters<sup>151–154</sup>. Thus, E2 has enhancing effects on antibody production by B cells.

Although less well characterized, testosterone can also affect vaccine-induced immunity by reducing the production of IFN $\gamma$  by T cells and by inhibiting antibody production as has been demonstrated in mice that received live vaccine against malaria<sup>100</sup>. Furthermore, in humans, reduced neutralizing antibody responses to influenza vaccination correlate with the higher expression of genes involved in lipid metabolism and serum testosterone concentrations in males<sup>148</sup>.

Sex-specific effects of ageing on humoral immunity also exist (Box 1). Few reports exist of sex differences in the humoral immune response of children prior to puberty. The numbers of B cells and levels of IgG and IgM are comparable in males and females in early life, with limited evidence suggesting that infant males have greater IgA and IgE levels than females<sup>2</sup>. However, compared with pre-pubertal boys, girls have greater levels of BAFF at birth<sup>155</sup>, suggesting that sex differences in B cell maturation and survival probably begin in utero. By contrast, boys have greater proportions of immature or naive CD5<sup>+</sup> B cells during the first 3 years of life than girls<sup>155</sup>. At the other end of the spectrum, aged females have comparatively higher numbers of age-associated B cells, which have been shown to take over antibody production from long-lived plasma cells at older ages<sup>156,157</sup>. In mice, the induction and activity of age-associated B cells is dependent on TLR signalling in B cells<sup>157,158</sup>, which is also higher in females than in males<sup>15</sup>.

Thus, sex differences in measures of humoral immunity are apparent across the life course, with females typically exhibiting greater B cell activation and antibody responses than males. Estrogens, in particular E2, are potent mediators of greater humoral immunity in females, which seems to be beneficial for vaccine-induced immunity and protection against infectious diseases, but as discussed below, is detrimental for autoimmunity and organ transplantation outcomes.

## Consequences of sex differences Autoimmune diseases

Autoimmune diseases affect about 5% of the human population and almost 80% of those affected are females<sup>159</sup>. Some autoimmune

diseases such as SLE affect females ten times more frequently than males, whereas a female preponderance is present, but less prominent, for diseases such as multiple sclerosis<sup>160</sup>. By contrast, some autoimmune diseases, including type 1 diabetes mellitus (T1DM), affect males and females equally<sup>160</sup>. T1DM and multiple sclerosis are both classic T cell-mediated autoimmune diseases that target pancreatic  $\beta$ -cells and central nervous system myelin, respectively. These diseases are each associated with the inheritance of specific HLA alleles and share common risk genes involved in T cell regulation and homeostasis, such as *IL2RA*, *IL7RA*, *CD226* and *PTPN22* (ref. 161). Despite these similarities, multiple sclerosis exhibits a female predominance, whereas T1DM does not<sup>160</sup>. This difference may be because autoimmunity typically initiates prior to puberty in T1DM, and after puberty in multiple sclerosis<sup>162,163</sup>.

There is strong evidence from experimental murine models that sex steroids influence the progression of T cell-mediated autoimmune diseases. Studies in the NOD model of T1DM, which unlike human T1DM, exhibits a post-pubertal onset, and studies in models of EAE in post-pubertal mice have demonstrated that male castration accelerates, whereas treatment of castrated males with androgens protects, against these diseases<sup>164–166</sup>. Androgens limit autoreactive T cell responses in EAE and multiple sclerosis by inhibiting the expansion of T<sub>H</sub>1 cells and by promoting the production of immunoregulatory cytokines such as IL-10 and TGF $\beta$ <sup>164,167</sup>. In NOD mice, androgens also regulate the composition of the microbiota, which feeds back to enhance androgen levels and limit the development of insulinitis<sup>168,169</sup>. By contrast, oestrogens and P4 have a more limited impact on these diseases and can either enhance or suppress autoimmunity depending on the timing of exposure and the concentration of hormones administered. For example, ovariectomy protects against diabetes in NOD mice and EAE, but only when the surgery is performed prior to the onset of puberty<sup>165,170</sup>. This protection against EAE correlates with a reduction in the expression of co-stimulatory markers and MHC class II on DCs and B cells in ovariectomized female mice<sup>170</sup>. By contrast, supraphysiological (pregnancy) levels of E2 inhibit EAE<sup>171,172</sup> and prevent relapses in patients with multiple sclerosis<sup>173</sup>. In humans, multiple sclerosis and T1DM are over-represented in trans females as compared with cisgender males, highlighting the importance of sex steroids as modifiers of these human diseases<sup>174,175</sup>.

SLE typically presents in the reproductive years, with the development of anti-nuclear antibodies occurring a decade prior to diagnosis<sup>176,177</sup>. The development of anti-nuclear antibodies, which are a prelude to SLE, are present in 20% of healthy human females<sup>178</sup>. The female preponderance of SLE is apparent prior to puberty<sup>179</sup>, suggesting that sex chromosomes regulate these immune mechanisms prior to the pubertal surge in gonadal hormone levels. However, SLE is 14-fold more frequent in males with Klinefelter syndrome (XXY) than in XY males<sup>180</sup> and is less frequent in females with Turner syndrome (XO) than XX females<sup>180,181</sup>, suggesting a contribution of X chromosome gene dosage to disease risk. The effect of X chromosome gene dosage on SLE may relate to incomplete X-inactivation at gene loci leading to the increased expression of certain X-encoded genes in female immune cells<sup>182</sup>. X-escapee genes that may contribute to autoimmunity include *TLR7* and *CD40*, which are important for providing T help to B cells, and *OGT*, which regulates T cell activation through glycosylation<sup>183</sup>.

The higher female-to-male ratio among individuals with SLE in those with post-pubertal onset than in those with pre-pubertal onset also implicates sex steroids as enhancers of this disease. In lupus-prone mice, androgens limit the development of autoimmunity, whereas oestrogen treatment promotes disease development<sup>184</sup>. Oestrogens may

also promote autoimmunity in human SLE. Oral contraceptive use was associated with increased disease risk in several studies<sup>185,186</sup>, but does not induce disease flares in those with established SLE<sup>187</sup>. By contrast, pregnancy, which is associated with higher levels of circulating oestrogens, causes SLE flares<sup>185,188</sup>. E2 promotes key disease mechanisms by increasing the production of IFN $\alpha$  by pDCs<sup>51</sup> and the expression of BAFF and AID<sup>126,189</sup>. Studies in an anti-DNA BCR transgenic mouse model of SLE also demonstrated that E2 promotes autoantibody production by increasing the threshold of BCR signalling in autoreactive B cells, thereby preventing their deletion<sup>190</sup>. Thus, sex steroids have a dominant role in modulating sex differences in T cell-mediated autoimmune diseases that manifest after puberty; however, evidence also suggests that X chromosome gene dosage modulates the development of systemic autoimmunity.

## Renal inflammation and hypertension

Although adaptive immune responses are stronger overall in women than in men with SLE (described above), glomerular damage is more common among men<sup>191</sup>. Specifically, lupus nephritis is a more frequent complication in males with SLE and male sex is a risk factor for progression to kidney failure<sup>191</sup>. This male predisposition to kidney damage is also observed in other kidney diseases, including chronic kidney disease<sup>192,193</sup>, IgA nephropathy<sup>194</sup>, hypertension<sup>195</sup> and in natural ageing<sup>196</sup>. The mechanisms underlying this male bias are not well understood, but may involve sex differences in the immune system. In particular, strong evidence supports a role for sex differences in T cell immunity in mediating the male bias in hypertension.

In the spontaneously hypertensive rat (SHR), males develop greater increases in blood pressure with age than females up until the time of ovarian senescence<sup>197</sup>. The greater blood pressure in males is caused by both testosterone levels and XY chromosome complement, and is preceded by T cell infiltration in the kidney<sup>197</sup>. Although male SHRs exhibit higher frequencies of circulating T<sub>reg</sub> cells than female SHRs<sup>198</sup>, kidneys of male SHRs exhibit a lower frequency of T<sub>reg</sub> and a higher overall accumulation of T<sub>H17</sub> cells than those of females<sup>198</sup>. A similar propensity for male rodents to accumulate T<sub>H17</sub> cells and for females to maintain T<sub>reg</sub> cells in the spleen or kidney has been described in Sprague–Dawley rats with DOCA-salt-induced hypertension<sup>199</sup> and in C57BL/6J mice with angiotensin II (ANGII)-induced hypertension<sup>200</sup>. Depletion of lymphocytes lowers blood pressure in male rats<sup>201</sup>, providing evidence that the infiltration of pro-inflammatory T cells in the kidney causes hypertension development in male rats.

The mouse model of ANGIO-induced hypertension has established a central role for T cells in the regulation of blood pressure in males<sup>200,202,203</sup>. Regardless of sex, *Rag1*<sup>-/-</sup> mice, which lack lymphocytes, are resistant to the development of hypertension, but exhibit an increase in blood pressure after the adoptive transfer of male, but not female, CD3<sup>+</sup> T cells<sup>200</sup>. The elevated blood pressure of male *Rag1*<sup>-/-</sup> mice transplanted with T cells from male donors is associated with the increased production of IL-17A and TNF by T cells in the spleen<sup>200</sup>. Although the total abundance of CD4<sup>+</sup> and CD8<sup>+</sup> T cells in the kidneys of male *Rag1*<sup>-/-</sup> mice is not differentially affected by the transfer of donor male or female T cells, mRNAs encoding IFN $\gamma$ , IL-10 and FoxP3 are expressed to a higher level in *Rag1*<sup>-/-</sup> mice with adoptively transferred female T cells. This finding is consistent with the notion that female T cells are biased towards T<sub>H1</sub> and hints at the development of a more regulatory T phenotype in *Rag1*<sup>-/-</sup> mice following transfer of female, rather than male, T cells<sup>200</sup>. The sex of the *Rag1*<sup>-/-</sup> mice also affects the development of hypertension, as the transfer of male

T cells increases blood pressure only in male hosts<sup>203</sup>. The effect of male host sex is not associated with alterations in T cell infiltration in the kidneys, but instead with increased renal expression of mRNAs encoding TNF and MCP1, suggesting an interaction of the male T cells with the host male innate immune system<sup>203</sup>. Compared with female *Rag1*<sup>-/-</sup> hosts, male *Rag1*<sup>-/-</sup> hosts that receive male T cells also exhibit greater T cell infiltration in the subfornical organ, which is a brain region that is involved in mediating the sympathetic control of blood pressure<sup>203</sup>. These findings support the notion that T<sub>H17</sub> and other pro-inflammatory T cells interact with the innate immune system and the sympathetic nervous system to promote hypertension in males.

Advances have also been made in our understanding of how IL-17 induces hypertension. Hypertensive stimuli, such as ANGIO and salt, enhance T<sub>H17</sub> differentiation<sup>200,204</sup>. T<sub>H</sub> cells and  $\gamma\delta$ T cells that produce IL-17 also promote oxidative damage of the kidney and vasculature<sup>205,206</sup>. T<sub>H17</sub> cells can promote hypertension by limiting the bioavailability of nitric oxide (NO), which regulates blood vessel relaxation and limits T cell proliferation, inflammation and oxidative stress<sup>207</sup>. Females have greater dependence than males on NO production for limiting hypertension<sup>207</sup>. Pharmacological inhibition of NO synthase (NOS) activity increases blood pressure to a greater extent in female than in male rats, accompanied by increased T<sub>H17</sub> and reduced T<sub>reg</sub> infiltration in the kidney and kidney injury<sup>207</sup>.

T<sub>reg</sub> cells have a direct role in actively restraining hypertension, particularly in female rats. DOCA salt increases blood pressure in male and female Sprague–Dawley rats, but to a greater extent in males than in females<sup>208</sup>. These sex differences in blood pressure correlate with a greater increase in the frequency of T<sub>H17</sub> cells in the kidneys of males and maintenance of T<sub>reg</sub> cells in the kidneys of females<sup>208</sup>. Treatment with anti-CD25, which depletes T<sub>reg</sub> cells, increases blood pressure and albumin excretion (indicative of glomerular injury) in female but not in male rats<sup>208</sup>.

Differences in innate inflammatory mechanisms also contribute to sex differences in experimental hypertension. Salt drives macrophages towards a more pro-inflammatory state<sup>209</sup>. More macrophages infiltrate the kidney in male than in female SHRs, which correlates with increased blood pressure and oxidative stress<sup>210</sup>. Male SHRs also have greater levels of circulating mitochondrial DNA than females, which increases the in vitro contractile response of aorta from male rats, but not that of female rats, to phenylephrine<sup>211</sup>. In Dahl salt-sensitive rats, deletion of CD14 is associated with increased macrophage infiltration in the kidney, and increases blood pressure in female rats only<sup>212</sup>. Ovariectomy of female rats inhibits this hyper-inflammatory response in CD14-knockout female rats, highlighting that ovarian steroids protect against hypertension by repressing the development of innate inflammation in the kidney<sup>212</sup>. These studies thus illustrate that sex differences in the development of IL-17-mediated inflammation, innate inflammation and differential protection by T<sub>reg</sub> cells influence the development of kidney and blood vessel inflammation in hypertension. A role for IL-17 in kidney injury has been recognized in other kidney conditions as well<sup>213</sup>, and could be a common mechanism underlying sex differences in kidney disease.

Sex differences in the expression of the membrane oestrogen receptor GPER1 may also contribute to sex differences in hypertension development<sup>214–216</sup>. For example, hypertensive-prone congenic Lewis female rats that express one copy of the mouse DBA/2J Ren-2 gene develop hypertension on a high-salt diet, which is associated with higher GPER1 expression in the renal cortex<sup>214</sup>. Treatment with the GPER1 agonist G1 reduces kidney injury and tubular oxidative stress in



female rats<sup>214</sup>. Medullary infusion with G1 was found in another study to promote the excretion of sodium from the kidney in female, but not in male, Sprague–Dawley rats<sup>216</sup>. These effects of G1 may be mediated by increased activity of the endothelin system and promotion of sodium excretion from the kidney<sup>215</sup>. G1 treatment also reduces pulse pressure in female, but not in male, C57BL/6/J mice, through the decreased generation of reactive oxygen species and Nox4 expression in vascular smooth muscle cells<sup>216</sup>. Thus, GPER1 activity preferentially protects females from hypertension, but the factors that cause sex differences in GPER1 expression in the kidney or the extent to which GPER1 modulates immune responses associated with hypertension remains unknown.

## Infectious diseases

Sex differences have been reported for many infectious diseases and have been extensively studied in the context of viral infection. For some infections, such as SARS-CoV-2, males experience more severe outcomes, whereas for others, such as HIV, females are more severely affected. Both biological and sociocultural factors are factors in these sex differences. On a global scale, SARS-CoV-2 infection rates were similar between males and females<sup>217</sup>; however, severe COVID-19 disease outcomes, including hospitalization, admission to the intensive care unit and death, were more likely to occur in males<sup>218–221</sup>. Similarly, more severe disease outcomes were seen in males following infection in other outbreaks of  $\beta$ -coronavirus, including SARS-CoV and Middle-East respiratory syndrome virus (MERS)<sup>222,223</sup>. Whether sex differences in susceptibility to infection or disease outcomes are driven by gendered behaviours that affect viral exposure or sex differences in susceptibility to infection remains to be clarified. For example, the prevalence of infection is greater among adult males than among females for Dengue virus, hantaviruses and hepatitis B (HBV) and HCV viruses<sup>224</sup>. Assuming that differences in disease prevalence reflect differences in exposure, these differences suggest that behavioural or occupational exposures may contribute to sex differences in infection with these viruses. By contrast, biological factors that regulate susceptibility to infection are likely to drive the comparably more severe disease outcomes seen in males infected with Epstein–Barr virus, HBV, HCV and West Nile virus<sup>224,225</sup>. For some viruses, including HBV and HCV, both gender and biological sex contribute to male-biased outcomes. Although less common, some viruses (for example, cytomegalovirus, herpes simplex virus type 2 and human T cell leukaemia virus type 1) are more prevalent in females, whereas other viruses (for example, hantaviruses, HSV2, HIV, IAV and measles virus) cause more severe disease following infection in females than in males<sup>225</sup>.

Sex differences in influenza pathogenesis have been evaluated in small-animal models<sup>43,226–229</sup>. Several studies in mice have demonstrated that the lethal dose of either H1N1 or H3N2 viruses required to kill 50% of mice (LD<sub>50</sub>) is significantly lower for females than males<sup>230</sup>, despite similar pulmonary viral loads and viral clearance during infection<sup>226,230,231</sup>. Consistent with findings in humans<sup>232</sup>, adult female mice experience more extensive pulmonary inflammation and more severe outcomes upon infection with diverse strains of IAV compared with males, despite similar virus titres<sup>226,227,230,233</sup>. The ‘cytokine storm’ is a classic example of an over-reactive inflammatory response to a viral infection, and pulmonary concentrations of pro-inflammatory cytokines (for example, TNF, IFN $\gamma$ , IL-6 and IL-12) and chemokines (for example, CCL2, CCL5 and CCL12) are higher in female than in male mice infected with IAV<sup>226,234</sup>. Not only do adult male mice have less pulmonary inflammation, but their damaged pulmonary tissue repairs faster than that of females<sup>231</sup>. In response to pulmonary tissue damage, epithelial cells release factors,

including AREG, that can promote the repair and integrity of lung tissue damaged during IAV infection<sup>235</sup>. Expression of AREG in lung tissue and respiratory epithelial cells is greater in male mice than in female mice during IAV infection<sup>231</sup>. Studies in *Areg*-knockout mice also show that males exhibit a greater dependence on AREG expression than females to support their faster recovery from IAV infection<sup>231</sup>. Interestingly, infection of young adult female mice with IAVs reduces ovarian function and concentrations of sex hormones<sup>226,236</sup>, indicating that inhibition of sex hormone production, including oestrogens and P4 may also contribute to the more severe outcomes seen in females during IAV infection.

Mouse studies of IAV infection have highlighted the important immunomodulatory effects of sex steroids in the lung and airways<sup>43,228,229,236,237</sup>. Exogenous treatment with P4 or a synthetic form of progestin found in contraceptives, attenuates pulmonary inflammation and expedites tissue repair following IAV infection, in part through upregulation of AREG<sup>43,238</sup>. Oestradiol protects females against severe IAV infection by dampening inflammation, promoting the recruitment of neutrophils and virus-specific CD8<sup>+</sup> T cells to the lungs, and increasing overall tolerance<sup>228</sup>. These protective effects occur mainly through ER $\alpha$  signalling<sup>226</sup>. Testosterone, which is also broadly anti-inflammatory, protects male mice from severe IAV by enhancing the contraction phase of the immune response in the lungs after viral clearance<sup>239</sup>. These protective effects of testosterone occur through AR signalling; inhibition of AR signalling through the co-administration of flutamide and testosterone prior to infection reverses, whereas treatment with DHT (which as mentioned above, is a non-aromatizable androgen that cannot signal through the ER) mimics, the protective effects of testosterone in limiting IAV infection morbidity<sup>239</sup>. Lower testosterone concentrations – as a consequence of either gonadectomy or ageing – are associated with increased pulmonary inflammation, morbidity and mortality from IAV infection, at least in mice<sup>228,229,231</sup>.

## Transplant rejection

Among recipients of solid organ transplants, clinically important differences in graft survival exist based on sex and age. The ability to parse out the independent roles of recipient sex and age is complicated by concomitant variability in other key factors, including organ transplant type, immunosuppressive medication, donor sex and/or issues relating to donor–recipient sex mismatch including size, and preceding sensitization as a consequence of prior organ transplantation, blood transfusion or pregnancy.

A comparison of graft survival among nearly 408,000 kidney transplant recipients in the Scientific Registry of Transplant Recipients (SRTR) found that 5-year death-censored graft survival was worse among young female recipients (aged 15–34 years) compared with that among similarly aged males. Graft survival was greater in individuals aged 35–54 years, with no difference between sexes, and was greatest among individuals aged 55–75 years, but slightly worse in male recipients of this age group compared with that in similarly aged females<sup>240</sup>. Moreover, young female recipients (aged 15–34 years) had a twofold higher risk of acute rejection within 1 year of transplantation than females in the oldest age group (aged 55–75 years). The interaction between sex and age in relation to graft survival highlights the probable role of the sex steroid milieu in modifying alloimmunity. Similar results have been demonstrated among kidney transplant recipients in at least two other large transplant registries outside the USA, and in studies that have considered age at time of organ rejection as a variable<sup>241,242</sup>.



Heart transplant recipients also demonstrate differences in graft failure rates and rejection according to age and sex. One multicentre study showed that females experience, on average, more episodes of acute rejection in the first year after transplantation than males and are more likely than males to have rejection of moderate or severe grade, requiring hospitalization<sup>243</sup>. Although that study did not stratify recipients by age, the mean age of females in the cohort (50 years) was about 6 years younger than that of males. A separate study that used SRTR data found that female recipients of male donor hearts had higher rates of graft failure across all age categories than male recipients of male donor hearts, with the largest absolute difference among recipients aged 13–24 years<sup>244</sup>. When the donor was female, no difference was observed in graft failure rates between sexes at any age<sup>244</sup>, suggesting that the higher rates of graft failure in female recipients of male donor organs was due to recognition of the HY antigen (present only in male tissues) by female (XX) immune cells<sup>244</sup>.

Oestrogen also seems to be a central factor in graft rejection. In contrast to young naive female mice, which show reduced graft survival following skin or heart transplantation, young ovariectomized females demonstrate graft survival similar to that of young males<sup>240</sup>. Ovariectomy in young mice extends skin and cardiac graft survival to a duration comparable to that in old ovariectomized, old naive and old sham control mice<sup>240</sup>. Young ovariectomized females also show reduced numbers of CD4<sup>+</sup> and CD8<sup>+</sup> T cells, dampened T<sub>H</sub>1 and T<sub>H</sub>17 responses, and increased numbers of FoxP3<sup>+</sup> T<sub>reg</sub> cells relative to sham controls.

## Box 2

### Historical perspective

The history of excluding females from biomedical and clinical studies began in 1977 when the FDA published guidelines advising that females of childbearing potential be excluded from drug trials<sup>256</sup>. The primary goal of the FDA guidelines was to protect pregnant females and their fetuses from adverse drug outcomes, but the unintended consequence was complete exclusion of females from clinical trials for decades<sup>257</sup>. In the early 1990s, the FDA and the National Institutes of Health (NIH), with advocacy from US Congresswomen, recommended that clinical trials include female participants<sup>257,258</sup>. For the next three decades, although females were included in clinical trials of drugs, devices and biologics, analysis of whether outcomes differ between females and males was inadequate<sup>259</sup>. In 2016, the NIH implemented a policy that sex as a biological variable (SABV) be factored into research design, analyses and reporting in vertebrate animal and human studies<sup>260</sup>. However, the policy did not require that researchers use methods, double sample sizes or power studies to detect sex differences. Rather, the policy requested that investigators know the existing data, balance the sexes in experimental design and consider SABV. The policy has led to mixed results across disciplines, with some disciplines, including immunology, making considerable progress in reporting the use of both sexes in animal studies<sup>263</sup>. However, the SABV policy has been criticized as it has had an unintended consequence of promoting research that lacks power to statistically compare the sexes or does not include appropriate statistical analyses to rigorously analyse sex differences<sup>261</sup>.

A more detailed analysis of the T<sub>reg</sub> cell population in a clinical cohort of young kidney transplant recipients (aged 3–29 years) showed that the frequency of highly suppressive, Helios<sup>+</sup> T<sub>reg</sub> cells is lower among female recipients than male recipients<sup>245</sup>. Thus, E2 shapes CD4<sup>+</sup> T cell activity and contributes to age-related alloimmunity.

Antibody-mediated immunity also contributes to post-transplantation outcomes. Panel-reactive antibodies (PRAs) are anti-HLA antibodies against donor MHC class I and II antigens expressed on cells present in the transplanted allograft and are a risk factor for adverse transplantation outcomes, including rejection, chronic allograft vasculopathy and death<sup>246</sup>. Pre-transplantation PRA levels are greater in women than in men, particularly in those with a history of pregnancy<sup>240,246</sup>. Indeed, ~50–75% of women are sensitized by pregnancy according to contemporary assessments of PRA; moreover, 63% of transplant candidates with an extremely high calculated PRA (that is, ≥98%) are women<sup>247–249</sup>. Higher sensitization among women reduces their likelihood of receiving a living donor transplant and prolongs their waitlist time compared with men<sup>250</sup>. Sex is a modifier of the association between high peak PRA levels and poor transplant outcomes<sup>246</sup>. Female heart transplant recipients with high peak PRA have an increased risk of graft failure compared with male recipients, whereas male recipients with high peak PRA have an increased risk of chronic allograft vasculopathy compared with female recipients<sup>246</sup>.

A complex relationship exists between pregnancy-induced alloimmunization and the contributions of T and B cell memory and function. Pregnancy induces the development of anti-HLA antibodies, and multiple pregnancies increase the risk of retaining detectable antibodies<sup>251</sup>. Over time, anti-HLA antibodies may become undetectable but a substantial proportion of parous women retain detectable antibodies<sup>248</sup>. Alloreactive memory B cells also persist and carry the potential to rapidly produce anti-HLA antibodies following transplantation<sup>252</sup>. In addition to humoral sensitization, pregnancy primes maternal T cell responses for immune tolerance to the semi-allogeneic fetus through the selective expansion of fetus-specific T<sub>reg</sub> cells, hypofunctional CD8<sup>+</sup> T cells that exhibit an 'exhaustion' phenotype, and anergic CD4<sup>+</sup> T cells<sup>248,249</sup>. Memory T cells persist after pregnancy but do not cause fetal loss upon a subsequent antigen-positive pregnancy, potentially as a consequence of chemokine gene silencing in decidual stromal cells, which prevents the T cells from accessing the maternal–fetal interface<sup>249</sup>. By contrast, few data are available to inform whether pregnancy-primed T cells are functionally tolerant to fetal antigens presented in the setting of organ transplantation. Emerging data from studies in animal models support the hypothesis that pregnancy-induced T cell tolerance is overridden by pregnancy-primed memory B cells and allo-antibodies in response to fetal antigens present on an organ allograft. These data also suggest that if the presence of B cells and alloantibody are negated, T cell tolerance to fetal antigens present on an allograft is permitted<sup>249</sup>. More work is needed to characterize the T cell repertoire among parous women in the transplantation setting. Animal studies suggest that humoral desensitization may be a key strategy in reducing pregnancy-induced alloimmunity<sup>249</sup>.

### Conclusions

Historically there has been a lack of inclusion of females in biomedical, clinical and public health research, which has hindered our understanding of sex differences in disease presentation, progression and responses to treatment. In the USA, the implementation of the National Institutes of Health requirement for consideration of sex as a biological variable (Box 2) has stimulated studies of sex differences in some scientific fields,

including immunology<sup>253</sup>. Broadly speaking, however, sex-related and gender-related reporting remains low within the biomedical sciences, with female researchers being more likely than male researchers to report sex-based or gender-based differences, and the findings of such studies largely published in low-impact journals<sup>254</sup>. Regardless, a large and growing body of work has highlighted important sex differences in immune cell function and immune responses, with important consequences for disease epidemiology and outcomes. Greater understanding of these differences may lead to the identification of new targetable pathways for novel treatment discoveries.

Published online: 22 November 2023

## References

- Mauvais-Jarvis, F. et al. Sex and gender: modifiers of health, disease, and medicine. *Lancet* **396**, 565–582 (2020).
- Klein, S. L. & Flanagan, K. L. Sex differences in immune responses. *Nat. Rev. Immunol.* **16**, 626–638 (2016).
- Haupt, S., Caramia, F., Klein, S. L., Rubin, J. B. & Haupt, Y. Sex disparities matter in cancer development and therapy. *Nat. Rev. Cancer* **21**, 393–407 (2021).
- Laffont, S. & Guery, J. C. Deconstructing the sex bias in allergy and autoimmunity: from sex hormones and beyond. *Adv. Immunol.* **142**, 35–64 (2019).
- Melk, A. et al. Equally interchangeable? How sex and gender affect transplantation. *Transplantation* **103**, 1094–1110 (2019).
- Klein, S. L. & Morgan, R. The impact of sex and gender on immunotherapy outcomes. *Biol. Sex Differ.* **11**, 24 (2020).
- Fish, E. N. The X-files in immunity: sex-based differences predispose immune responses. *Nat. Rev. Immunol.* **8**, 737–744 (2008).
- Migeon, B. R. Why females are mosaics, X-chromosome inactivation, and sex differences in disease. *Genet. Med.* **4**, 97–105 (2007).
- Oktenli, C. et al. Study of autoimmunity in Klinefelter's syndrome and idiopathic hypogonadotropic hypogonadism. *J. Clin. Immunol.* **22**, 137–143 (2002).
- Cacciarri, E. et al. Serum immunoglobulins and lymphocyte subpopulations derangement in Turner's syndrome. *J. Immunogenet.* **8**, 337–344 (1981).
- Gravholt, C. H. Turner syndrome in adulthood. *Horm. Res.* **64**, 86–93 (2005).
- Schmiedel, B. J. et al. Impact of genetic polymorphisms on human immune cell gene expression. *Cell* **175**, 1701–1715.e6 (2018).
- Itoh, Y. et al. The X-linked histone demethylase Kdm6a in CD4<sup>+</sup> T lymphocytes modulates autoimmunity. *J. Clin. Invest.* **129**, 3852–3863 (2019).
- Souyris, M. et al. TLR7 escapes X chromosome inactivation in immune cells. *Sci. Immunol.* **3**, eaap8855 (2018).
- Fink, A. L., Engle, K., Ursin, R. L., Tang, W. Y. & Klein, S. L. Biological sex affects vaccine efficacy and protection against influenza in mice. *Proc. Natl Acad. Sci. USA* **115**, 12477–12482 (2018).
- van der Made, C. I. et al. Presence of genetic variants among young men with severe COVID-19. *JAMA* **324**, 663–673 (2020).
- Cheng, M. I. et al. The X-linked epigenetic regulator UTX controls NK cell-intrinsic sex differences. *Nat. Immunol.* **24**, 780–791 (2023).
- Sharma, S. & Eghbali, M. Influence of sex differences on microRNA gene regulation in disease. *Biol. Sex Differ.* **5**, 3 (2014).
- Malmhall, C., Weidner, J. & Radinger, M. MicroRNA-155 expression suggests a sex disparity in innate lymphoid cells at the single-cell level. *Cell Mol. Immunol.* **17**, 544–546 (2020).
- Arnold, A. P. & Chen, X. What does the “four core genotypes” mouse model tell us about sex differences in the brain and other tissues? *Front. Neuroendocrinol.* **30**, 1–9 (2009).
- Golden, L. C. et al. Parent-of-origin differences in DNA methylation of X chromosome genes in T lymphocytes. *Proc. Natl Acad. Sci. USA* **116**, 26779–26787 (2019).
- Case, L. K. et al. The Y chromosome as a regulatory element shaping immune cell transcriptomes and susceptibility to autoimmune disease. *Genome Res.* **23**, 1474–1485 (2013).
- Kremontsov, D. N. et al. Genetic variation in chromosome Y regulates susceptibility to influenza A virus infection. *Proc. Natl Acad. Sci. USA* **114**, 3491–3496 (2017).
- Thompson, D. J. et al. Genetic predisposition to mosaic Y chromosome loss in blood. *Nature* **575**, 652–657 (2019).
- Weinstein, Y., Ran, S. & Segal, S. Sex-associated differences in the regulation of immune responses controlled by the MHC of the mouse. *J. Immunol.* **132**, 656–661 (1984).
- Eidinger, D. & Garrett, T. J. Studies of the regulatory effects of the sex hormones on antibody formation and stem cell differentiation. *J. Exp. Med.* **136**, 1098–1116 (1972).
- Bjornstrom, L. & Sjoberg, M. Mechanisms of estrogen receptor signaling: convergence of genomic and nongenomic actions on target genes. *Mol. Endocrinol.* **19**, 833–842 (2005).
- Cvoro, A. et al. Distinct roles of unliganded and liganded estrogen receptors in transcriptional repression. *Mol. Cell* **21**, 555–564 (2006).
- Blanquart, E., Laffont, S. & Guery, J. C. Sex hormone regulation of innate lymphoid cells. *Biomed. J.* **44**, 144–156 (2021).
- Kovats, S. Estrogen receptors regulate innate immune cells and signaling pathways. *Cell Immunol.* **294**, 63–69 (2015).
- Notas, G., Kampa, M. & Castanas, E. G protein-coupled estrogen receptor in immune cells and its role in immune-related diseases. *Front. Endocrinol.* **11**, 579420 (2020).
- Phiel, K. L., Henderson, R. A., Adelman, S. J. & Elloso, M. M. Differential estrogen receptor gene expression in human peripheral blood mononuclear cell populations. *Immunol. Lett.* **97**, 107–113 (2005).
- Brundin, P. M. A. et al. Expression of sex hormone receptor and immune response genes in peripheral blood mononuclear cells during the menstrual cycle. *Front. Endocrinol.* **12**, 721813 (2021).
- Fox, H. S., Bond, B. L. & Parslow, T. G. Estrogen regulates the IFN-gamma promoter. *J. Immunol.* **146**, 4362–4367 (1991).
- Stein, B. & Yang, M. X. Repression of the interleukin-6 promoter by estrogen receptor is mediated by NF-κB and C/EBPβ. *Mol. Cell Biol.* **15**, 4971–4979 (1995).
- Wang, C. et al. Membrane estrogen receptor regulates experimental autoimmune encephalomyelitis through up-regulation of programmed death 1. *J. Immunol.* **182**, 3294–3303 (2009).
- Wang, C. et al. GPR30 contributes to estrogen-induced thymic atrophy. *Mol. Endocrinol.* **22**, 636–648 (2008).
- Tan, I. J., Peeva, E. & Zandman-Goddard, G. Hormonal modulation of the immune system – a spotlight on the role of progestogens. *Autoimmun. Rev.* **14**, 536–542 (2015).
- Griekspoor, A., Zwart, W., Neefjes, J. & Michalides, R. Visualizing the action of steroid hormone receptors in living cells. *Nucl. Recept. Signal.* **5**, e003 (2007).
- Christin-Maitre, S. History of oral contraceptive drugs and their use worldwide. *Best. Pract. Res. Clin. Endocrinol. Metab.* **27**, 3–12 (2013).
- Meier, A. et al. Sex differences in the toll-like receptor-mediated response of plasmacytoid dendritic cells to HIV-1. *Nat. Med.* **15**, 955–959 (2009).
- Su, S. et al. Modulation of innate immune response to viruses including SARS-CoV-2 by progesterone. *Signal. Transduct. Target. Ther.* **7**, 137 (2022).
- Hall, O. J. et al. Progesterone-based therapy protects against influenza by promoting lung repair and recovery in females. *PLoS Pathog.* **12**, e1005840 (2016).
- Park, C. J. et al. Progesterone receptor serves the ovary as a trigger of ovulation and a terminator of inflammation. *Cell Rep.* **31**, 107496 (2020).
- Butts, C. L. et al. Progesterone inhibits mature rat dendritic cells in a receptor-mediated fashion. *Int. Immunol.* **19**, 287–296 (2007).
- Pauklin, S. & Petersen-Mahrt, S. K. Progesterone inhibits activation-induced deaminase by binding to the promoter. *J. Immunol.* **183**, 1238–1244 (2009).
- Gubbels Bupp, M. R. & Jorgensen, T. N. Androgen-induced immunosuppression. *Front. Immunol.* **9**, 794 (2018).
- Jacquelot, N., Luong, K. & Seillet, C. Physiological regulation of innate lymphoid cells. *Front. Immunol.* **10**, 405 (2019).
- Berghofer, B. et al. TLR7 ligands induce higher IFN-α production in females. *J. Immunol.* **177**, 2088–2096 (2006).
- Ziegler, S. M. et al. Human pDCs display sex-specific differences in type I interferon subtypes and interferon α/β receptor expression. *Eur. J. Immunol.* **47**, 251–256 (2017).
- Seillet, C. et al. The TLR-mediated response of plasmacytoid dendritic cells is positively regulated by estradiol in vivo through cell-intrinsic estrogen receptor α signaling. *Blood* **119**, 454–464 (2012).
- Wang, J. P., Zhang, L., Madera, R. F., Woda, M. & Libraty, D. H. Plasmacytoid dendritic cell interferon-α production to R-848 stimulation is decreased in male infants. *BMC Immunol.* **13**, 35 (2012).
- Webb, K. et al. Sex and pubertal differences in the type 1 interferon pathway associate with both X chromosome number and serum sex hormone concentration. *Front. Immunol.* **9**, 3167 (2018).
- Griesbeck, M. et al. Sex differences in plasmacytoid dendritic cell levels of IRF5 drive higher IFN-α production in women. *J. Immunol.* **195**, 5327–5336 (2015).
- Panchanathan, R., Liu, H. & Choubey, D. Expression of murine Unc93b1 is up-regulated by interferon and estrogen signaling: implications for sex bias in the development of autoimmunity. *Int. Immunol.* **25**, 521–529 (2013).
- Laffont, S. et al. X-chromosome complement and estrogen receptor signaling independently contribute to the enhanced TLR7-mediated IFN-α production of plasmacytoid dendritic cells from women. *J. Immunol.* **193**, 5444–5452 (2014).
- Hagen, S. H. et al. Heterogeneous escape from X chromosome inactivation results in sex differences in type I IFN responses at the single human pDC level. *Cell Rep.* **33**, 108485 (2020).
- MacCall, C. A., Ritchie, G. & Sood, M. Oral fluid testing as an alternative to urine testing for drugs of abuse in inpatient forensic settings: giving patients choice. *Scott. Med. J.* **58**, 99–103 (2013).
- Congy-Jolivet, N. et al. Monocytes are the main source of STING-mediated IFN-α production. *EBioMedicine* **80**, 104047 (2022).
- Gal-Oz, S. T. et al. ImmGen report: sexual dimorphism in the immune system transcriptome. *Nat. Commun.* **10**, 4295 (2019).
- Kim, S. & Voskuhl, R. R. Decreased IL-12 production underlies the decreased ability of male lymph node cells to induce experimental autoimmune encephalomyelitis. *J. Immunol.* **162**, 5561–5568 (1999).
- Cua, D. J., Hinton, D. R. & Stohlman, S. A. Self-antigen-induced Th2 responses in experimental allergic encephalomyelitis (EAE)-resistant mice. Th2-mediated suppression of autoimmune disease. *J. Immunol.* **155**, 4052–4059 (1995).

63. Drohomyrecky, P. C. et al. Peroxisome proliferator-activated receptor- $\delta$  acts within peripheral myeloid cells to limit Th cell priming during experimental autoimmune encephalomyelitis. *J. Immunol.* **203**, 2588–2601 (2019).
64. Wilcoxon, S. C., Kirkman, E., Dowdell, K. C. & Stohlman, S. A. Gender-dependent IL-12 secretion by APC is regulated by IL-10. *J. Immunol.* **164**, 6237–6243 (2000).
65. Zhang, M. A. et al. Peroxisome proliferator-activated receptor (PPAR) $\alpha$  and  $\gamma$  regulate IFN $\gamma$  and IL-17A production by human T cells in a sex-specific way. *Proc. Natl Acad. Sci. USA* **109**, 9505–9510 (2012).
66. Delpy, L. et al. Estrogen enhances susceptibility to experimental autoimmune myasthenia gravis by promoting type 1-polarized immune responses. *J. Immunol.* **175**, 5050–5057 (2005).
67. Kovats, S. Estrogen receptors regulate an inflammatory pathway of dendritic cell differentiation: mechanisms and implications for immunity. *Horm. Behav.* **62**, 254–262 (2012).
68. Douin-Echinard, V. et al. Estrogen receptor  $\alpha$ , but not  $\beta$ , is required for optimal dendritic cell differentiation and CD40-induced cytokine production. *J. Immunol.* **180**, 3661–3669 (2008).
69. Scotland, R. S., Stables, M. J., Madalli, S., Watson, P. & Gilroy, D. W. Sex differences in resident immune cell phenotype underlie more efficient acute inflammatory responses in female mice. *Blood* **118**, 5918–5927 (2011).
70. Bain, C. C. et al. Rate of replenishment and microenvironment contribute to the sexually dimorphic phenotype and function of peritoneal macrophages. *Sci. Immunol.* **5**, eabc4466 (2020).
71. Jaillon, S., Berthenet, K. & Garlanda, C. Sexual dimorphism in innate immunity. *Clin. Rev. Allergy Immunol.* **56**, 308–321 (2019).
72. Li, K. et al. Differential macrophage polarization in male and female BALB/c mice infected with coxsackievirus B3 defines susceptibility to viral myocarditis. *Circ. Res.* **105**, 353–364 (2009).
73. Denny, M. et al. Sex difference in innate inflammatory response and macrophage polarization in *Streptococcus agalactiae*-induced pneumonia and potential role of microRNA-223-3p. *Sci. Rep.* **12**, 17126 (2022).
74. Melgert, B. N. et al. Macrophages: regulators of sex differences in asthma? *Am. J. Respir. Cell Mol. Biol.* **42**, 595–603 (2010).
75. Huber, S. & Sartini, D. T cells expressing the V $\gamma$ 1 T-cell receptor enhance virus-neutralizing antibody response during coxsackievirus B3 infection of BALB/c mice: differences in male and female mice. *Viral Immunol.* **18**, 730–739 (2005).
76. Ballard, M. S. et al. The changing epidemiology of group B streptococcus bloodstream infection: a multi-national population-based assessment. *Infect. Dis.* **48**, 386–391 (2016).
77. Tribbels-Smeulders, M. et al. Epidemiology of neonatal group B streptococcal disease in the Netherlands before and after introduction of guidelines for prevention. *Arch. Dis. Child. Fetal Neonatal Ed.* **92**, F271–276 (2007).
78. Dick, S. A., Zaman, R. & Epelman, S. Using high-dimensional approaches to probe monocytes and macrophages in cardiovascular disease. *Front. Immunol.* **10**, 2146 (2019).
79. Lu, R. J. et al. Multi-omic profiling of primary mouse neutrophils predicts a pattern of sex and age-related functional regulation. *Nat. Aging* **1**, 715–733 (2021).
80. Chuang, K. H. et al. Neutropenia with impaired host defense against microbial infection in mice lacking androgen receptor. *J. Exp. Med.* **206**, 1181–1199 (2009).
81. Markman, J. L. et al. Loss of testosterone impairs anti-tumor neutrophil function. *Nat. Commun.* **11**, 1613 (2020).
82. Zhang, M. A. et al. Antagonizing peroxisome proliferator-activated receptor  $\alpha$  activity selectively enhances Th1 immunity in male mice. *J. Immunol.* **195**, 5189–5202 (2015).
83. Hrushesky, W. J. et al. Natural killer cell activity: age, estrous- and circadian-stage dependence and inverse correlation with metastatic potential. *J. Natl Cancer Inst.* **80**, 1232–1237 (1988).
84. Lee, S. et al. Fluctuation of peripheral blood T, B, and NK cells during a menstrual cycle of normal healthy women. *J. Immunol.* **185**, 756–762 (2010).
85. Zychlinsky Scharff, A. et al. Sex differences in IL-17 contribute to chronicity in male versus female urinary tract infection. *JCI Insight* **5**, e122998 (2019).
86. Darboe, A. et al. Age-related dynamics of circulating innate lymphoid cells in an African population. *Front. Immunol.* **11**, 594107 (2020).
87. Kadel, S. et al. A major population of functional KLRG1<sup>+</sup> ILC2s in female lungs contributes to a sex bias in ILC2 numbers. *Immunohorizons* **2**, 74–86 (2018).
88. Cephus, J. Y. et al. Testosterone attenuates group 2 innate lymphoid cell-mediated airway inflammation. *Cell Rep.* **21**, 2487–2499 (2017).
89. Laffont, S. et al. Androgen signaling negatively controls group 2 innate lymphoid cells. *J. Exp. Med.* **214**, 1581–1592 (2017).
90. Gandhi, V. D. et al. Androgen receptor signaling promotes Treg suppressive function during allergic airway inflammation. *J. Clin. Invest.* **132**, e153397 (2022).
91. Russi, A. E., Ebel, M. E., Yang, Y. & Brown, M. A. Male-specific IL-33 expression regulates sex-dimorphic EAE susceptibility. *Proc. Natl Acad. Sci. USA* **115**, E1520–E1529 (2018).
92. Terabe, M. & Berzofsky, J. A. Tissue-specific roles of NKT cells in tumor immunity. *Front. Immunol.* **9**, 1838 (2018).
93. Sandberg, J. K., Bhardwaj, N. & Nixon, D. F. Dominant effector memory characteristics, capacity for dynamic adaptive expansion, and sex bias in the innate V $\alpha$ 24 NKT cell compartment. *Eur. J. Immunol.* **33**, 588–596 (2003).
94. Bernin, H., Fehling, H., Marggraft, C., Tannich, E. & Lotter, H. The cytokine profile of human NKT cells and PBMCs is dependent on donor sex and stimulus. *Med. Microbiol. Immunol.* **205**, 321–332 (2016).
95. Gourdy, P. et al. Relevance of sexual dimorphism to regulatory T cells: estradiol promotes IFN- $\gamma$  production by invariant natural killer T cells. *Blood* **105**, 2415–2420 (2005).
96. Lotter, H., Jacobs, T., Gaworski, I. & Tannich, E. Sexual dimorphism in the control of amebic liver abscess in a mouse model of disease. *Infect. Immun.* **74**, 118–124 (2006).
97. Lotter, H. et al. Natural killer T cells activated by a lipopeptidophosphoglycan from *Entamoeba histolytica* are critically important to control amebic liver abscess. *PLoS Pathog.* **5**, e1000434 (2009).
98. Palaszynski, K. M. et al. A yin-yang effect between sex chromosome complement and sex hormones on the immune response. *Endocrinology* **146**, 3280–3285 (2005).
99. Ma, L. J. et al. Local cytokine levels associated with delayed-type hypersensitivity responses: modulation by gender, ovariectomy, and estrogen replacement. *J. Endocrinol.* **193**, 291–297 (2007).
100. Vom Steeg, L. G., Flores-Garcia, Y., Zavala, F. & Klein, S. L. Irradiated sporozoite vaccination induces sex-specific immune responses and protection against malaria in mice. *Vaccine* **37**, 4468–4476 (2019).
101. Peacock, J. W. et al. Gender differences in human immunodeficiency virus type 1-specific CD8 responses in the reproductive tract and colon following nasal peptide priming and modified vaccinia virus Ankara boosting. *J. Virol.* **78**, 13163–13172 (2004).
102. Kondo, H. et al. Markers of memory CD8 T cells depicting the effect of the BNT162b2 mRNA COVID-19 vaccine in Japan. *Front. Immunol.* **13**, 836923 (2022).
103. Stanberry, L. R. et al. Glycoprotein-D-adjunct vaccine to prevent genital herpes. *N. Engl. J. Med.* **347**, 1652–1661 (2002).
104. Umlauf, B. J. et al. Associations between demographic variables and multiple measles-specific innate and cell-mediated immune responses after measles vaccination. *Viral Immunol.* **25**, 29–36 (2012).
105. Amadori, A. et al. Genetic control of the CD4/CD8 T-cell ratio in humans. *Nat. Med.* **1**, 1279–1283 (1995).
106. Breznik, J. A., Schulz, C., Ma, J., Sloboda, D. M. & Bowdish, D. M. E. Biological sex, not reproductive cycle, influences peripheral blood immune cell prevalence in mice. *J. Physiol.* **599**, 2169–2195 (2021).
107. Pido-Lopez, J., Imami, N. & Aspinall, R. Both age and gender affect thymic output: more recent thymic migrants in females than males as they age. *Clin. Exp. Immunol.* **125**, 409–413 (2001).
108. Aguirre-Gamboa, R. et al. Differential effects of environmental and genetic factors on T and B cell immune traits. *Cell Rep.* **17**, 2474–2487 (2016).
109. Gui, J., Mustachio, L. M., Su, D. M. & Craig, R. W. Thymus size and age-related thymic involution: early programming, sexual dimorphism, progenitors and stroma. *Aging Dis.* **3**, 280–290 (2012).
110. Olsen, N. J., Olson, G., Viselli, S. M., Gu, X. & Kovacs, W. J. Androgen receptors in thymic epithelium modulate thymus size and thymocyte development. *Endocrinology* **142**, 1278–1283 (2001).
111. Utsuyama, M. & Hirokawa, K. Hypertrophy of the thymus and restoration of immune functions in mice and rats by gonadectomy. *Mech. Ageing Dev.* **47**, 175–185 (1989).
112. Sutherland, J. S. et al. Activation of thymic regeneration in mice and humans following androgen blockade. *J. Immunol.* **175**, 2741–2753 (2005).
113. Proekt, I., Miller, C. N., Lionakis, M. S. & Anderson, M. S. Insights into immune tolerance from AIRE deficiency. *Curr. Opin. Immunol.* **49**, 71–78 (2017).
114. Zhu, M. L. et al. Sex bias in CNS autoimmune disease mediated by androgen control of autoimmune regulator. *Nat. Commun.* **7**, 11350 (2016).
115. Dragin, N. et al. Estrogen-mediated downregulation of AIRE influences sexual dimorphism in autoimmune diseases. *J. Clin. Invest.* **126**, 1525–1537 (2016).
116. Dumont-Lagace, M., St-Pierre, C. & Perreault, C. Sex hormones have pervasive effects on thymic epithelial cells. *Sci. Rep.* **5**, 12895 (2015).
117. Hun, M. L. et al. Gender disparity impacts on thymus aging and LHRH receptor antagonist-induced thymic reconstitution following chemotherapeutic damage. *Front. Immunol.* **11**, 302 (2020).
118. Huang, Z. et al. Effects of sex and aging on the immune cell landscape as assessed by single-cell transcriptomic analysis. *Proc. Natl Acad. Sci. USA* **118**, e2023216118 (2021).
119. Sankaran-Walters, S. et al. Sex differences matter in the gut: effect on mucosal immune activation and inflammation. *Biol. Sex. Differ.* **4**, 10 (2013).
120. Hewagama, A., Patel, D., Yarlaga, S., Strickland, F. M. & Richardson, B. C. Stronger inflammatory/cytotoxic T-cell response in women identified by microarray analysis. *Genes. Immun.* **10**, 509–516 (2009).
121. Guan, X. et al. Androgen receptor activity in T cells limits checkpoint blockade efficacy. *Nature* **606**, 791–796 (2022).
122. Park, H. J., Park, H. S., Lee, J. U., Bothwell, A. L. & Choi, J. M. Gender-specific differences in PPAR $\gamma$  regulation of follicular helper T cell responses with estrogen. *Sci. Rep.* **6**, 28495 (2016).
123. Maret, A. et al. Estradiol enhances primary antigen-specific CD4 T cell responses and Th1 development in vivo. Essential role of estrogen receptor  $\alpha$  expression in hematopoietic cells. *Eur. J. Immunol.* **33**, 512–521 (2003).
124. Bao, M., Yang, Y., Jun, H. S. & Yoon, J. W. Molecular mechanisms for gender differences in susceptibility to T cell-mediated autoimmune diabetes in nonobese diabetic mice. *J. Immunol.* **168**, 5369–5375 (2002).
125. Kissick, H. T. et al. Androgens alter T-cell immunity by inhibiting T-helper 1 differentiation. *Proc. Natl Acad. Sci. USA* **111**, 9887–9892 (2014).
126. Pauklin, S., Hernandez, I. V., Bachmann, G., Ramiro, A. R. & Petersen-Mahrt, S. K. Estrogen directly activates AID transcription and function. *J. Exp. Med.* **206**, 99–111 (2009).



127. Dunn, S. E. et al. Peroxisome proliferator-activated receptor (PPAR) $\alpha$  expression in T cells mediates gender differences in development of T cell-mediated autoimmunity. *J. Exp. Med.* **204**, 321–330 (2007).
128. Cephus, J. Y. et al. Estrogen receptor- $\alpha$  signaling increases allergen-induced IL-33 release and airway inflammation. *Allergy* **76**, 255–268 (2021).
129. Dimitrijevic, M. et al. Sex differences in Tfh cell help to B cells contribute to sexual dimorphism in severity of rat collagen-induced arthritis. *Sci. Rep.* **10**, 1214 (2020).
130. Foster, A. D. et al. Donor CD8 T cell activation is critical for greater renal disease severity in female chronic graft-vs.-host mice and is associated with increased splenic ICOS<sup>hi</sup> host CD4 T cells and IL-21 expression. *Clin. Immunol.* **136**, 61–73 (2010).
131. Kim, S. J., Zou, Y. R., Goldstein, J., Reizis, B. & Diamond, B. Tolerogenic function of Blimp-1 in dendritic cells. *J. Exp. Med.* **208**, 2193–2199 (2011).
132. Robinson, G. A. et al. Investigating sex differences in T regulatory cells from cisgender and transgender healthy individuals and patients with autoimmune inflammatory disease: a cross-sectional study. *Lancet Rheumatol.* **4**, e710–e724 (2022).
133. Afshan, G., Afzal, N. & Qureshi, S. CD4<sup>+</sup>CD25<sup>hi</sup> regulatory T cells in healthy males and females mediate gender difference in the prevalence of autoimmune diseases. *Clin. Lab.* **58**, 567–571 (2012).
134. Arruvito, L., Sanz, M., Banham, A. H. & Fainboim, L. Expansion of CD4<sup>+</sup>CD25<sup>+</sup> and FOXP3<sup>+</sup> regulatory T cells during the follicular phase of the menstrual cycle: implications for human reproduction. *J. Immunol.* **178**, 2572–2578 (2007).
135. Cho, J. et al. Sex bias in experimental immune-mediated, drug-induced liver injury in BALB/c mice: suggested roles for Tregs, estrogen, and IL-6. *PLoS ONE* **8**, e61186 (2013).
136. Peterson, R. A. Regulatory T-cells: diverse phenotypes integral to immune homeostasis and suppression. *Toxicol. Pathol.* **40**, 186–204 (2012).
137. Hussain, S., Kirwin, S. J. & Stohman, S. A. Increased T regulatory cells lead to development of Th2 immune response in male SJL mice. *Autoimmunity* **44**, 219–228 (2011).
138. Yee Mon, K. J. et al. Differential sensitivity to IL-12 drives sex-specific differences in the CD8<sup>+</sup> T cell response to infection. *Immunohorizons* **3**, 121–132 (2019).
139. Rukavina, D. et al. Age-related decline of perforin expression in human cytotoxic T lymphocytes and natural killer cells. *Blood* **92**, 2410–2420 (1998).
140. Yang, C. et al. Androgen receptor-mediated CD8<sup>+</sup> T cell stemness programs drive sex differences in antitumor immunity. *Immunity* **55**, 1268–1283.e9 (2022).
141. Kwon, H. et al. Androgen conspires with the CD8<sup>+</sup> T cell exhaustion program and contributes to sex bias in cancer. *Sci. Immunol.* **7**, eabq2630 (2022).
142. Mohanram, V. et al. B cell responses associated with vaccine-induced delayed SIV<sub>mac251</sub> acquisition in female rhesus macaques. *J. Immunol.* **197**, 2316–2324 (2016).
143. Tuero, I. et al. Mucosal B cells are associated with delayed SIV acquisition in vaccinated female but not male rhesus macaques following SIV<sub>mac251</sub> rectal challenge. *PLoS Pathog.* **11**, e1005101 (2015).
144. Zivkovic, I. et al. Sex bias in mouse humoral immune response to influenza vaccine depends on the vaccine type. *Biologicals* **52**, 18–24 (2018).
145. Zivkovic, I. et al. Sexual diergism in antibody response to whole virus trivalent inactivated influenza vaccine in outbred mice. *Vaccine* **33**, 5546–5552 (2015).
146. Ursin, R. L. et al. Greater breadth of vaccine-induced immunity in females than males is mediated by increased antibody diversity in germinal center B cells. *mBio* **13**, e0183922 (2022).
147. Engler, R. J. et al. Half- vs full-dose trivalent inactivated influenza vaccine (2004–2005): age, dose, and sex effects on immune responses. *Arch. Intern. Med.* **168**, 2405–2414 (2008).
148. Furman, D. et al. Systems analysis of sex differences reveals an immunosuppressive role for testosterone in the response to influenza vaccination. *Proc. Natl. Acad. Sci. USA* **111**, 869–874 (2014).
149. Potluri, T. et al. Age-associated changes in the impact of sex steroids on influenza vaccine responses in males and females. *NPJ Vaccines* **4**, 29 (2019).
150. Nguyen, D. C. et al. 17 $\beta$ -Estradiol restores antibody responses to an influenza vaccine in a postmenopausal mouse model. *Vaccine* **29**, 2515–2518 (2011).
151. Jones, B. G. et al. Binding of estrogen receptors to switch sites and regulatory elements in the immunoglobulin heavy chain locus of activated B cells suggests a direct influence of estrogen on antibody expression. *Mol. Immunol.* **77**, 97–102 (2016).
152. Gearhart, P. J. Immunology: the roots of antibody diversity. *Nature* **419**, 29–31 (2002).
153. Hurwitz, J. L. et al. Hotspots for vitamin-steroid-thyroid hormone response elements within switch regions of immunoglobulin heavy chain loci predict a direct influence of vitamins and hormones on B cell class switch recombination. *Viral Immunol.* **29**, 132–136 (2016).
154. Cunningham, M. A., Wirth, J. R., Naga, O., Eudaly, J. & Gilkeson, G. S. Estrogen receptor alpha binding to ERE is required for full Tlr7- and Tlr9-induced inflammation. *SOJ Immunol.* **2**, 4 (2014).
155. Lundell, A. C. et al. Higher B-cell activating factor levels at birth are positively associated with maternal dairy farm exposure and negatively related to allergy development. *J. Allergy Clin. Immunol.* **136**, 1074–1082.e3 (2015).
156. Hao, Y., O'Neill, P., Naradikian, M. S., Scholz, J. L. & Cancro, M. P. A B-cell subset uniquely responsive to innate stimuli accumulates in aged mice. *Blood* **118**, 1294–1304 (2011).
157. Rubtsov, A. V. et al. Toll-like receptor 7 (TLR7)-driven accumulation of a novel CD11c<sup>+</sup> B-cell population is important for the development of autoimmunity. *Blood* **118**, 1305–1315 (2011).
158. Rubtsov, A. V., Rubtsova, K., Kappler, J. W. & Marrack, P. TLR7 drives accumulation of ABCs and autoantibody production in autoimmune-prone mice. *Immunol. Res.* **55**, 210–216 (2013).
159. Fairweather, D., Frisnacho-Kiss, S. & Rose, N. R. Sex differences in autoimmune disease from a pathological perspective. *Am. J. Pathol.* **173**, 600–609 (2008).
160. Whitacre, C. C. Sex differences in autoimmune disease. *Nat. Immunol.* **2**, 777–780 (2001).
161. Handel, A. E., Handunnetthi, L., Ebers, G. C. & Ramagopalan, S. V. Type 1 diabetes mellitus and multiple sclerosis: common etiological features. *Nat. Rev. Endocrinol.* **5**, 655–664 (2009).
162. Ziegler, A. G. et al. Seroconversion to multiple islet autoantibodies and risk of progression to diabetes in children. *JAMA* **309**, 2473–2479 (2013).
163. Ucciferri, C. C. & Dunn, S. E. Effect of puberty on the immune system: relevance to multiple sclerosis. *Front. Pediatrics* **10**, 1059083 (2022).
164. Dalal, M., Kim, S. & Voskuhl, R. R. Testosterone therapy ameliorates experimental autoimmune encephalomyelitis and induces a T helper 2 bias in the autoantigen-specific T lymphocyte response. *J. Immunol.* **159**, 3–6 (1997).
165. Makino, S., Kunimoto, K., Muraoka, Y. & Katagiri, K. Effect of castration on the appearance of diabetes in NOD mouse. *Jikken Dobutsu* **30**, 137–140 (1981).
166. Fox, H. S. Androgen treatment prevents diabetes in nonobese diabetic mice. *J. Exp. Med.* **175**, 1409–1412 (1992).
167. Gold, S. M., Chalifoux, S., Giesser, B. S. & Voskuhl, R. R. Immune modulation and increased neurotrophic factor production in multiple sclerosis patients treated with testosterone. *J. Neuroinflammation* **5**, 32 (2008).
168. Markle, J. G. et al. Sex differences in the gut microbiome drive hormone-dependent regulation of autoimmunity. *Science* **339**, 1084–1088 (2013).
169. Yurkovetskiy, L. et al. Gender bias in autoimmunity is influenced by microbiota. *Immunity* **39**, 400–412 (2013).
170. Ahn, J. J. et al. Puberty in females enhances the risk of an outcome of multiple sclerosis in children and the development of central nervous system autoimmunity in mice. *Mult. Scler.* **21**, 735–748 (2015).
171. Bebo, B. F. Jr et al. Low-dose estrogen therapy ameliorates experimental autoimmune encephalomyelitis in two different inbred mouse strains. *J. Immunol.* **166**, 2080–2089 (2001).
172. McClain, M. A. et al. Pregnancy suppresses experimental autoimmune encephalomyelitis through immunoregulatory cytokine production. *J. Immunol.* **179**, 8146–8152 (2007).
173. Confavreux, C., Hutchinson, M., Hours, M. M., Cortinovis-Tourniaire, P. & Moreau, T. Rate of pregnancy-related relapse in multiple sclerosis. Pregnancy in Multiple Sclerosis Group. *N. Engl. J. Med.* **339**, 285–291 (1998).
174. Logel, S. N., Bekx, M. T. & Rehm, J. L. Potential association between type 1 diabetes mellitus and gender dysphoria. *Pediatr. Diabetes* **21**, 266–270 (2020).
175. Pakpoor, J., Wotton, C. J., Schmierer, K., Giovannoni, G. & Goldacre, M. J. Gender identity disorders and multiple sclerosis risk: a national record-linkage study. *Mult. Scler.* **22**, 1759–1762 (2016).
176. Arbucl, M. R. et al. Development of autoantibodies before the clinical onset of systemic lupus erythematosus. *N. Engl. J. Med.* **349**, 1526–1533 (2003).
177. Eriksson, C. et al. Autoantibodies predate the onset of systemic lupus erythematosus in northern Sweden. *Arthritis Res. Ther.* **13**, R30 (2011).
178. Wither, J. et al. Presence of an interferon signature in individuals who are anti-nuclear antibody positive lacking a systemic autoimmune rheumatic disease diagnosis. *Arthritis Res. Ther.* **19**, 41 (2017).
179. Hiraki, L. T. et al. Prevalence, incidence, and demographics of systemic lupus erythematosus and lupus nephritis from 2000 to 2004 among children in the US Medicaid beneficiary population. *Arthritis Rheum.* **64**, 2669–2676 (2012).
180. Scofield, R. H. et al. Klinefelter's syndrome (47,XXY) in male systemic lupus erythematosus patients: support for the notion of a gene-dose effect from the X chromosome. *Arthritis Rheum.* **58**, 2511–2517 (2008).
181. Cooney, C. M. et al. 46,X,del(X)(q13) Turner's syndrome women with systemic lupus erythematosus in a pedigree multiplex for SLE. *Genes. Immun.* **10**, 478–481 (2009).
182. Syrett, C. M. et al. Altered X-chromosome inactivation in T cells may promote sex-biased autoimmune diseases. *JCI Insight* **4**, e126751 (2019).
183. Hewagama, A. et al. Overexpression of X-linked genes in T cells from women with lupus. *J. Autoimmun.* **41**, 60–71 (2013).
184. Roubinian, J. R., Papoian, R. & Talal, N. Androgenic hormones modulate autoantibody responses and improve survival in murine lupus. *J. Clin. Invest.* **59**, 1066–1070 (1977).
185. Costenbader, K. H., Feskanich, D., Stampfer, M. J. & Karlson, E. W. Reproductive and menopausal factors and risk of systemic lupus erythematosus in women. *Arthritis Rheum.* **56**, 1251–1262 (2007).
186. Bernier, M. O., Mikaeloff, Y., Hudson, M. & Suissa, S. Combined oral contraceptive use and the risk of systemic lupus erythematosus. *Arthritis Rheum.* **61**, 476–481 (2009).
187. Petri, M. et al. Combined oral contraceptives in women with systemic lupus erythematosus. *N. Engl. J. Med.* **353**, 2550–2558 (2005).
188. Tomer, Y., Viegas, O. A., Swissa, M., Koh, S. C. & Shoenfeld, Y. Levels of lupus autoantibodies in pregnant SLE patients: correlations with disease activity and pregnancy outcome. *Clin. Exp. Rheumatol.* **14**, 275–280 (1996).
189. Panchanathan, R. & Choubey, D. Murine BAFF expression is up-regulated by estrogen and interferons: implications for sex bias in the development of autoimmunity. *Mol. Immunol.* **53**, 15–23 (2013).
190. Cohen-Solal, J. F. et al. Hormonal regulation of B-cell function and systemic lupus erythematosus. *Lupus* **17**, 528–532 (2008).



191. Schwartzman-Morris, J. & Putterman, C. Gender differences in the pathogenesis and outcome of lupus and of lupus nephritis. *Clin. Dev. Immunol.* **2012**, 604892 (2012).
192. Ricardo, A. C. et al. Sex-related disparities in CKD progression. *J. Am. Soc. Nephrol.* **30**, 137–146 (2019).
193. Carrero, J. J., Hecking, M., Chesnaye, N. C. & Jager, K. J. Sex and gender disparities in the epidemiology and outcomes of chronic kidney disease. *Nat. Rev. Nephrol.* **14**, 151–164 (2018).
194. Deng, W. et al. Gender-related differences in clinicopathological characteristics and renal outcomes of Chinese patients with IgA nephropathy. *BMC Nephrol.* **19**, 31 (2018).
195. Connelly, P. J., Currie, G. & Delles, C. Sex differences in the prevalence, outcomes and management of hypertension. *Curr. Hypertens. Rep.* **24**, 185–192 (2022).
196. Baylis, C. Sexual dimorphism: the aging kidney, involvement of nitric oxide deficiency, and angiotensin II overactivity. *J. Gerontol. A Biol. Sci. Med. Sci.* **67**, 1365–1372 (2012).
197. Elmarakby, A. A. & Sullivan, J. C. Sex differences in hypertension: lessons from spontaneously hypertensive rats (SHR). *Clin. Sci.* **135**, 1791–1804 (2021).
198. Tipton, A. J., Baban, B. & Sullivan, J. C. Female spontaneously hypertensive rats have greater renal anti-inflammatory T lymphocyte infiltration than males. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **303**, R359–R367 (2012).
199. Zimmerman, M. A., Baban, B., Tipton, A. J., O'Connor, P. M. & Sullivan, J. C. Chronic ANG II infusion induces sex-specific increases in renal T cells in Sprague-Dawley rats. *Am. J. Physiol. Ren. Physiol.* **308**, F706–F712 (2015).
200. Ji, H. et al. Sex-specific T-cell regulation of angiotensin II-dependent hypertension. *Hypertension* **64**, 573–582 (2014).
201. Rodriguez-Isturbe, B. et al. Reduction of renal immune cell infiltration results in blood pressure control in genetically hypertensive rats. *Am. J. Physiol. Ren. Physiol.* **282**, F191–F201 (2002).
202. Sandberg, K., Ji, H. & Hay, M. Sex-specific immune modulation of primary hypertension. *Cell Immunol.* **294**, 95–101 (2015).
203. Follow, D. P. et al. Sex differences in T-lymphocyte tissue infiltration and development of angiotensin II hypertension. *Hypertension* **64**, 384–390 (2014).
204. Wu, C. et al. Induction of pathogenic TH17 cells by inducible salt-sensing kinase SGK1. *Nature* **496**, 513–517 (2013).
205. Drummond, G. R., Vinh, A., Guzik, T. J. & Sobey, C. G. Immune mechanisms of hypertension. *Nat. Rev. Immunol.* **19**, 517–532 (2019).
206. Mikolajczyk, T. P. & Guzik, T. J. Adaptive immunity in hypertension. *Curr. Hypertens. Rep.* **21**, 68 (2019).
207. Brinson, K. N. et al. Female SHR have greater blood pressure sensitivity and renal T cell infiltration following chronic NOS inhibition than males. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **305**, R701–R710 (2013).
208. Belanger, K. M. et al. Greater T regulatory cells in females attenuate DOCA-salt-induced increases in blood pressure versus males. *Hypertension* **75**, 1615–1623 (2020).
209. Fehrenbach, D. J. & Mattson, D. L. Inflammatory macrophages in the kidney contribute to salt-sensitive hypertension. *Am. J. Physiol. Ren. Physiol.* **318**, F544–F548 (2020).
210. Sullivan, J. C., Semprun-Prieto, L., Boesen, E. I., Pollock, D. M. & Pollock, J. S. Sex and sex hormones influence the development of albuminuria and renal macrophage infiltration in spontaneously hypertensive rats. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **293**, R1573–R1579 (2007).
211. Echem, C. et al. Mitochondrial DNA: a new driver for sex differences in spontaneous hypertension. *Pharmacol. Res.* **144**, 142–150 (2019).
212. Fehrenbach, D. J. et al. Sexual dimorphic role of CD14 (cluster of differentiation 14) in salt-sensitive hypertension and renal injury. *Hypertension* **77**, 228–240 (2021).
213. Basile, D. P., Abais-Battad, J. M. & Mattson, D. L. Contribution of Th17 cells to tissue injury in hypertension. *Curr. Opin. Nephrol. Hypertens.* **30**, 151–158 (2021).
214. Lindsey, S. H., Yamaleyeva, L. M., Brosnihan, K. B., Gallagher, P. E. & Chappell, M. C. Estrogen receptor GPR30 reduces oxidative stress and proteinuria in the salt-sensitive female mRen2.Lewis rat. *Hypertension* **58**, 665–671 (2011).
215. Gohar, E. Y. et al. Evidence for G-protein-coupled estrogen receptor as a pronatriuretic factor. *J. Am. Heart Assoc.* **9**, e015110 (2020).
216. Ogola, B. O. et al. G protein-coupled estrogen receptor protects from angiotensin II-induced increases in pulse pressure and oxidative stress. *Front. Endocrinol.* **10**, 586 (2019).
217. Scully, E. P., Haverfield, J., Ursin, R. L., Tannenbaum, C. & Klein, S. L. Considering how biological sex impacts immune responses and COVID-19 outcomes. *Nat. Rev. Immunol.* **20**, 442–447 (2020).
218. Scully, E. P. et al. Sex and gender differences in testing, hospital admission, clinical presentation, and drivers of severe outcomes from COVID-19. *Open. Forum Infect. Dis.* **8**, ofab448 (2021).
219. Kharroubi, S. A. & Diab-El-Harake, M. Sex-differences in COVID-19 diagnosis, risk factors and disease comorbidities: a large US-based cohort study. *Front. Public Health* **10**, 1029190 (2022).
220. Torres, C. et al. Identifying age- and sex-specific COVID-19 mortality trends over time in six countries. *Int. J. Infect. Dis.* **128**, 32–40 (2023).
221. Gebhard, C., Regitz-Zagrosek, V., Neuhauser, H. K., Morgan, R. & Klein, S. L. Impact of sex and gender on COVID-19 outcomes in Europe. *Biol. Sex. Differ.* **11**, 29 (2020).
222. Karlberg, J., Chong, D. S. & Lai, W. Y. Do men have a higher case fatality rate of severe acute respiratory syndrome than women do? *Am. J. Epidemiol.* **159**, 229–231 (2004).
223. Alghamdi, I. G. et al. The pattern of Middle East respiratory syndrome coronavirus in Saudi Arabia: a descriptive epidemiological analysis of data from the Saudi Ministry of Health. *Int. J. Gen. Med.* **7**, 417–423 (2014).
224. Klein, S. L. Sex influences immune responses to viruses, and efficacy of prophylaxis and treatments for viral diseases. *BioEssays* **34**, 1050–1059 (2012).
225. vom Steeg, L. G. & Klein, S. L. SexX matters in infectious disease pathogenesis. *PLoS Pathog.* **12**, e1005374 (2016).
226. Robinson, D. P., Lorenzo, M. E., Jian, W. & Klein, S. L. Elevated 17 $\beta$ -estradiol protects females from influenza A virus pathogenesis by suppressing inflammatory responses. *PLoS Pathog.* **7**, e1002149 (2011).
227. Robinson, D. P. et al. Sex chromosome complement contributes to sex differences in Coxsackievirus B3 but not Influenza A virus pathogenesis. *Biol. Sex. Differ.* **2**, 8 (2011).
228. Robinson, D. P., Hall, O. J., Nilles, T. L., Bream, J. H. & Klein, S. L. 17 $\beta$ -Estradiol protects females against influenza by recruiting neutrophils and increasing virus-specific CD8 T cell responses in the lungs. *J. Virol.* **88**, 4711–4720 (2014).
229. Vom Steeg, L. G. et al. Age and testosterone mediate influenza pathogenesis in male mice. *Am. J. Physiol. Lung Cell Mol. Physiol.* **311**, L1234–L1244 (2016).
230. Lorenzo, M. E. et al. Antibody responses and cross protection against lethal influenza A viruses differ between the sexes in C57BL/6 mice. *Vaccine* **29**, 9246–9255 (2011).
231. Vermillion, M. S. et al. Production of amphiregulin and recovery from influenza is greater in males than females. *Biol. Sex. Differ.* **9**, 24 (2018).
232. Giurgea, L. T. et al. Sex differences in influenza: the challenge study experience. *J. Infect. Dis.* **225**, 715–722 (2022).
233. Larcombe, A. N. et al. Sexual dimorphism in lung function responses to acute influenza A infection. *Influenza Other Respir. Viruses* **5**, 334–342 (2011).
234. Hoffmann, J. et al. Sex differences in H7N9 influenza A virus pathogenesis. *Vaccine* **33**, 6949–6954 (2015).
235. Monticelli, L. A. et al. Innate lymphoid cells promote lung-tissue homeostasis after infection with influenza virus. *Nat. Immunol.* **12**, 1045–1054 (2011).
236. Vermillion, M. S., Ursin, R. L., Attreed, S. E. & Klein, S. L. Estriol reduces pulmonary immune cell recruitment and inflammation to protect female mice from severe influenza. *Endocrinology* **159**, 3306–3320 (2018).
237. Robinson, D. P. & Klein, S. L. Pregnancy and pregnancy-associated hormones alter immune responses and disease pathogenesis. *Horm. Behav.* **62**, 263–271 (2012).
238. Hall, O. J. et al. Progesterone-based contraceptives reduce adaptive immune responses and protection against sequential influenza A virus infections. *J. Virol.* **91**, e02160-16 (2017).
239. Vom Steeg, L. G. et al. Androgen receptor signaling in the lungs mitigates inflammation and improves the outcome of influenza in mice. *PLoS Pathog.* **16**, e1008506 (2020).
240. Maenosono, R. et al. Recipient sex and estradiol levels affect transplant outcomes in an age-specific fashion. *Am. J. Transpl.* **21**, 3239–3255 (2021).
241. Lepeyre, F. et al. Association of sex with risk of kidney graft failure differs by age. *J. Am. Soc. Nephrol.* **28**, 3014–3023 (2017).
242. Vinson, A. J. et al. Age-dependent sex differences in graft loss after kidney transplantation. *Transplantation* **106**, 1473–1484 (2022).
243. Hickey, K. T. et al. Clinical and gender differences in heart transplant recipients in the NEW HEART study. *Eur. J. Cardiovasc. Nurs.* **16**, 222–229 (2017).
244. Foster, B. J. et al. Differences in heart graft survival by recipient sex. *Transpl. Direct* **7**, e749 (2021).
245. Dziarmaga, R. et al. Age- and sex-mediated differences in T lymphocyte populations of kidney transplant recipients. *Pediatr. Transpl.* **26**, e14150 (2022).
246. Magnussen, C. et al. Sex differences in preformed panel-reactive antibody levels and outcomes in patients undergoing heart transplantation. *Clin. Transpl.* **33**, e13572 (2019).
247. Redfield, R. R. et al. The mode of sensitization and its influence on allograft outcomes in highly sensitized kidney transplant recipients. *Nephrol. Dial. Transpl.* **31**, 1746–1753 (2016).
248. Porrett, P. M. Biologic mechanisms and clinical consequences of pregnancy alloimmunization. *Am. J. Transpl.* **18**, 1059–1067 (2018).
249. Durgam, S. S., Alegre, M. L. & Chong, A. S. Toward an understanding of allogeneic conflict in pregnancy and transplantation. *J. Exp. Med.* **219**, e20211493 (2022).
250. Bromberger, B. et al. Pregnancy-induced sensitization promotes sex disparity in living donor kidney transplantation. *J. Am. Soc. Nephrol.* **28**, 3025–3033 (2017).
251. Honger, G. et al. Frequency and determinants of pregnancy-induced child-specific sensitization. *Am. J. Transpl.* **13**, 746–753 (2013).
252. Higgins, R. et al. Pregnancy-induced HLA antibodies respond more vigorously after renal transplantation than antibodies induced by prior transplantation. *Hum. Immunol.* **76**, 546–552 (2015).
253. Witowich, N. C., Beery, A. & Woodruff, T. A 10-year follow-up study of sex inclusion in the biological sciences. *Elife* **9**, e56344 (2020).
254. Sugimoto, C. R., Ahn, Y. Y., Smith, E., Macaluso, B. & Lariviere, V. Factors affecting sex-related reporting in medical research: a cross-disciplinary bibliometric analysis. *Lancet* **393**, 550–559 (2019).
255. Marquez, E. J. et al. Sexual-dimorphism in human immune system aging. *Nat. Commun.* **11**, 751 (2020).
256. U.S. Department of Health, Education, and Welfare. General considerations for the clinical evaluation of drugs. *FDA* <https://www.fda.gov/media/71495/download> (1977).
257. Parekh, A., Fadiran, E. O., Uhl, K. & Throckmorton, D. C. Adverse effects in women: implications for drug development and regulatory policies. *Expert. Rev. Clin. Pharmacol.* **4**, 453–466 (2011).
258. Pinn, V. W. The role of the NIH's Office of Research on Women's Health. *Acad. Med.* **69**, 698–702 (1994).

259. Beery, A. K. & Zucker, I. Sex bias in neuroscience and biomedical research. *Neurosci. Biobehav. Rev.* **35**, 565–572 (2011).
260. Clayton, J. A. Applying the new SABV (sex as a biological variable) policy to research and clinical care. *Physiol. Behav.* **187**, 2–5 (2018).
261. DiMarco, M., Zhao, H., Boulicault, M. & Richardson, S. S. Why “sex as a biological variable” conflicts with precision medicine initiatives. *Cell Rep. Med.* **3**, 100550 (2022).

## Acknowledgements

The authors thank K. L. Sandberg, Georgetown University, DC, USA, for critical review of the section on hypertension during the revisions process. S.E.D.'s work is supported by a Canadian Institutes of Health Research (CIHR) Sex and Gender Chair in Immunity and operating grants from CIHR, MS Canada and the BranchOut Foundation. S.L.K.'s work is supported by the Johns Hopkins Specialized Center of Research Excellence in Sex Differences (U54AG062333).

## Author contributions

The authors contributed equally to all aspects of the article.

## Competing interests

The authors declare no competing interests.

## Additional information

**Peer review information** *Nature Reviews Nephrology* thanks Annet Kirabo, Susan Kovats and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© Springer Nature Limited 2023



## OPEN ACCESS

## EDITED BY

Axel Cloeckaert,  
Institut National de recherche pour  
l'agriculture, l'alimentation et l'environnement  
(INRAE), France

## REVIEWED BY

Madjid Morsli,  
Aix Marseille Université, France  
Vladimir Lazarevic,  
Hôpitaux Universitaires de Genève (HUG),  
Switzerland  
Jessica E. Manning,  
National Institute of Allergy and Infectious  
Diseases (NIH), United States

## \*CORRESPONDENCE

Jessica Galloway-Peña  
✉ jgallowaypena@cvm.tamu.edu

RECEIVED 14 March 2023

ACCEPTED 12 June 2023

PUBLISHED 28 June 2023

## CITATION

Batool M and Galloway-Peña J (2023) Clinical  
metagenomics—challenges and future  
prospects.

*Front. Microbiol.* 14:1186424.

doi: 10.3389/fmicb.2023.1186424

## COPYRIGHT

© 2023 Batool and Galloway-Peña. This is an  
open-access article distributed under the terms  
of the [Creative Commons Attribution License  
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction  
in other forums is permitted, provided the  
original author(s) and the copyright owner(s)  
are credited and that the original publication in  
this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted which  
does not comply with these terms.

# Clinical metagenomics— challenges and future prospects

Maliha Batool and Jessica Galloway-Peña\*

Department of Veterinary Pathobiology, College of Veterinary Medicine and Biomedical Sciences, Texas  
A&M University, College Station, TX, United States

Infections lacking precise diagnosis are often caused by a rare or uncharacterized pathogen, a combination of pathogens, or a known pathogen carrying undocumented or newly acquired genes. Despite medical advances in infectious disease diagnostics, many patients still experience mortality or long-term consequences due to undiagnosed or misdiagnosed infections. Thus, there is a need for an exhaustive and universal diagnostic strategy to reduce the fraction of undocumented infections. Compared to conventional diagnostics, metagenomic next-generation sequencing (mNGS) is a promising, culture-independent sequencing technology that is sensitive to detecting rare, novel, and unexpected pathogens with no preconception. Despite the fact that several studies and case reports have identified the effectiveness of mNGS in improving clinical diagnosis, there are obvious shortcomings in terms of sensitivity, specificity, costs, standardization of bioinformatic pipelines, and interpretation of findings that limit the integration of mNGS into clinical practice. Therefore, physicians must understand the potential benefits and drawbacks of mNGS when applying it to clinical practice. In this review, we will examine the current accomplishments, efficacy, and restrictions of mNGS in relation to conventional diagnostic methods. Furthermore, we will suggest potential approaches to enhance mNGS to its maximum capacity as a clinical diagnostic tool for identifying severe infections.

## KEYWORDS

metagenomics, sepsis, diagnostics, meningitis, next-generation sequencing

## 1. Introduction

### 1.1. Overview of the limitations of routine diagnostics for pathogen detection

Conventionally, the clinical detection of pathogens is based on the isolation and cultivation of organisms (Fournier et al., 2014). Once cultivated, these organisms are typically characterized using biochemical tests, mass spectrometry, nuclear magnetic resonance (NMR) spectrometry, nucleic acid amplification, or immunological testing (Carroll et al., 2019). Culture-dependent methods are considered the “gold standard” for diagnosis of infectious diseases in clinics but it may take several days to weeks to cultivate slow-growing organisms. Also, prior exposure to antibiotics can impair the sensitivity of culturing, thus missing cases of treatable diseases (Govender et al., 2021).

Polymerase chain reaction (PCR) is a widely used molecular diagnostic method in clinical laboratories that can rapidly detect the presence or absence of DNA and RNA from a clinical specimen without the need for microbial cultivation (Carroll et al., 2019). PCR-based tests have been further developed into real-time PCR, allowing the amplification, quantification of expression, and thus identification of specific pathogen genetic content with high sensitivity and

specificity. However, PCR-based methodologies typically detect the presence or absence of a single gene at a time, offering low sensitivity, and potentially providing false negatives in cases containing low gene copy numbers (Huanyu Wang, 2021). To enhance the diagnostic capacity of PCR, multiplex PCR was developed to allow the simultaneous detection of multiple targets in a single PCR reaction (Huanyu Wang, 2021), although, it requires prior knowledge about pathogens of interest in order to identify them (Gu et al., 2021). Broad-range PCR is another effective method for hypothesis-independent detection of bacterial or fungal species, but has limitations. It has lower sensitivity than species-specific PCRs, cannot detect viral or parasitic infections, is only suitable for sterile bodily fluids and tissues, and can be more expensive than traditional methods (Rampini et al., 2011; Tkadlec et al., 2019; Aggarwal et al., 2020). Different PCR tests have varying diagnostic accuracy. In-house PCR tests are cheaper but require more time and training, while commercial PCR tests are automated and faster with higher sensitivity (Venter et al., 2019).

Alternatively, although antigen-detection is inexpensive and can be used in point-of-care setting due to promptness of the assay, diagnosis based on immunological tests are inherently less sensitive and may not provide accurate information (Govender et al., 2021). Furthermore, since it may take 1–2 weeks for antibodies to develop, antibody testing is not recommended for the diagnosis of acute disease (Govender et al., 2021).

In recent years, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) has become a tool of choice for bacterial and fungal identification in clinical laboratories (Dingle and Butler-Wu, 2013). Although identification can be provided in minutes, MALDI-TOF requires bacterial cultivation prior to analysis. Furthermore, it is not a quantitative approach and presents low specificity (Roux-Dalvai et al., 2019).

Peptide nucleic acid fluorescent *in situ* hybridization (PNA-FISH) is a recently introduced rapid and reliable method for the detection of bacteria and fungi responsible for blood stream infections. It provides more timely results compared to traditional culturing-based methods (Calderaro et al., 2014). PNA-FISH is validated by the U.S. Food and Drug Administration (FDA) for diagnosis of blood samples, however, it is not available to use at the tissue level (Weaver et al., 2019). A summary of the current technologies used in routine diagnostics in the clinical setting is provided in Table 1.

## 1.2. Overview of clinical need and advantages of metagenomics

Clinical metagenomics using next-generation sequencing (mNGS) has the potential to surpass the limitations of conventional diagnostics and make a seismic shift in the care of patients suffering from various infections (Simner et al., 2018). Unlike other diagnostic methods, mNGS does not require background knowledge of a suspected pathogen (Duan et al., 2021). mNGS can capture millions to billions of nucleic acids sequences at once and detect multiple organisms including novel pathogens that may be present in a clinical specimen (John et al., 2021). The time required for sample preparation, sequencing, and preliminary bioinformatic analysis depends on the nature of sequencing platform being used (Morsli et al., 2021a,b, 2022a,b). For example, newly available long-read

sequencing platforms, such as Oxford Nanopore sequencing, provide real time pathogen detection within minutes and additional information regarding genotyping and bacterial profiling in less than 6 h (Morsli et al., 2021a,b, 2022a,b). Also, Oxford Nanopore Technologies is currently the most prevalent and cost-effective mNGS platform in low- and middle-income countries (Yek et al., 2022). Clinical NGS includes two sequencing strategies: targeted amplicon sequencing and untargeted shotgun metagenomic sequencing. The targeted amplicon sequencing targets the universally conserved regions among bacteria (16S or 23S rRNA gene) or fungi and parasites (internal transcribed spacer (ITS), 18S rRNA, 28S rRNA gene) for pathogen detection (Salipante et al., 2013; Wagner et al., 2018). As an example, PCR-amplified 16S rRNA gene sequencing targets and amplifies one or more selected hypervariable regions (V1–V9) of the 16S rRNA gene. However, the choice of a particular hypervariable region targeted in 16S rRNA gene sequencing appears to be one of the biggest factors underlying technical variation in microbiome composition (Hiergeist et al., 2015; Tremblay et al., 2015; Gohl et al., 2016).

Unlike targeted amplicon sequencing, which only targets specific genes or gene regions, shotgun metagenomic sequencing targets the entire genetic content of a clinical sample, thus permitting the detection of all potential pathogens (Chiu and Miller, 2019). The capability to simultaneously identify viruses, bacteria, fungi, and parasites in a sample makes it broadly appealing for co-infection cases (Chen et al., 2021a). Furthermore, the amount of information derived from shotgun mNGS sequencing can potentially be used for additional analyses, such as antibiotic resistance profiling, virulence gene information, metabolic function profiling, and analyses of human host responses via transcriptome profiling (Chiu and Miller, 2019).

## 1.3. Commercially available clinical metagenomics platforms

Recently mNGS testing and analyses have become commercially available. For example, Charles Chiu and colleagues from the University of California, San Francisco (UCSF) are pioneers in the development of mNGS testing for the diagnosis of central nervous system (CNS) infections (Wilson et al., 2014). In 2014, the first use of mNGS was reported for the diagnosis of neuro-leptospirosis on CSF from a 14-year-old boy presenting the signs of meningoencephalitis (Wilson et al., 2014). This was the first report demonstrating the use of mNGS with medically actionable information and successful clinical diagnosis that led to the appropriate treatment of the patient (Wilson et al., 2014). Since then, UCSF provides validated mNGS DNA and RNA testing for patients with meningitis and/or encephalitis. The UCSF diagnostic lab also offers mNGS DNA testing for patients with sepsis and disseminated infections (Chiu and Miller, n.d.). UCSF software analyzes sequence reads, identifies those reads which align to pathogens in the GenBank database, and issues a report showing the presence of pathogens in a clinical sample, along with clinical interpretation. At least 66.7% of pathogens detected from CSF were true positives, and only 5.6% were found to be false positives (Miller et al., 2019). The mNGS test at UCSF for pathogen detection from CSF specimens showed a sensitivity of 86.1% and a specificity of 97.9%. The mNGS test for pathogen detection from plasma samples showed a sensitivity of 77% and specificity of 86%. The turnaround



TABLE 1 Routine diagnostic assays in a clinical setting for the detection of pathogens.

Technology	Manufacturer	Trade name	Detection time	Organisms	Limitations
<b>Respiratory infections</b>					
Multiplex PCR	Biofire Diagnostics, GenMark Diagnostics, Luminex, Roche Diagnostics	FilmArray, ePlex, VERIGENE, Lyra, Cobas Liat PCR system, NxTAG	15 min–2 h	Bacteria	<ul style="list-style-type: none"> <li>Requires a separate primer set for each target gene</li> <li>Low application efficiency</li> </ul>
				Viruses	
				Adenovirus, coronavirus, human metapneumovirus, human rhinovirus/enterovirus, influenza A, influenza B, parainfluenza virus, respiratory syncytial virus, influenza A/B	
Nuclear magnetic resonance (NMR)	T2 biosystems	T2SARS-CoV-2	<2h	SARS-CoV-2	<ul style="list-style-type: none"> <li>Magnetic field drift that may be detrimental to NMR spectra</li> </ul>
<b>Bloodstream infections</b>					
Multiplex PCR	Biofire Diagnostics, GenMark Diagnostics	FilmArray BCID, ePlex BCID	1–1.5 h	Gram-negative bacteria	<ul style="list-style-type: none"> <li>Requires background knowledge of suspected pathogen</li> </ul>
				<i>Acinetobacter baumannii</i> , <i>Bacteroides fragilis</i> , <i>Citrobacter</i> , <i>Escherichia coli</i> , <i>Enterobacter</i> spp., <i>Enterococcus</i> spp., <i>Fusobacterium</i> spp., <i>Haemophilus influenzae</i> , <i>Klebsiella</i> spp., <i>Neisseria meningitidis</i> , <i>Pseudomonas aeruginosa</i> , <i>Proteus</i> spp., <i>Serratia marcescens</i> , <i>Salmonella</i> spp., <i>Stenotrophomonas maltophilia</i>	
				Gram-positive bacteria	<ul style="list-style-type: none"> <li>Unable to detect pathogens containing low copy number of genes in a clinical sample</li> </ul>
				<i>Bacillus</i> spp., <i>Corynebacterium</i> spp., <i>Cutibacterium acnes</i> , <i>Enterococcus</i> spp., <i>Listeria monocytogenes</i> , <i>Micrococcus</i> spp., <i>Staphylococcus</i> spp., <i>Streptococcus</i> spp.	
Fungal pathogens	<i>Candida</i> spp., <i>Fusarium</i> spp., <i>Rhodotorula</i> spp.				
PNA-FISH	Accelerate Diagnostics, AdvanDx,	Accelerate Pheno, QuickFISH	20 min–1.5 h	Gram-negative bacteria	<ul style="list-style-type: none"> <li>Detects limited number of targets</li> </ul>
				<i>Acinetobacter baumannii</i> , <i>Citrobacter</i> spp., <i>Escherichia coli</i> , <i>Enterobacter</i> spp., <i>Klebsiella</i> spp., <i>Pseudomonas aeruginosa</i> , <i>Proteus</i> spp., <i>Serratia marcescens</i>	
				Gram-positive bacteria	
				<i>Enterococcus</i> spp., <i>Staphylococcus</i> spp., <i>Streptococcus</i> spp.,	
				Fungal pathogens	<i>Candida</i> spp.
MALDI-TOF	BioMerieux and Bruker	VITEK MS, MALDI Biotyper	30 min		<ul style="list-style-type: none"> <li>Has proprietary databases</li> <li>Limited in the differentiation of closely related species</li> </ul>

(Continued)

TABLE 1 (Continued)

Technology	Manufacturer	Trade name	Detection time	Organisms	Limitations
rRNA/PCR	Karius	Karius test		Detects greater than 1,000 pathogens	<ul style="list-style-type: none"> <li>Give false positives</li> </ul>
Nuclear magnetic resonance	T2 Biosystems	T2Bacteria,		Gram positive and gram-negative bacteria	<ul style="list-style-type: none"> <li>Low limit of detection</li> </ul>
		T2Candida		<i>Escherichia coli</i> , <i>Pseudomonas aeruginosa</i> , <i>Klebsiella pneumoniae</i> , <i>Staphylococcus aureus</i>	
				Fungal pathogens <i>Candida</i> spp.	
<b>Central nervous system</b>					
Multiplex PCR	Biofire diagnostics	FilmArray BCID	1 h	Gram positive and gram-negative bacteria	<ul style="list-style-type: none"> <li>Cannot differentiate between live and dead organisms</li> </ul>
				<i>Streptococcus species</i> , <i>Haemophilus influenzae</i> , <i>Neisseria meningitidis</i>	
				Viruses	
				Cytomegalovirus, enterovirus, herpes simplex viruses 1 and 2, human herpesvirus 6, human parechovirus, and varicella-zoster virus	

time from shipping samples to delivery of a report is generally 1–2 weeks (Chiu and Miller, n.d.).

The Karius test (Karius, California, United States) is another example of how mNGS is useful for the diagnosis of bloodstream infections (BSIs) and sepsis (Blauwkamp et al., 2019). The Karius test involves extraction of cell-free DNA (cfDNA) from plasma, then a sequencing library is created and sequenced using Illumina technology. The sequence data is compared to an internal reference database encompassing a number of microbial genomes (Blauwkamp et al., 2019). A published study by Thair et al. confirmed that the Karius test detected approximately three times more positive cases than culture-based detection (Thair et al., 2017). However, the limitation is that the test can give false positive results. The Arizona-based Fry Laboratories also provides DNA sequencing diagnostic services for cutaneous, gastrointestinal, hematologic, musculoskeletal, and pulmonary infections (Fry, n.d.).

The Beijing Genomic Institute (BGI Genomics), a China-based company, is one of the largest companies that provide clinical mNGS services for the detection of pathogens causing respiratory infections such as Coronavirus and other pathogenic microorganisms (Jiang, n.d.). The sequencing services by the Zhejiang, China-based IngeniGen XunMinKang Biotechnology company also provide the detection of undiagnosed pathogens in patients with respiratory diseases (Wang et al., 2019; Li et al., 2022).

## 2. Detection of pathogens via clinical metagenomics

mNGS is an unbiased culture-independent and hypothesis-free sequencing technology that has shown tremendous clinical application particularly in the diagnosis of CNS infections, bloodstream infections, and respiratory tract infections (Blauwkamp et al., 2019; Miller et al., 2019; Wilson et al., 2019;

Chen H. et al., 2020; Hasan et al., 2020; Haston et al., 2020; Li et al., 2020; Chen et al., 2021b; Hogan et al., 2021; Jing et al., 2021; Liu et al., 2021; Mu et al., 2021; Pollock et al., 2021; Zhou et al., 2021; Deng et al., 2022; Fu et al., 2022; Guo et al., 2022; Li et al., 2022; Wang et al., 2022; Zhang et al., 2022; Morsli et al., 2022b). Examples of recent applications of mNGS in the diagnosis of these infections are provided in Table 2. Below is a brief overview of the areas where mNGS has made considerable impact and the implications.

### 2.1. Respiratory infections

Pneumonia is considered among the top 10 causes of death in the United States, especially among immunocompromised patients such as those with hematologic malignancy or undergoing hematopoietic stem cell transplant (Lippert et al., 2022). The identification of the causative agent of pneumonia is difficult and often inaccurate due to the pathogen diversity, heterogeneity of sampling, and limited detection methods (Buchan et al., 2022). Traditional molecular diagnosis for pneumonia is pathogen-specific but unreliable for novel or unexpected pathogens (Diao et al., 2022). The ability of mNGS to provide a comprehensive view of pathogens makes it useful in the diagnosis of unexplained pneumonia and disease of unknown etiology (Ramesh et al., 2019; Diao et al., 2022). Recently, mNGS has improved the diagnosis of pulmonary infections over traditional methods by detecting a broad range of organisms including bacteria, viruses and fungi in a number of recent investigations (Chen H. et al., 2020; Li et al., 2020; Mu et al., 2021; Deng et al., 2022; Li et al., 2022; Zhang et al., 2022). Remarkably, the causative agent was identified only by mNGS in two recent studies (Zhou et al., 2021; Guo et al., 2022). Importantly, mNGS led to the treatment modifications and guided treatment decisions for 127 patients with pulmonary infections (Mu et al., 2021; Zhou et al., 2021; Li et al., 2022). Moreover,

TABLE 2 Examples of the potential impact of clinical shotgun metagenomics (2019–2022) on infectious disease diagnosis.

Study details	Type of the study (n=subjects)	Samples and target population	Sequencing platform	Main findings
<b>Respiratory infections</b>				
Chen H. et al. (2020)	Prospective, observational study (n = 93)	Bronchoalveolar lavage fluid from patients with lower respiratory tract infections	Illumina Nextseq 550	The detection rate of mNGS for causative pathogen of lower respiratory infection was significantly higher (65% vs. 20%) than traditional culture method.
Li et al. (2020)	Prospective study (n = 121)	Lung biopsies from patients with peripheral pulmonary lesions and lung infection	BGISEQ-50	The percentage of mNGS-positive samples in radial endobronchial ultrasound (R-EBUS)-guided transbronchial lung biopsy (TBLB) was 78.8% that was significantly greater than TBLB (60.0%).
Zhou et al. (2021)	Multi-center, prospective, observational study (n = 159)	Bronchoalveolar lavage fluid from patients with pulmonary infections	Illumina NextSeq 550	mNGS detected more organisms (117 vs. 72) when compared with standard methods including bacteria (89 vs. 54), viruses (10 vs. 3), and fungi (18 vs. 15). Importantly, the bacteria known to cause pneumonia was detected only by mNGS that included <i>Haemophilus influenzae</i> , <i>Legionella pneumoniae</i> , <i>Mycobacterium avium</i> , <i>Mycobacteroides abscessus</i> , <i>Chlamydia psittaci</i> , and <i>Actinomyces</i> species. mNGS also led to the treatment modification for 59 patients.
Azar et al. (2021)	Prospective, observational study, (n = 30)	Bronchoalveolar lavage fluid from immunocompromised adults with pneumonia	Illumina NextSeq500 or NextSeq550	A combination of mNGS and conventional testing improved the diagnostic rate of pneumonia from 35% to 58%.
Chen S. et al. (2021)	Single center, retrospective, observation study (n = 408)	Blood, sputum, urine and bronchoalveolar lavage fluid from COVID-19 patients	BGISEQ-50	mNGS showed positive detection rate of 92.3% in bronchoalveolar lavage and 66.7% in sputum. Overall, mNGS results were comparable with conventional culture.
Chen Y. et al. (2021)	Retrospective study (n = 90)	Bronchoalveolar lavage fluid, transbronchial brushing from patients with focal pulmonary infections	Illumina Nextseq 550	The analysis of patients with focal pulmonary infections revealed sensitivity of mNGS in bronchoalveolar lavage fluid, transbronchial brushing group, and pathological specimen was 50%, 66.7%, and 90%, respectively.
Deng et al. (2022)	Retrospective, observational study (n = 103)	Bronchoalveolar lavage fluid from children with pneumonia	Illumina NextSeq CN500 sequencer	Out of 52 monomicrobial and 44 polymicrobial cases, mNGS detected 48 and 29 cases, respectively. Overall, the pathogen detection rate of mNGS was higher than conventional detection methods.
Zhang et al. (2022)	Retrospective, observational study (n = 47)	Bronchoalveolar lavage fluid from patients with lower respiratory tract infections	MGISEQ-2000	As compared to conventional culturing, mNGS increased the detection rate for causative pathogens of lower respiratory tract infections with a diagnostic sensitivity of 80% and specificity of 35.13%.
Xu et al. (2022)	Retrospective, observational study (n = 35)	Alveolar lavage fluid or venous blood from patients with severe psittacosis pneumonia	DA8600	mNGS detected DNA of <i>chlamydia psittaci</i> in alveolar lavage fluid of 30 patients and blood of 5 patients.
Pollock et al. (2021)	Single-center, proof-of-concept study (n = 30)	Plasma samples from patients with pulmonary tuberculosis	Illumina NextSeq 550	<i>Mycobacterium tuberculosis</i> cell-free DNA was detected from the plasma of 50% of pediatric and 60% of adult patients. Furthermore, it was also detected in an additional 25% of pediatric and 40% of adult patients when the relaxed research use statistical threshold was applied.
Mu et al. (2021)	Single-center, prospective study (n = 292)	Bronchoalveolar lavage fluid and sputum from patients with different kinds of lower respiratory tract infections	Nanopore	Compared with conventional testing, mNGS showed 96.6% sensitivity and 80% specificity and detected pathogens in 63 out of 161 culture-negative cases. Furthermore, mNGS proposed antibiotic de-escalation for 34 patients.

(Continued)

TABLE 2 (Continued)

Study details	Type of the study (n=subjects)	Samples and target population	Sequencing platform	Main findings
Li et al. (2022)	Single-center, prospective study (n = 138)	Bronchoalveolar lavage fluid from patients with pulmonary infections	Illumina Miniseq	mNGS improved diagnosis by detecting more pathogens such as bacteria (53 vs. 27) and viruses (16 vs. 1) than conventional methods. Importantly, mNGS led to the treatment modification for 34 out of 138 patients.
Guo et al. (2022)	Single-center, retrospective study (n = 121)	Bronchoalveolar lavage fluid from children with community-acquired pneumonia	Illumina Novaseq	The causative pathogens of pneumonia were only detected by mNGS. These organisms included <i>Streptococcus pneumoniae</i> , <i>Mycoplasma pneumoniae</i> , <i>Haemophilus influenzae</i> , Human bocavirus 1, and <i>Mycobacterium tuberculosis</i> . Moreover, mNGS identified 50% of human bocavirus-infected cases which were co-infected with other bacteria of respiratory origin.
<b>Bloodstream infections</b>				
Hogan et al. (2021)	Multicenter, retrospective study (n = 82)	Plasma samples from patients with suspicion of several infections	Illumina	The positivity rate for Karius-based mNGS was 61.0%. Of which 50% of cases were detected with monomicrobial infections and 50% of them were infected with 2 or more organisms. Overall, Karius-based mNGS showed a positive impact on 7.3% of cases, a negative impact on 3.7% of cases, and showed no impact on 86.6% of cases.
Blauwkamp et al. (2019)	Prospective study (n = 350)	Plasma samples from patients with clinical suspicion of sepsis	Illumina NextSeq 500	In contrast to culture, mNGS identified much more bacteria. 62 out of 166 samples were negative by traditional testing but sequencing identified these microorganisms in cell-free DNA.
Kalantar et al. (2022)	Prospective study (n = 221)	Blood and plasma from critically ill patients	Illumina Novaseq 6000	The pathogen detection in plasma by mNGS and traditional testing varied by organism. For example, mNGS showed 100% sensitivity for <i>Staphylococcus aureus</i> and <i>Escherichia coli</i> . However, mNGS missed the detection of <i>Streptococcus pyogenes</i> . Furthermore, the findings suggest that detection of a pathogen alone is not sufficient for sepsis diagnosis, instead when combined with hosts transcriptional profiling it may provide promising diagnostic utility.
Wang et al. (2022)	Retrospective, observational study (n = 435)	Blood, tissues, urine, sputum and different types of body fluids from patients with clinical suspicion of infections	Illumina NextSeq CN500 sequencer	The overall sensitivity of mNGS results were significantly higher than traditional methods. However, there was no difference in specificity of two methods. The sensitivity of mNGS for bronchoalveolar lavage fluid was 72.6% that was higher than blood that showed mNGS sensitivity of 39.3%.
Liu et al. (2021)	Prospective study (n = 24)	Blood samples from patients with hematological malignancies and sepsis	MGISEQ-200	The pathogen detection rate of mNGS was comparable with conventional testing for 9 out of 24 patients. However, for 10 patients, mNGS identified additional pathogens as compared to traditional methods most of the identified pathogens were viruses.
Jing et al. (2021)	Retrospective study (n = 209)	Blood samples from patients with suspected bloodstream infections	Illumina NextSeq 550	mNGS of plasma improved the clinical sensitivity (87.1%) and specificity (80.2%) as compared to conventional testing.
Fu et al. (2022)	Single center, retrospective study (n = 175)	Blood samples from patients with fever of unknown origin	BGISEQ-2000	mNGS increased the detection rate of new organisms in patients with fever of unknown origin by 22.9 and 19.79% than culture and standard detection methods, respectively. Specifically, it improved the detection rate of bloodstream infections by 38 and 32% respectively, as compared to culture and conventional testing.
<b>Central nervous system infections</b>				
Wilson et al. (2019)	Multicenter, prospective study (n = 204)	Severely ill pediatric and adult patients admitted to the intensive care unit	Illumina Hiseq	mNGS improved diagnosis over traditional methods of neurologic infections by identifying 22% (13 out of 58) of unique pathogens that were missed by clinical testing. The identification of these pathogens led to the treatment modification of 50% (7 out of 13) of these patients.

(Continued)



TABLE 2 (Continued)

Study details	Type of the study (n=subjects)	Samples and target population	Sequencing platform	Main findings
Miller et al. (2019)	Development and prospective study (n = 115)	CSF samples from patients with meningitis and/or encephalitis and patients with suspected neurological infections	Illumina Hiseq and Illumina MiSeq	For 95 samples, mNGS revealed 73% sensitivity and 99% specificity as compared to conventional testing. Moreover, for 20 CSF samples collected from pediatric patients 92% sensitivity and 96% specificity was observed relative to microbiological testing of CSF.
Hasan et al. (2020)	Retrospective study (n = 83)	Hospitalized children with suspected CNS infections	Illumina Miseq	In contrast to conventional methods, mNGS showed 100% diagnostic accuracy, 95% sensitivity, and 96% specificity for cerebrospinal fluid samples for hospitalized patients.
Morsli et al. (2022b)	Prospective and proof-of-concept study (n = 52)	Patients with community-acquired meningitis	MinION	Out of 52 subjects enrolled, 47 patients showed positive results on CSF samples via routine diagnostics and MinION sequencer. However, in addition to pathogen detection MinION sequencer provided additional information about genotype and antibiotic susceptibility of pathogens.
Haston et al. (2020)	Prospective study (n = 20)	Children with encephalitis of unidentified etiology	Illumina Miseq or NextSeq 500	mNGS detected sequence reads of pathogens such as <i>Mycoplasma bovis</i> , <i>Neisseria meningitidis</i> , parvovirus B19, and <i>Balamuthia mandrillaris</i> in 6 out of 20 patients. Furthermore, mNGS also detected some nonpathogenic organisms such as <i>Cladophialophora</i> species, human bocavirus, and tobacco mosaic virus. The patients with detectable pathogens via mNGS presented immune-mediated phenomena than patients for whom mNGS did not make any diagnosis.
Chen et al. (2021b)	Retrospective study (n = 88)	Patients suspected of encephalitis and meningitis	BGISEQ-50 and MGISEQ-2000	mNGS of cerebrospinal fluid detected pathogens in 56.81% (50 out of 88) of patients. The outcomes of mNGS helped in the treatment modification for 23.9% of patients and provided confidence in the continuation of original treatment for 34.1% of patients.

We used predefined filters to refine PubMed search on “classical article,” “clinical study,” “observational study,” “randomized controlled trial,” and “validation study” from 2019 to 2023. We included only retrospective or prospective clinical studies focusing on hospitalized patients using shotgun metagenomics. We have used the keys words “metagenomics AND respiratory infections” for respiratory diseases. Similarly, “metagenomics AND blood infections” and “metagenomics AND central nervous system infections” for blood stream infections and central nervous system infections, respectively. We also search these same terms in google scholar search and added clinical studies and studies focusing on hospitalized patients using shotgun metagenomics not found in PubMed.

mNGS detected 50% of cases coinfecting with bacteria of different respiratory origin in another study (Guo et al., 2022).

A mNGS approach can be superior to traditional methods for pathogen detection and confirmation of respiratory infections, particularly for *Mycobacterium tuberculosis* (Jin et al., 2022). *Mycobacterium tuberculosis*, can be quite challenging to detect, however, it has been shown in the last few years that mNGS could potentially be used as the first-line diagnostic test for tuberculosis. Karius-based mNGS testing of plasma samples detected suspected tuberculosis in 60% of adults and 50% of pediatric patients (Pollock et al., 2021).

Lastly, RNA viruses are also considered one of the primary causes of respiratory infections (Miller and Chiu, 2022). mNGS can detect a number of viruses that are usually not screened for in respiratory infections using routine diagnostic assays (Prachayangprecha et al., 2014; Bohl et al., 2022). It has shown good sensitivity and specificity compared to conventional testing and can identify viruses such as Influenza, Rhinovirus, and HIV (Jia et al., 2021). An additional advantage to mNGS is the potential to document and describe emerging, and re-emerging viral infections associated with outbreaks (Quer et al., 2022). For example, RNA-based viral metagenomics has detected the presence of novel human coronavirus variants from patients with respiratory symptoms (Wu et al., 2020; Castañeda-Mogollón et al., 2021).

## 2.2. Bloodstream infections

In 2017, it was estimated that 48.9 million cases and 11 million deaths were related to sepsis globally (Rudd et al., 2020). Thus, the early and accurate diagnosis of BSI is critical to initiate appropriate antibiotic therapy and for patient survival. Recent findings indicate sequencing microbial cfDNA using mNGS is a valuable approach for the detection of BSI pathogens when the conventional diagnostics fail to detect the etiological agent (Hogan et al., 2021; Eichenberger et al., 2022). A retrospective multi-center study utilizing the cfDNA and RNA showed that mNGS had a positive impact in 7.3% of cases, a negative impact in 3.7% of cases, and no impact in 86.6% of cases in patients with suspicion of multiple infections (Hogan et al., 2021). Another study applied mNGS on cfDNA in septic and non-septic intensive care unit (ICU) patients and was able to diagnose sepsis and predicted mortality as soon as the first day (Jing et al., 2022). Similarly, cfDNA of relevant pathogens was detected in the blood plasma of cystic fibrosis patients (Barrett et al., 2020). mNGS testing improved the detection rate of BSI in patients having fever of unknown origin or patients with suspected BSI from 38% to 87.1% when compared to conventional methods (Jing et al., 2021; Fu et al., 2022; Wang et al., 2022). However, no difference was observed in specificity between two methods for patients with clinical suspicion of infections (Wang et al., 2022). In one report, mNGS pathogen detection rate was comparable with routine diagnostics in

37% of cases (Liu et al., 2021). In some scenarios, the pathogen detection rate for mNGS varied by organism. For instance, mNGS was 100% sensitive for the detection of *Staphylococcus aureus* and *Escherichia coli*. However, mNGS test missed the presence of *Streptococcus pyogenes* (Kalantar et al., 2022). In contrast, 37% of BSI cases were found to be positive by only mNGS test in patients with clinical suspicion of sepsis in another study (Blauwkamp et al., 2019).

### 2.3. Central nervous system infections

Neuroinflammatory diseases such as meningitis and encephalitis can be diagnostically challenging due to the requirement of invasive procedures for CSF collection, limited availability/low volume of CNS samples, and difficulty of detection by traditional culture (Vetter et al., 2020; Heming et al., 2022; Mokhtari et al., 2022). Furthermore, meningoencephalitis is related with increased risk of morbidity and mortality and thus needs prompt diagnosis and disease management (Ramachandran and Wilson, 2020). CSF culture is considered the gold-standard method for the diagnosis of meningitis. However, prior antibiotic therapy may reduce the sensitivity of CSF cultures, increasing the possibility of false-negatives (Greenberg and Herrera, 2019). mNGS has potential to detect pathogens of unknown etiology as evidenced by clinical series demonstrating the success of mNGS in the detection of hard-to-diagnose CNS infection cases (Miller et al., 2019; Wilson et al., 2019; Hasan et al., 2020; Haston et al., 2020; Chen et al., 2021b; Morsli et al., 2022b). Recently published studies have confirmed the diagnostic sensitivity of mNGS ranging 22% to 95% in patients with CNS infections (Miller et al., 2019; Wilson et al., 2019; Hasan et al., 2020). mNGS testing guided treatment decisions and clinical actionable management for 34.1%–53% of patients in these reports (Wilson et al., 2019; Chen et al., 2021b). However, contaminants from skin flora can lead to false positive bacterial sequences in CSF specimens obtained by lumbar puncture.

## 3. Limitations, knowledge gaps, and potential solutions of clinical metagenomics

As demonstrated in the previous sections, mNGS is a promising diagnostic technology. However, in the present state of knowledge, metagenomics is not always well-positioned to assist clinicians in rapid clinical decision-making due to expertise required for sample preparation, sequencing, bioinformatic analysis, and the high variability in methodologies and interpretation.

Pathogen detection by mNGS depends on the proportion of pathogen sequences in the total sequencing library. Essentially, the diagnostic performance of mNGS is optimal when the sequencing library contains a nominal fraction of host DNA or there is an enrichment of pathogenic sequences (Olausson et al., 2022). While mNGS tends to be more sensitive than traditional methods as evidenced by several studies (Duan et al., 2021; Guo et al., 2022; Li et al., 2022; Zhu et al., 2022; Wang et al., 2023), one of the caveats to being more sensitive is that mNGS can pick up microbial contamination derived from the environment, containers, reagents, and colonizing microorganisms in the human body, thus giving false-positive results. The use of negative controls is recommended for

reagents and containers (López-Labrador et al., 2021). Moreover, depending on the sample source and bacterial load, in most cases the majority of reads in a mNGS data set can be derived from human DNA, while the proportion of pathogens tends to be very low (Gu et al., 2019). In order to overcome this challenge, depletion of host background DNA or targeted sequencing approaches such as depletion of abundant sequences by hybridization (DASH), and finding low-abundance sequences by hybridization (FLASH) in a combination of mNGS could be used (Gu et al., 2016; Hasan et al., 2016; Gu et al., 2019; Quan et al., 2019). For example, a positive selection probe-based system called virome capture sequencing platform for vertebrate viruses (VirCapSeq-VERT) has been established by a research group from the United States to increase the sensitivity of detecting viral sequences in clinical samples (Briese et al., 2015). Compact aggregation of targets for comprehensive hybridization (CATCH) is another available method to capture diverse targets from diverse patient samples (Viral Hemorrhagic Fever Consortium et al., 2019). The sequencing depth required to detect and characterize the genome of interest is a major influencer of metagenomics sensitivity. However, there is no consensus for how sequencing depth should be reported. For now, the choice of sequencing depth is dependent on budget and desired outcomes (Greninger, 2018). Retrospective experiments testing different human specimens with known infection at various sequencing depths are required to determine the ideal sequencing coverage for the diagnosis of different human specimens and infectious agents.

On that note, NGS tests do not equally detect all pathogens. For example, the detection rate of intracellular pathogens in clinical samples, such as a *Mycobacterium*, is relatively low as the amount of cell-free DNA of intracellular pathogen released into extracellular body fluids is relatively minute (Chen P. et al., 2020). The use of higher sequencing depth may help in identifying the presence of less abundant pathogens in a clinical sample. However, as a result of higher sequencing depth, a large amount of data would be generated, and thus, require more time for analysis. To avoid this challenge, rapid and more advanced bioinformatic tools must be developed (Chiu and Miller, 2019; Miller et al., 2020).

The human is a host to several commensal organisms, thus separating organisms associated with true bloodstream infection from transient gastrointestinal or oral flora in blood/plasma samples is an obstacle to the interpretation of mNGS results (Chiu and Miller, 2019; Chen et al., 2022). Retrospective studies using different threshold levels (such as a cut-off value for the abundance of pathogens, number of sequencing reads to detect a specific pathogen, and sequencing read normalization) are required to differentiate potential pathogens from commensal organisms. Moreover, the clinical significance of identified organisms needs to be further confirmed by conventional testing and the condition of the host (Chen et al., 2022). If a patient presents with an infection and receives appropriate curative therapy, mNGS can still detect lingering DNA from dead pathogens. It is unknown how long the detectable half-life of a pathogen is once the patient receives appropriate treatment and is conceivably circumstantial. A potential solution is the detection of RNA, whose abundance is directly correlated with the degree of gene transcription activity, thus, it can distinguish dead and live organisms in a clinical sample (d'Humières et al., 2021). Compared with DNA sequencing, a combination of DNA and RNA sequencing may have additional benefits (Arroyo Mühr et al., 2021). However, the detection of RNA through mNGS still has its challenges because of the

higher abundance of human-derived RNA in a clinical sample as well as the labile nature of RNA. To overcome this challenge, depletion of the host background is needed (Zheng et al., 2021).

Another limitation of the technology is the actual determination of what is detected (López-Labrador et al., 2021). For example, incomplete databases, mis-annotated sequences, databases containing contaminating organisms, bias in databases, and misclassification of organisms all affect the actual determination of the pathogen after sequencing (López-Labrador et al., 2021; Diao et al., 2022). Moreover, differences in pipelines, reproducibility, quality control, and workflow may lead to different and inaccurate pathogen identification between hospitals and commercially available mNGS tools (López-Labrador et al., 2021). Other logistical challenges include patient privacy, bioinformatic data storage, and lack of standardization (Diao et al., 2022). Continuous effort is needed to improve academically and commercially available tools and make them more accessible to the public (Chen et al., 2022). For example, efforts for improving crowdsourcing of bioinformatics pipelines, software, and taxonomic metagenome profilers could be made. Furthermore, stringent quality controls in the laboratory such as unidirectional workflow, strict decontamination methods during nucleic acid processing, and use of negative controls can help in reducing the detection of exogenous DNA contamination derived from reagents and laboratory environment (Chen et al., 2022).

Currently, the role of mNGS is mostly derived from case reports and small cohort studies (Zhang et al., 2020). Large-scale clinical and cross-institutional studies are required to validate the clinical efficacy of mNGS and to provide a better understanding of how metagenomic approaches could help us to improve patient outcomes over the current standard of care (Zhang et al., 2023). While mNGS may be more expensive than routine diagnostics, the incremental cost is minimum when compared to the cost of invasive diagnostic procedures, a series of diagnostic tests, and cost of intensive care units in hospital; thus, it may help in reducing the overall health care resources (Miller et al., 2020).

## 4. Conclusions and future perspectives

In conclusion, the diversity in clinical metagenomics methods, although allowing flexibility, translates to variability in application and performance relative to conventional diagnostics. Given sensitivity and specificity of mNGS is influenced by a number of factors including the sample type, quantity of host DNA, the sequencing platform used, number of reads generated, selected reference database, and data analysis tools, as well as issues surrounding the current cost and expertise limitations, it is unlikely clinical metagenomics will be utilized as a first-line approach unless these issues are resolved.

## References

- Aggarwal, D., Kanitkar, T., Narouz, M., Azadian, B. S., Moore, L. S. P., and Mughal, N. (2020). Clinical utility and cost-effectiveness of bacterial 16S rRNA and targeted PCR based diagnostic testing in a UK microbiology laboratory network. *Sci. Rep.* 10:7965. doi: 10.1038/s41598-020-64739-1
- Arroyo Mühr, L. S., Dillner, J., Ure, A. E., Sundström, K., and Hultin, E. (2021). Comparison of DNA and RNA sequencing of total nucleic acids from human cervix for metagenomics. *Sci. Rep.* 11:18852. doi: 10.1038/s41598-021-98452-4

Thus, an important question in clinical metagenomics remains: Should mNGS only be used as a last resort when gold-standard culture-based or current molecular diagnostic methods fail? Or are there benefits to using it at earlier stages in certain populations?

At this juncture, clinical settings where copious resources are already expended or where patients are at risk for mortality from rare or difficult to treat pathogens, such as in critically ill or immunosuppressed patients (transplant, cancer etc.), could exemplify practical target populations for implementing clinical metagenomics as a standard of care. Despite unresolved application questions, mNGS has undeniable benefits over traditional testing and potentially provides a more complete picture of the state of any clinical infection. Thus, the combination of mNGS and conventional diagnostic methods could be a superior diagnostic strategy to improve overall public health and healthcare associated costs. Moreover, mNGS may be vital for time-sensitive diagnostics in life-threatening infections. With advancements in sequencing technology, clinical metagenomic sequencing can not only identify the pathogen but also predict antimicrobial resistance and virulence, enabling prompt and effective treatment. However, to fully utilize mNGS as a clinical diagnostic tool, it is essential to standardize the methods, bioinformatics, and databases used in practice.

## Author contributions

JG-P conceived and supervised the review topic. All authors contributed to the article and approved the submitted version.

## Funding

This work was partially supported by NIH 1K01AI143881 to JG-P.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Azar, M. M., Schlaberg, R., Malinis, M. F., Bermejo, S., Schwarz, T., Xie, H., et al. (2021). Added diagnostic utility of clinical metagenomics for the diagnosis of pneumonia in immunocompromised adults. *Chest* 159, 1356–1371. doi: 10.1016/j.chest.2020.11.008

- Barrett, S. L. R., Holmes, E. A., Long, D. R., Shean, R. C., Bautista, G. E., Ravishanker, S., et al. (2020). Cell free DNA from respiratory pathogens is detectable in the blood plasma of cystic fibrosis patients. *Sci. Rep.* 10, 1–6. doi: 10.1038/s41598-020-63970-0

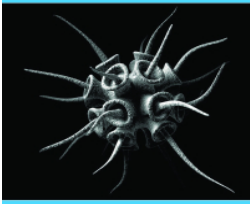


- Blauwkamp, T. A., Thair, S., Rosen, M. J., Blair, L., Lindner, M. S., Vilfan, I. D., et al. (2019). Analytical and clinical validation of a microbial cell-free DNA sequencing test for infectious disease. *Nat. Microbiol.* 4, 663–674. doi: 10.1038/s41564-018-0349-6
- Bohl, J. A., Lay, S., Chea, S., Ahyoung, V., Parker, D. M., Gallagher, S., et al. (2022). Discovering disease-causing pathogens in resource-scarce Southeast Asia using a global metagenomic pathogen monitoring system. *Proc. Natl. Acad. Sci.* 119:e2115285119. doi: 10.1073/pnas.2115285119
- Briese, T., Kapoor, A., Mishra, N., Jain, K., Kumar, A., Jabado, O. J., et al. (2015). Virome capture sequencing enables sensitive viral diagnosis and comprehensive virome analysis. *MBio* 6, e01491–e01415. doi: 10.1128/mBio.01491-15
- Buchan, B. W., Armand-Lefevre, L., and Anderson, N. (2022). Molecular diagnosis of pneumonia (including multiplex panels). *Clin. Chem.* 68, 59–68. doi: 10.1093/clinchem/hvab143
- Calderaro, A., Martinelli, M., Motta, F., Larini, S., Arcangeletti, M. C., Medici, M. C., et al. (2014). Comparison of peptide nucleic acid fluorescence in situ hybridization assays with culture-based matrix-assisted laser desorption/ionization-time of flight mass spectrometry for the identification of bacteria and yeasts from blood cultures and cerebrospinal fluid cultures. *Clin. Microbiol. Infect.* 20, O468–O475. doi: 10.1111/1469-0691.12490
- Carroll, K. C., Hobden, J. A., Miller, S., Morse, S. A., Mietzner, T. A., Detrick, B., et al. (eds). (2019). “Principles of diagnostic medical microbiology” in *Jawetz, Melnick, & Adelberg’s medical microbiology, 27e* (New York, NY: McGraw-Hill Education)
- Castañeda-Mogollón, D., Kamaliddin, C., Oberding, L., Liu, Y., Mohon, A. N., Faridi, R. M., et al. (2021). A metagenomics workflow for SARS-CoV-2 identification, co-pathogen detection, and overall diversity. *J. Clin. Virol.* 145:105025. doi: 10.1016/j.jcv.2021.105025
- Chen, Y., Fan, L. C., Chai, Y. H., and Xu, J. F. (2022). Advantages and challenges of metagenomic sequencing for the diagnosis of pulmonary infectious diseases. *Clin. Respir. J.* 16, 646–656. doi: 10.1111/crj.13538
- Chen, P., Sun, W., and He, Y. (2020). Comparison of metagenomic next-generation sequencing technology, culture and GeneXpert MTB/RIF assay in the diagnosis of tuberculosis. *J. Thorac. Dis.* 12, 4014–4024. doi: 10.21037/jtd-20-1232
- Chen, Y., Yan, X., Li, N., Zhang, Q., Lv, Y., and Ruan, T. (2021). Metagenomic next-generation sequencing of radial ultrasound bronchoscopy-guided “cocktail” specimens as an efficient method for the diagnosis of focal pulmonary infections: a randomised study. *Annals Palliat Med* 10, 2080–2088. doi: 10.21037/apm-20-2578
- Chen, H., Yin, Y., Gao, H., Guo, Y., Dong, Z., Wang, X., et al. (2020). Clinical utility of in-house metagenomic next-generation sequencing for the diagnosis of lower respiratory tract infections and analysis of the host immune response. *Clin. Infect. Dis.* 71, S416–S426. doi: 10.1093/cid/ciaa1516
- Chen, J., Zhang, R., Liu, L., Qi, T., Wang, Z., Song, W., et al. (2021b). Clinical usefulness of metagenomic next-generation sequencing for the diagnosis of central nervous system infection in people living with HIV. *Int. J. Infect. Dis.* 107, 139–144. doi: 10.1016/j.ijid.2021.04.057
- Chen, J., Zhao, Y., Shang, Y., Lin, Z., Xu, G., Bai, B., et al. (2021a). The clinical significance of simultaneous detection of pathogens from bronchoalveolar lavage fluid and blood samples by metagenomic next-generation sequencing in patients with severe pneumonia. *J. Med. Microbiol.* 70:001259. doi: 10.1099/jmm.0.001259
- Chen, S., Zhu, Q., Xiao, Y., Wu, C., Jiang, Z., Liu, L., et al. (2021). Clinical and etiological analysis of co-infections and secondary infections in COVID-19 patients: an observational study. *Clin. Respir. J.* 15, 815–825. doi: 10.1111/crj.13369
- Chiu, C. Y., and Miller, S. A. (2019). Clinical metagenomics. *Nat. Rev. Genet.* 20, 341–355. doi: 10.1038/s41576-019-0113-7
- Chiu, C. Y., and Miller, S. A. (n.d.). University of California San Francisco: The next generation of diagnostic tests. Available at: <https://nextgendiagnosics.ucsf.edu/our-diagnostic-lab/>.
- d’Humières, C., Salmons, M., Dellière, S., Leo, S., Rodriguez, C., Angebault, C., et al. (2021). The potential role of clinical metagenomics in infectious diseases: therapeutic perspectives. *Drugs* 81, 1453–1466. doi: 10.1007/s40265-021-01572-4
- Deng, W., Xu, H., Wu, Y., and Li, J. (2022). Diagnostic value of bronchoalveolar lavage fluid metagenomic next-generation sequencing in pediatric pneumonia. *Front. Cell. Infect. Microbiol.* 12:1536. doi: 10.3389/fcimb.2022.950531
- Diao, Z., Han, D., Zhang, R., and Li, J. (2022). Metagenomics next-generation sequencing tests take the stage in the diagnosis of lower respiratory tract infections. *J. Adv. Res.* 38, 201–212. doi: 10.1016/j.jare.2021.09.012
- Dingle, T. C., and Butler-Wu, S. M. (2013). MALDI-TOF mass spectrometry for microorganism identification. *Clin. Lab. Med.* 33, 589–609. doi: 10.1016/j.cl.2013.03.001
- Duan, H., Li, X., Mei, A., Li, P., Liu, Y., Li, X., et al. (2021). The diagnostic value of metagenomic next-generation sequencing in infectious diseases. *BMC Infect. Dis.* 21, 1–13. doi: 10.1186/s12879-020-05746-5
- Eichenberger, E. M., de Vries, C. R., Ruffin, F., Sharma-Kuinkel, B., Park, L., Hong, D., et al. (2022). Microbial cell-free DNA identifies etiology of bloodstream infections, persists longer than conventional blood cultures, and its duration of detection is associated with metastatic infection in patients with staphylococcus aureus and gram-negative bacteremia. *Clin. Infect. Dis.* 74, 2020–2027. doi: 10.1093/cid/ciab742
- Fournier, P.-E., Dubourg, G., and Raoult, D. (2014). Clinical detection and characterization of bacterial pathogens in the genomics era. *Genome Med.* 6, 1–15. doi: 10.1186/s13073-014-0114-2
- Fry, S. E. (n.d.). Fry laboratories: Next-generation sequencing for infectious diseases Available at: <https://frylabs.com/dna-sequencing/>.
- Fu, Z.-F., Zhang, H. C., Zhang, Y., Cui, P., Zhou, Y., Wang, H. Y., et al. (2022). Evaluations of clinical utilization of metagenomic next-generation sequencing in adults with fever of unknown origin. *Front. Cell. Infect. Microbiol.* 11:1250. doi: 10.3389/fcimb.2021.745156
- Gohl, D. M., Vangay, P., Garbe, J., MacLean, A., Hauge, A., Becker, A., et al. (2016). Systematic improvement of amplicon marker gene methods for increased accuracy in microbiome studies. *Nat. Biotechnol.* 34, 942–949. doi: 10.1038/nbt.3601
- Govender, K. N., Street, T. L., Sanderson, N. D., and Eyre, D. W. (2021). Metagenomic sequencing as a pathogen-agnostic clinical diagnostic tool for infectious diseases: a systematic review and meta-analysis of diagnostic test accuracy studies. *J. Clin. Microbiol.* 59, e02916–e02920. doi: 10.1128/JCM.02916-20
- Greenberg, R. G., and Herrera, T. I. (2019). “Chapter 8—when to perform lumbar puncture in infants at risk for meningitis in the neonatal intensive care unit” in *Infectious disease and pharmacology*. eds. W. E. Benitz and P. B. Smith (Philadelphia: Elsevier), 87–102.
- Greninger, A. L. (2018). The challenge of diagnostic metagenomics. *Expert Rev. Mol. Diagn.* 18, 605–615. doi: 10.1080/14737159.2018.1487292
- Gu, W., Crawford, E. D., O’Donovan, B. D., Wilson, M. R., Chow, E. D., Retallack, H., et al. (2016). Depletion of abundant sequences by hybridization (DASH): using Cas9 to remove unwanted high-abundance species in sequencing libraries and molecular counting applications. *Genome Biol.* 17, 1–13. doi: 10.1186/s13059-016-0904-5
- Gu, W., Deng, X., Lee, M., Sucu, Y. D., Arevalo, S., Stryke, D., et al. (2021). Rapid pathogen detection by metagenomic next-generation sequencing of infected body fluids. *Nat. Med.* 27, 115–124. doi: 10.1038/s41591-020-1105-z
- Gu, W., Miller, S., and Chiu, C. Y. (2019). Clinical metagenomic next-generation sequencing for pathogen detection. *Annu. Rev. Pathol.* 14, 319–338. doi: 10.1146/annurev-pathmechdis-012418-012751
- Guo, W., Cui, X., Wang, Q., Wei, Y., Guo, Y., Zhang, T., et al. (2022). Clinical evaluation of metagenomic next-generation sequencing for detecting pathogens in bronchoalveolar lavage fluid collected from children with community-acquired pneumonia. *Front. Med.* 9:9. doi: 10.3389/fmed.2022.952636
- Hasan, M. R., Rawat, A., Tang, P., Jithesh, P. V., Thomas, E., Tan, R., et al. (2016). Depletion of human DNA in spiked clinical specimens for improvement of sensitivity of pathogen detection by next-generation sequencing. *J. Clin. Microbiol.* 54, 919–927. doi: 10.1128/JCM.03050-15
- Hasan, M. R., Sundararaju, S., Tang, P., Tsui, K. M., Lopez, A. P., Janahi, M., et al. (2020). A metagenomics-based diagnostic approach for central nervous system infections in hospital acute care setting. *Sci. Rep.* 10:11194. doi: 10.1038/s41598-020-68159-z
- Haston, J. C., Rostad, C. A., Jerriss, R. C., Milla, S. S., McCracken, C., Pratt, C., et al. (2020). Prospective cohort study of next-generation sequencing as a diagnostic modality for unexplained encephalitis in children. *J. Pediatr Infect Diseases Soc* 9, 326–333. doi: 10.1093/jpids/piz032
- Heming, M., Börsch, A. L., Wiendl, H., and Meyer zu Hörste, G. (2022). High-dimensional investigation of the cerebrospinal fluid to explore and monitor CNS immune responses. *Genome Med.* 14, 1–12. doi: 10.1186/s13073-022-01097-9
- Hiergeist, A., Gläser, J., Reischl, U., and Gessner, A. (2015). Analyses of intestinal microbiota: culture versus sequencing. *ILAR J.* 56, 228–240. doi: 10.1093/ilar/ilv017
- Hogan, C. A., Yang, S., Garner, O. B., Green, D. A., Gomez, C. A., Dien Bard, J., et al. (2021). Clinical impact of metagenomic next-generation sequencing of plasma cell-free DNA for the diagnosis of infectious diseases: a multicenter retrospective cohort study. *Clin. Infect. Dis.* 72, 239–245. doi: 10.1093/cid/ciaa035
- Huanyu Wang, S. J. (2021). *Next-generation sequencing for infectious diseases diagnostics—Is it worth the hype?* Washington, DC: American Association for clinical chemistry.
- Jia, X., Hu, L., Wu, M., Ling, Y., Wang, W., Lu, H., et al. (2021). A streamlined clinical metagenomic sequencing protocol for rapid pathogen identification. *Sci. Rep.* 11, 1–10. doi: 10.1038/s41598-021-83812-x
- Jiang, W. (n.d.). Beijing genomic institute (BGI genomics): Metagenomic sequencing. Available at: <https://www.bgi.com/global/science-detail/metagenomic-sequencing>.
- Jin, X., Li, J., Shao, M., Lv, X., Ji, N., Zhu, Y., et al. (2022). Improving suspected pulmonary infection diagnosis by Bronchoalveolar lavage fluid metagenomic next-generation sequencing: a multicenter retrospective study. *Microbiol Spectr* 10, e02473–e02421. doi: 10.1128/spectrum.02473-21
- Jing, C., Chen, H., Liang, Y., Zhong, Y., Wang, Q., Li, L., et al. (2021). Clinical evaluation of an improved metagenomic next-generation sequencing test for the diagnosis of bloodstream infections. *Clin. Chem.* 67, 1133–1143. doi: 10.1093/clinchem/hvab061
- Jing, Q., Leung, C. H. C., and Wu, A. R. (2022). Cell-free DNA as biomarker for sepsis by integration of microbial and host information. *Clin. Chem.* 68, 1184–1195. doi: 10.1093/clinchem/hvac097



- John, G., Sahajpal, N. S., Mondal, A. K., Ananth, S., Williams, C., Chaubey, A., et al. (2021). Next-generation sequencing (NGS) in COVID-19: a tool for SARS-CoV-2 diagnosis, monitoring new strains and phylodynamic modeling in molecular epidemiology. *Curr. Issues Mol. Biol.* 43, 845–867. doi: 10.3390/cimb43020061
- Kalantar, K. L., Neyton, L., Abdelghany, M., Mick, E., Jauregui, A., Caldera, S., et al. (2022). Integrated host-microbe plasma metagenomics for sepsis diagnosis in a prospective cohort of critically ill adults. *Nat. Microbiol.* 7, 1–12. doi: 10.1038/s41564-022-01237-2
- Li, G., Huang, J., Li, Y., and Feng, J. (2020). The value of combined radial endobronchial ultrasound-guided transbronchial lung biopsy and metagenomic next-generation sequencing for peripheral pulmonary infectious lesions. *Can. Respir. J.* 2020, 1–9. doi: 10.1155/2020/2367505
- Li, N., Ma, X., Zhou, J., Deng, J., Gu, C., Fei, C., et al. (2022). Clinical application of metagenomic next-generation sequencing technology in the diagnosis and treatment of pulmonary infection pathogens: a prospective single-center study of 138 patients. *J. Clin. Lab. Anal.* 36:e24498. doi: 10.1002/jcla.24498
- Lippert, J. F., Buscemi, J., Saiyed, N., Silva, A., and Benjamins, M. R. (2022). Influenza and pneumonia mortality across the 30 biggest US cities: assessment of overall trends and racial inequities. *J. Racial Ethn. Health Disparities* 9, 1152–1160. doi: 10.1007/s40615-021-01056-x
- Liu, W.-D., Yen, T. Y., Liu, P. Y., Wu, U. I., Bhan, P., Li, Y. C., et al. (2021). Clinical application of metagenomic next-generation sequencing in patients with hematologic malignancies suffering from sepsis. *Microorganisms* 9:2309. doi: 10.3390/microorganisms9112309
- López-Labrador, F. X., Brown, J. R., Fischer, N., Harvala, H., van Boheemen, S., Cinek, O., et al. (2021). Recommendations for the introduction of metagenomic high-throughput sequencing in clinical virology, part I: wet lab procedure. *J. Clin. Virol.* 134:104691. doi: 10.1016/j.jcv.2020.104691
- Miller, S., and Chiu, C. (2022). The role of metagenomics and next-generation sequencing in infectious disease diagnosis. *Clin. Chem.* 68, 115–124. doi: 10.1093/clinchem/hvab173
- Miller, S., Chiu, C., Rodino, K. G., and Miller, M. B. (2020). Point-counterpoint: should we be performing metagenomic next-generation sequencing for infectious disease diagnosis in the clinical laboratory? *J. Clin. Microbiol.* 58, e01739–e01719. doi: 10.1128/JCM.01739-19
- Miller, S., Naccache, S. N., Samayoa, E., Messacar, K., Arevalo, S., Federman, S., et al. (2019). Laboratory validation of a clinical metagenomic sequencing assay for pathogen detection in cerebrospinal fluid. *Genome Res.* 29, 831–842. doi: 10.1101/gr.238170.118
- Mokhtari, M., Alizadeh, A., Monabati, A., and Safaei, A. (2022). Comparison of flowcytometry and conventional cytology for diagnosis of CNS involvement in hematologic malignancies. *Pediatr Hematol Oncol J* 7, 146–154. doi: 10.1016/j.phoj.2022.08.005
- Morsli, M., Boudet, A., Kerharo, Q., Stephan, R., Salipante, F., Dunyach-Remy, C., et al. (2022b). Real-time metagenomics-based diagnosis of community-acquired meningitis: a prospective series, southern France. *EBioMedicine* 84:104247. doi: 10.1016/j.ebiom.2022.104247
- Morsli, M., Kerharo, Q., Delerce, J., Roche, P. H., Troude, L., and Drancourt, M. (2021a). *Haemophilus influenzae* meningitis direct diagnosis by metagenomic next-generation sequencing: a case report. *Pathogens* 10:461. doi: 10.3390/pathogens10040461
- Morsli, M., Lavigne, J. P., and Drancourt, M. (2022a). Direct metagenomic diagnosis of community-acquired meningitis: state of the art. *Front. Microbiol.* 13:13. doi: 10.3389/fmicb.2022.926240
- Morsli, M., Vincent, J. J., Milliere, L., Colson, P., and Drancourt, M. (2021b). Direct next-generation sequencing diagnosis of echovirus 9 meningitis, France. *Eur. J. Clin. Microbiol. Infect. Dis.* 40, 2037–2039. doi: 10.1007/s10096-021-04205-6
- Mu, S., Hu, L., Zhang, Y., Liu, Y., Cui, X., Zou, X., et al. (2021). Prospective evaluation of a rapid clinical metagenomics test for bacterial pneumonia. *Front. Cell. Infect. Microbiol.* 11:953. doi: 10.3389/fcimb.2021.684965
- Olausson, J., Brunet, S., Vracar, D., Tian, Y., Abrahamsson, S., Meghadri, S. H., et al. (2022). Optimization of cerebrospinal fluid microbial DNA metagenomic sequencing diagnostics. *Sci. Rep.* 12:3378. doi: 10.1038/s41598-022-07260-x
- Pollock, N., MacIntyre, A., Blauwkamp, T. A., Blair, L., Ho, C., Calderon, R., et al. (2021). Detection of *Mycobacterium tuberculosis* cell-free DNA to diagnose TB in pediatric and adult patients. *Int. J. Tuberc. Lung Dis.* 25, 403–405. doi: 10.5588/ijtld.21.0055
- Prachayangprecha, S., Schapendonk, C. M. E., Koopmans, M. P., Osterhaus, A. D. M. E., Schürch, A. C., Pas, S. D., et al. (2014). Exploring the potential of next-generation sequencing in detection of respiratory viruses. *J. Clin. Microbiol.* 52, 3722–3730. doi: 10.1128/JCM.01641-14
- Quan, J., Langelier, C., Kuchta, A., Batson, J., Teyssier, N., Lyden, A., et al. (2019). FLASH: a next-generation CRISPR diagnostic for multiplexed detection of antimicrobial resistance sequences. *Nucleic Acids Res.* 47:e83. doi: 10.1093/nar/gkz418
- Quer, J., Colomer-Castell, S., Campos, C., Andrés, C., Piñana, M., Cortese, M. F., et al. (2022). Next-generation sequencing for confronting virus pandemics. *Viruses* 14:600. doi: 10.3390/v14030600
- Ramachandran, P. S., and Wilson, M. R. (2020). Metagenomics for neurological infections—expanding our imagination. *Nat. Rev. Neurol.* 16, 547–556. doi: 10.1038/s41582-020-0374-y
- Ramesh, A., Nakiely, S., Hsu, J., Kyohere, M., Byaruhanga, O., de Bourcy, C., et al. (2019). Metagenomic next-generation sequencing of samples from pediatric febrile illness in Tororo, Uganda. *PLoS One* 14:e0218318. doi: 10.1371/journal.pone.0218318
- Rampini, S. K., Bloemberg, G. V., Keller, P. M., Büchler, A. C., Dollenmaier, G., Speck, R. F., et al. (2011). Broad-range 16S rRNA gene polymerase chain reaction for diagnosis of culture-negative bacterial infections. *Clin. Infect. Dis.* 53, 1245–1251. doi: 10.1093/cid/cir692
- Roux-Dalvai, F., Gotti, C., Leclercq, M., Hélie, M. C., Boissinot, M., Arrey, T. N., et al. (2019). Fast and accurate bacterial species identification in urine specimens using LC-MS/MS mass spectrometry and machine learning\*[S]. *Mol. Cell. Proteomics* 18, 2492–2505. doi: 10.1074/mcp.TIR119.001559
- Rudd, K. E., Johnson, S. C., Agesa, K. M., Shackelford, K. A., Tsoi, D., Kievlan, D. R., et al. (2020). Global, regional, and national sepsis incidence and mortality, 1990–2017: analysis for the global burden of disease study. *Lancet* 395, 200–211. doi: 10.1016/S0140-6736(19)32989-7
- Salipante, S. J., Sengupta, D. J., Rosenthal, C., Costa, G., Spangler, J., Sims, E. H., et al. (2013). Rapid 16S rRNA next-generation sequencing of polymicrobial clinical samples for diagnosis of complex bacterial infections. *PLoS One* 8:e65226. doi: 10.1371/journal.pone.0065226
- Simner, P. J., Miller, S., and Carroll, K. C. (2018). Understanding the promises and hurdles of metagenomic next-generation sequencing as a diagnostic tool for infectious diseases. *Clin. Infect. Dis.* 66, 778–788. doi: 10.1093/cid/cix881
- Thair, S., Seng, H., Hollemon, D., Hong, D., Blauwkamp, T., Kertesz, M., et al. (2017). “The SEP-SEQ trial: clinical validation of the Karius plasma next-generation sequencing test for pathogen detection in sepsis” in *Open forum infectious diseases* (Oxford University Press), 4.
- Tkadlec, J., Peckova, M., Sramkova, L., Rohn, V., Jahoda, D., Raszka, D., et al. (2019). The use of broad-range bacterial PCR in the diagnosis of infectious diseases: a prospective cohort study. *Clin. Microbiol. Infect.* 25, 747–752. doi: 10.1016/j.cmi.2018.10.001
- Tremblay, J., Singh, K., Fern, A., Kirton, E. S., He, S., Woyke, T., et al. (2015). Primer and platform effects on 16S rRNA tag sequencing. *Front. Microbiol.* 6:771. doi: 10.3389/fmicb.2015.00771
- Venter, J. M., Mahlangu, P. M., Müller, E. E., Lewis, D. A., Rebe, K., Struthers, H., et al. (2019). Comparison of an in-house real-time duplex PCR assay with commercial HOLOGIC® APTIMA assays for the detection of *Neisseria gonorrhoeae* and *Chlamydia trachomatis* in urine and extra-genital specimens. *BMC Infect. Dis.* 19, 6–7. doi: 10.1186/s12879-018-3629-0
- Vetter, P., Schibler, M., Herrmann, J. L., and Boutolleau, D. (2020). Diagnostic challenges of central nervous system infection: extensive multiplex panels versus stepwise guided approach. *Clin. Microbiol. Infect.* 26, 706–712. doi: 10.1016/j.cmi.2019.12.013
- Viral Hemorrhagic Fever Consortium/Metsky, H. C., Siddle, K. J., Gladden-Young, A., Qu, J., Yang, D. K., et al. (2019). Capturing sequence diversity in metagenomes with comprehensive and scalable probe design. *Nat. Biotechnol.* 37, 160–168. doi: 10.1038/s41587-018-0006-x
- Wagner, K., Springer, B., Pires, V. P., and Keller, P. M. (2018). Molecular detection of fungal pathogens in clinical specimens by 18S rDNA high-throughput screening in comparison to ITS PCR and culture. *Sci. Rep.* 8, 1–7. doi: 10.1038/s41598-018-25129-w
- Wang, L., Li, S., Qin, J., Tang, T., Hong, J., Tung, T. H., et al. (2023). Clinical diagnosis application of metagenomic next-generation sequencing of plasma in suspected Sepsis. *Infect Drug Resist* 16, 891–901. doi: 10.2147/IDR.S395700
- Wang, G., Xu, N., Yang, L., Zheng, F., Sai, L., Zhou, J., et al. (2019). Community acquired *Stenotrophomonas maltophilia* discitis: diagnosis aided by shotgun metagenomic sequencing. *Int. J. Infect. Dis.* 81, 1–3. doi: 10.1016/j.ijid.2019.01.032
- Wang, C., Yan, D., Huang, J., Yang, N., Shi, J., Pan, S., et al. (2022). The clinical application of metagenomic next-generation sequencing in infectious diseases at a tertiary hospital in China. *Front. Cell. Infect. Microbiol.* 12:1853. doi: 10.3389/fcimb.2022.957073
- Weaver, A. J., Brandenburg, K. S., Sanjar, F., Wells, A. R., Peacock, T. J., and Leung, K. P. (2019). Clinical utility of PNA-FISH for burn wound diagnostics: a noninvasive, culture-independent technique for rapid identification of bacterial organisms in burn wounds. *J. Burn Care Res.* 40, 464–470. doi: 10.1093/jbcr/irz047
- Wilson, M. R., Naccache, S. N., Samayoa, E., Biagtan, M., Bashir, H., Yu, G., et al. (2014). Actionable diagnosis of neuroleptospirosis by next-generation sequencing. *N. Engl. J. Med.* 370, 2408–2417. doi: 10.1056/NEJMoa1401268
- Wilson, M. R., Sample, H. A., Zorn, K. C., Arevalo, S., Yu, G., Neuhaus, J., et al. (2019). Clinical metagenomic sequencing for diagnosis of meningitis and encephalitis. *N. Engl. J. Med.* 380, 2327–2340. doi: 10.1056/NEJMoa1803396
- Wu, F., Zhao, S., Yu, B., Chen, Y. M., Wang, W., Song, Z. G., et al. (2020). A new coronavirus associated with human respiratory disease in China. *Nature* 579, 265–269. doi: 10.1038/s41586-020-2008-3

- Xu, L., Zhao, Z., Mai, H., Tan, X., du, Y., and Fang, C. (2022). Clinical and chest computed tomography features associated with severe *Chlamydia psittaci* pneumonia diagnosed by metagenomic next-generation sequencing: a multicenter, retrospective, observational study. *Medicine* 101:e32117. doi: 10.1097/MD.00000000000032117
- Yek, C., Pacheco, A. R., Vanaerschot, M., Bohl, J. A., Fahsbender, E., Aranda-Díaz, A., et al. (2022). Metagenomic pathogen sequencing in resource-scarce settings: lessons learned and the road ahead. *Front. Epidemiol.* 2:2. doi: 10.3389/fepid.2022.926695
- Zhang, X., Chen, H., Han, D., and Wu, W. (2023). Clinical usefulness of metagenomic next-generation sequencing for *Rickettsia* and *Coxiella burnetii* diagnosis. *Eur. J. Clin. Microbiol. Infect. Dis.* 42, 681–689. doi: 10.1007/s10096-023-04586-w
- Zhang, P., Chen, Y., Li, S., Li, C., Zhang, S., Zheng, W., et al. (2020). Metagenomic next-generation sequencing for the clinical diagnosis and prognosis of acute respiratory distress syndrome caused by severe pneumonia: a retrospective study. *PeerJ* 8:e9623. doi: 10.7717/peerj.9623
- Zhang, D., Yang, X., Wang, J., Xu, J., and Wang, M. (2022). Application of metagenomic next-generation sequencing for bronchoalveolar lavage diagnostics in patients with lower respiratory tract infections. *J. Int. Med. Res.* 50:030006052210897. doi: 10.1177/03000605221089795
- Zheng, Y., Qiu, X., Wang, T., and Zhang, J. (2021). The diagnostic value of metagenomic next-generation sequencing in lower respiratory tract infection. *Front. Cell. Infect. Microbiol.* 11:694756. doi: 10.3389/fcimb.2021.694756
- Zhou, H., Larkin, P. M. K., Zhao, D., Ma, Q., Yao, Y., Wu, X., et al. (2021). Clinical impact of metagenomic next-generation sequencing of bronchoalveolar lavage in the diagnosis and management of pneumonia: a multicenter prospective observational study. *J. Mol. Diagn.* 23, 1259–1268. doi: 10.1016/j.jmoldx.2021.06.007
- Zhu, Y., Xu, M., Ding, C., Peng, Z., Wang, W., Sun, B., et al. (2022). Metagenomic next-generation sequencing vs. traditional microbiological tests for diagnosing varicella-zoster virus central nervous system infection. *Front. Public Health* 9:2165. doi: 10.3389/fpubh.2021.738412



## Clinical microbiology in detection and identification of emerging microbial pathogens: past, present and future

Hui Wang, Wenhong Zhang & Yi-Wei Tang

To cite this article: Hui Wang, Wenhong Zhang & Yi-Wei Tang (2022) Clinical microbiology in detection and identification of emerging microbial pathogens: past, present and future, Emerging Microbes & Infections, 11:1, 2579-2589, DOI: [10.1080/22221751.2022.2125345](https://doi.org/10.1080/22221751.2022.2125345)

To link to this article: <https://doi.org/10.1080/22221751.2022.2125345>



© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 04 Nov 2022.



Submit your article to this journal [↗](#)



Article views: 5906



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 5 View citing articles [↗](#)

## Clinical microbiology in detection and identification of emerging microbial pathogens: past, present and future

Hui Wang<sup>a</sup>, Wenhong Zhang<sup>b</sup> and Yi-Wei Tang <sup>c</sup>

<sup>a</sup>Department of Clinical Laboratories, Peking University People's Hospital, Beijing, People's Republic of China; <sup>b</sup>Department of Infectious Diseases, National Medical Center for Infectious Diseases, Huashan Hospital, Fudan University, Shanghai, People's Republic of China; <sup>c</sup>Medical Affairs, Danaher Diagnostic Platform China/Cepheid, Shanghai, People's Republic of China

### ABSTRACT

Clinical microbiology has possessed a marvellous past, an important present and a bright future. Western medicine modernization started with the discovery of bacterial pathogens, and from then, clinical bacteriology became a cornerstone of diagnostics. Today, clinical microbiology uses standard techniques including Gram stain morphology, *in vitro* culture, antigen and antibody assays, and molecular biology both to establish a diagnosis and monitor the progression of microbial infections. Clinical microbiology has played a critical role in pathogen detection and characterization for emerging infectious diseases as evidenced by the ongoing COVID-19 pandemic. Revolutionary changes are on the way in clinical microbiology with the application of “-omic” techniques, including transcriptomics and metabolomics, and optimization of clinical practice configurations to improve outcomes of patients with infectious diseases.

**ARTICLE HISTORY** Received 17 May 2022; Revised 6 September 2022; Accepted 12 September 2022

**KEYWORDS** Clinical microbiology; nucleic acid amplification; genomics; transcriptomics; proteomics; metabolomics; point of care; COVID-19



The unprecedented outbreak of the coronavirus disease 2019 (COVID-19) pandemic has highlighted the necessity for readily available, accurate and fast diagnostic testing methods to detect and characterize emerging pathogens. The basic work of the clinical microbiology laboratory is to provide evidence for the diagnosis, treatment, and control of infectious diseases by detecting the presence of specific pathogenic microorganisms in clinical specimens. When a pathogen is detected, it is then subjected to a series of further analyses, including identification, typing, quantification, and antimicrobial susceptibility testing [1,2].

### A long and uneven past

The discipline of clinical microbiology has been evolving for more than two hundred years since Dutchman, Anton van Leeuwenhoek invented the microscope (Table 1). The development of the light microscope was the foundational technological advancement in modern medicine for the direct visualization of microorganisms. By the 1830s, Lister introduced the “achromatic” lens to eliminate the blurring and colour distortion known as “chromatic aberration,” which had previously limited resolution

at higher magnification (i.e. bacterial level) [3]. These technical advancements in microscopy converged with seminal concepts in bacteriology from the work of Pasteur and Koch. Together, they allowed the microscope to serve as a powerful instrument for physicians and microbiologists to directly visualize pathogens in human specimens, especially with the application of (still universal) techniques such as the Gram and acid-fast staining procedures, also developed in the late nineteenth century [2,4].

Equally important during this period were new-found abilities to culture microorganisms from human sources. While early bacterial cultures were accomplished with slices of raw potato, the 1887 invention of the petri dish facilitated the direct observation of colonies with gelatine or agar, allowing for the morphological description of species in pure culture [5]. It likewise allowed for their biochemical characterization, creating the phenotypic profiles that have served as the basis for taxonomic identification in clinical laboratories until the twenty-first century. Culture is likewise a pre-requisite for phenotypic antimicrobial susceptibility testing (AST), which remains the gold standard for predicting the response









**CONTACT** Yi-Wei Tang  [yi-wei.tang@cepheid.com](mailto:yi-wei.tang@cepheid.com)  Medical Affairs, Danaher Diagnostic Platform China/Cepheid, 518 Fuquan North Road, Shanghai 200325, People's Republic of China

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Table 1.** Historical figures in clinical microbiology.

	Scientist	Year of birth and death	Major contributions in clinical microbiology
	Anton van Leeuwenhoek	1632–1723	Invented microscope to observe microorganisms
	Louis Pasteur	1822–1895	Found microorganisms in infected patients
	Robert Koch	1843–1910	Established three criteria for determining a causative relationship between a microbe and a disease
	Hans Christian Gram	1853–1938	Invented the Gram stain
	Julius Richard Petri	1852–1921	Invented Petri dishes to start in vitro culture
	John Franklin Enders	1897–1985	Applied cell culture techniques for isolation and growth of viruses
	Feifan Tang	1897–1958	Discovered and cultured the pathogen causing trachoma
	Kary Banks Mullis	1944–2019	Invented polymerase chain reaction for in vitro nucleic acid amplification

of a patient's bacterial/fungal infection to treatment [1,4,6].

Shortly after the development of axenic bacterial/fungal culture, methods for propagating human viruses using *ex vivo* cellular substrates, fertilized chicken embryos, and *in vitro* cell lines became available. In the late 1940s, John Franklin Enders first applied cell culture techniques to isolate and grow poliovirus, initiating modern clinical virology [7]. In 1956, Feifan Tang, a Chinese microbiologist and virologist, discovered and isolated *Chlamydia trachomatis*, which clarified the cause of trachoma [8]. For decades, tissue and cell culture-based methods have played critical roles in rapid identification of emerging pathogens and exploration of pathogenesis. This has been demonstrated in three unprecedented outbreaks of emerging human coronavirus (SARS-CoV, MERS-CoV and SARS-CoV-2) infections since the beginning of the twenty-first century [9]. More importantly, cell culture techniques, along with chicken egg embryos, have become the main tools to reproduce large quantities of viruses for vaccine manufacturing for fighting emerging infectious diseases [10].

Alongside these abilities to visualize and cultivate organisms came new techniques that assess the body's humoral response to infection. This principle was first established in 1892 by Sternberg using the vaccinia virus [11]. It is now recognized that antibodies within serum facilitate identification not just

of viruses, but multiple types of pathogens by the clinical laboratory [12,13]. The value of diagnostic serology for infectious diseases is founded on two criteria: (i) antibodies are specific for a particular cellular target and (ii) antibodies are induced by a specific stimulus. Ideally, the definitive evidence for infection by a pathogen would involve recovery of that agent from infected tissue, along with an increase in specific antibodies over time. With some infections, however, the recovery of the causative agent may be difficult, dangerous, impractical, or even impossible [13]. Certainly before the development nucleic acid assays (and in some cases still), the combination of appropriate clinical and serologic findings can serve as sufficient criteria for diagnosis, even if the organism itself is not detected [12]. Seropositivity can likewise function as a marker of functional immunity toward many (but not all) pathogens, and these tests facilitate population monitoring within the fields of epidemiology and public health [12,13]. Moreover, the purification of pathogen-specific antibodies allows for the development of "antigen-based" diagnostics that detect microbe-specific biomolecules other than nucleic acids, typically carbohydrates and proteins.

Kary Banks Mullis, an American biochemist and Nobel Prize winner, was credited with the invention of the polymerase chain reaction (PCR) in the 1980s for nucleic acid amplification, thus applying rapid, sensitive and specific molecular biology techniques

to the practice of clinical microbiology [14,15]. A quick application of reverse transcriptase (RT), an enzyme recovered in 1970 simultaneously by Baltimore and Temin [16,17], resulted in RT-PCR which has been widely used in the detection and characterization of RNA targets. Further modification of PCR into a quantitative format (qPCR) by Higuchi et al. in 1993 has enabled accurate determination of pathogen loads in clinical specimens [18]. Additionally, the automation of the Sanger sequencing method by Leroy Hood and Michael Hunkapillar at Applied Biosystems in 1987 rendered the rapid generation of complete genome sequences of *Haemophilus influenzae* [19] and *Mycoplasma genitalium* [20] a reality [21].



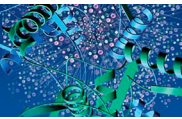

### A flourishing present

In the past decades, clinical microbiology laboratories have undergone important changes with the introduction of molecular biology techniques [22] and laboratory automation [23]. Diagnostic methods in clinical microbiology are currently divided into the following five categories (Table 2) [1]. The first one is the morphological observation under the naked eye or microscope. This is a fundamental method, which is currently used mainly for initial screening and for guiding the next step of testing. Specific morphological findings by microscopy can quickly identify the pathogenetic agent in some cases. For example, the finding of Gram-negative diplococci in urethral exudates from males is a reliable indication of *Neisseria gonorrhoeae* infection. The second is the antigen test for pathogenic microorganisms. It is widely used in clinical practice for outpatient testing because of its

rapidity, simplicity, and specificity [24,25]. Examples for this include laboratory testing for the urine pneumococcal and *Legionella* antigen tests [26]. On the other hand, the disadvantage of antigen tests lies in its often poor sensitivity. Specimens that are antigen-negative in clinical practice usually need to be retested with more sensitive methods, such as culture or/and molecular methods to avoid a missed diagnosis. Examples for this include laboratory testing for *Streptococcus pyogenes* [24] and influenza virus [25].

During the ongoing COVID-19 pandemic, the performance of rapid antigen tests for COVID-19 diagnosis has been widely evaluated and found to be varied in different settings [27]. Rapid antigen tests are cheaper and provide faster results, thus potentially enabling prompt isolation of positive cases and quarantine of close contacts. A recent literature review covering a total of 16 studies reported on the effectiveness of rapid antigen testing for screening of asymptomatic individuals to limit the transmission [9,28]. Eight included studies examining the effectiveness of rapid antigen testing for population-level screening, four for pre-event screening and four for serial testing. Overall, there was no evidence regarding the effectiveness of rapid antigen testing for the screening of asymptomatic individuals to limit the transmission of SARS-CoV-2. This uncertainty is due to the inconsistent results, the relatively low number of studies identified, the predominantly observational and/or uncontrolled nature of the study designs used, and concerns regarding methodological quality. Given this uncertainty, more real-world research evidence in relevant settings, which is of good quality and timely, as well as economic evaluation, is required to inform public policy on the widespread use of rapid antigen tests in asymptomatic individuals [28].

**Table 2.** Microbiomic technology main contents. MALDI-TOF MS, matrix assisted laser desorption ionization time of flight mass spectrometry.

	Technology	Target molecule	Question addressed	Output	Main methods	Selected references
	Genomics	DNA	Infection potential	DNA sequences	DNA sequencing	[69,71,72]
	Transcriptomics	RNA	Infection strategy	Transcription of text	RNA sequencing, quantitative transcription PCR	[77,82,84]
	Proteomics	Protein	Infection process	Protein profiles	MALDI-TOF MS	[83,89,90]
	Metabolomics	Metabolites	Infection outcome	Metabolic text	Gas or liquid phase mass spectrometry	[78–80]

The third category is the culture method which remains the gold standard for culturable pathogenic microorganisms [1]. It mainly includes inanimate media such as agar or broth and animate media such as tissues and cells. The former is mainly used for culturing extracellular pathogens such as bacteria and fungi, while the latter is mainly used for culturing intracellular pathogens such as viruses/chlamydia. Cultivation for bacteriology and mycology has gradually transitioned into a totally automated continuous monitoring mode of blood culture/liquid culture from the manual method. Viral cultures are still an important research and clinical laboratory tool when the potential viral agent is not known. Other application niches include documenting active infection, performing antiviral susceptibility testing and developing a vaccine and therapeutic agents.

Viral culture has been a valuable tool for studying COVID-19 pathogenesis, resulting in developing more effective disease prevention, diagnosis, and control. To combat COVID-19, since SARS-CoV-2 was first isolated from nasopharyngeal and oropharyngeal specimens from a patient with COVID-19 in Vero-CCL81 and Vero E6 cells [29], several *in vitro* and *ex vivo* cell culture systems have progressively been used and described [30]. Matsuyama et al engineered a Vero E6 cell line expressing TMPRSS2 for culturing of SARS-CoV-2 showing more than 100 times greater production of viral RNA copies than Vero E6 cells alone [31]. Cell culture remains the definitive assay to determine viral infectivity and transmission. Using cell culture methods, Wölfel et al revealed that infectious viruses were readily isolated from samples derived from the throat or lung, but not from stool samples, in spite of high concentrations of virus RNA. Blood and urine samples never yielded virus [32]. A recent review by Bhat et al summarized that infectious virus was generally not shed beyond 20 days of the onset of symptoms in most COVID-19 patients, including severely ill and immunocompromised, as indicated by failure to isolate replication-competent virus by viral culture [33]. Cell culture-based plaque reduction neutralization assays and derivatives are the most reliable and accurate methods to determine SARS-CoV-2 neutralization antibody activities [34,35].

The fourth category is the serological technique to detect the specific host response to pathogenic microorganisms causing infection. However, due to the lag time to a detectable host response and the cross-reactivity between similar pathogens, serological methods are rarely used for the rapid detection of infections. Successful examples are the detection of pathogen-specific IgM antibodies for hepatitis A virus and perinatal infections, such as Zika virus. In addition to antibodies (humoral immunity), clinical tests are now available for cellular immunity in infectious diseases,

including interferon- $\gamma$  release assays tuberculosis and cytomegalovirus infections [36,37].

Serological detection has been recognized for its sensitivity in convalescent patients with COVID-19 and plays an important role for understanding the epidemiology of SARS-CoV-2 and emerging variants, including the burden and role of asymptomatic infections [9]. A meta-analysis of diagnostic performance of the serological tests for COVID-19 revealed the impact of assay design and post-symptom-onset intervals. Using combined nucleocapsid (N) and spike (S) protein had a better sensitivity compared to either N or S protein only. Serological tests played an important role in the clinical diagnosis for the later-stage COVID-19 patients. Enzyme-linked immunosorbent assays (ELISA) for detecting total antibodies or targeting combined N and S proteins had a higher diagnostic sensitivity compared to other methods [38], which have been used successfully for contact tracing in the early COVID-19 pandemic [39]. Lichtenegger et al developed a faster live virus assay to quantitatively detect neutralizing antibodies through the early measurement of antibody-mediated intracellular virus reduction by SARS-CoV-2 real-time PCR [40]. Cell culture-based plaque reduction neutralization assays and their derivatives are the most reliable and accurate methods to determine SARS-CoV-2 neutralization antibody activities [34,35].

The last category of infectious disease diagnostics is molecular biology techniques for the detection of pathogen-specific nucleic acids. PCR, an enzyme-mediated *in vitro* nucleic acid amplification method, has become the dominant method in clinical microbiology services, especially for viral infections [15,22]. A group of simple, rapid and integrated molecular devices is gradually replacing antigen testing for immediate diagnosis at the point of care [41,42]. The “syndromic panel” kits incorporating multiplex-PCR have been widely used to identify a range of pathogens with similar symptoms [42,43]. Quantitative pathogenic testing plays a key role in the monitoring of infection treatment. Next-generation nucleic acid sequencing and gene editing technologies are also finding their way into the detection and identification of specific nucleic acids of pathogenic microorganisms [44,45].

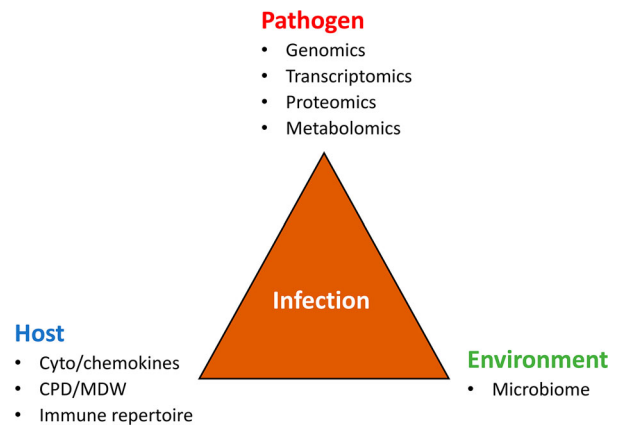
Beyond nucleic acids, matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS), which targets microbial proteomic profiles, has been widely used in the clinical microbiology laboratory for rapid and accurate identification [46,47]. In addition, MALDI-TOF MS has been explored for determining epidemic relatedness and antibiotic resistance of microbial isolates. The value of MALDI-TOF MS for microbial typing was investigated in several studies involving *Staphylococcus aureus* [48,49]. Garrigos and colleagues built a

database using MALDI-TOF MS allowing rapid and accurate species identification and determination of the multi-resistant epidemic clones of *Achromobacter* species in French cystic fibrosis centers [50]. These data indicated that this technology is a potential rapid screening tool for nosocomial infection investigations. Recently, the MALDI-TOF MS technique has been extended to functional identification including determination of antibiotic resistance [51,52]. Youn et al used the MALDI-TOF MS for the rapid detection of potentially *bla*<sub>KPC</sub>-containing carbapenem-resistant isolates, providing early and clinically actionable results [53]. MALDI-TOF MS profiling in combination with growth media containing isotopically labelled amino acids was reported for the detection of resistant microorganisms after hours of incubation [54]. Scientists in Bruker Daltonik and clinical collaborators reported the rapid detection of antibiotic resistance by MALDI-TOF MS using a direct-on-target microdroplet growth assay [55–57]. Detection of colistin resistance became rapid and reliable by use of the MALDI Biotyper Sirius system in both *E. coli* and *P. aeruginosa* [58,59]. These methods being developed based MALDI-TOF MS technology provide an alternative approach to timely monitoring microbial infections [52].

The COVID-19 pandemic ignited the development of numerous nucleic acid amplification methods for the diagnosis and monitoring of SARS-CoV-2 infections. Molecular tests such as real-time PCR are highly sensitive and specific at detecting viral RNA and are now recommended by World Health Organization (WHO) for confirming the diagnosis in individuals who are symptomatic as well as for informing public health decisions. Several newer molecular methods including digital droplet PCR [60] and CRISPR-based assays [61] have been used in COVID-19 diagnostics. Culture accompanied with mNGS increased pathogen diagnostic rate of secondary infections in severe and critical ill COVID-19 patients [62]. Integrated, random-access, point-of-care molecular devices have been developed for fast and accurate diagnosis of SARS-CoV-2 infections in local hospitals and clinics bearing the burden of identifying and treating patients [63,64]. Molecular methods have played significant roles in the discovery and characterization of emerging pathogens such as new world hantavirus [65], influenza A (H7N9) virus [66] and most recently Langya henipavirus [67].

### A bright and evolving future

The development and progression of an infectious disease can be thought of as the result of a tripartite interaction between the pathogen, the host, and the environment (Figure 1). On the pathogen side, techniques will continue to evolve to expand and enhance the capacity for detection and characterization of



**Figure 1.** Infection is the result of the interaction between the pathogen, the host and the environment. CPD, cell population data; MDW, monocyte distribution width.

emerging microbial pathogens. Since the genomes of an increasing number of microbial species/strains have been sequenced, the door has been opened wide for “omic” technologies to be used more broadly [68]. Metagenomic next-generation sequencing (mNGS)-enabled surveillance methods offer the opportunity to improve the detection of both known and yet-to-emerge pathogens such as SARS-CoV-2 and the new Langya henipavirus reported recently [67,69,70]. Wilson et al in 2014 reported the use of NGS to provide an actionable diagnosis of neuroleptospirosis [71]. NGS-based system has been commercially available to determine for HIV-1 genotypic resistance directly in clinical samples [72]. The mNGS testing is a powerful tool that can aid in aetiology diagnosis especially in complicated cases as a rule-out. With limitations understood, mNGS data can be useful with host response information incorporated. Currently, turn-around times remain a major hurdle, although same-day result can be produced using faster system such as Nanopore sequencing [73]. Finally, the clinical relevance of the presence of pathogen-specific nucleic acids in a clinical specimen, which may or may not indicate infection, can be adjudicated by a panel of clinical microbiologists and physicians, much like in done in other areas of pathology [74] or a sequencing stewardship panel [75].

Besides genomics, the laboratory may use transcriptomics, proteomics, and metabolomics [76], each of which may carry potential diagnostic utility. For instance, in *Mycobacterium tuberculosis* infections, because the half-life of mRNA is extremely short as compared with rRNA or genomic DNA, assays that target mycobacterial mRNA better reflect mycobacterial viability, which may be used to monitor the efficacy of anti-TB therapy [77]. Regarding metabolomics, Koo et al reported the use of breath fungal secondary metabolite signatures to diagnose invasive aspergillosis infections [78]. On 14 April 2022, the US FDA granted an emergency use authorization for



the InspectIR COVID-19 Breathalyzer (InspectIR Systems, Frisco, TX, USA), the first COVID-19 diagnostic test using gas chromatography/gas mass-spectrometry to identify chemical mixtures of five volatile organic compounds (VOCs) associated with SARS-CoV-2 infection in exhaled breath, which provides a SARS-CoV-2 identification result in less than three minutes (<https://www.fda.gov/media/157720/download>). A series of metabolomic profiles were reported to detect and characterize culturable and unculturable bacterial pathogens in clinical settings [79–81]. In these and numerous other research studies, multi-omic techniques have been leveraged to generate molecular profiles for the surveillance and management of emerging infections [82,83].

Host response markers also have been explored to facilitate diagnosis of microbial infections. Zhang et al. utilized metatranscriptomics of blood from COVID-19 patients and identified a transcriptional signature of differentially expressed genes significantly associated with immune response to SARS-CoV-2 [84]. Furthermore, Sweeney et al. conducted integrated multi-analyte profiling to yield a three-gene set for testing whole blood specimens that is very robust for diagnosing active tuberculosis (with potential relevance both for diagnosis and treatment monitoring) [82]. Cepheid (Sunnyvale, CA, USA) has recently developed a prototype assay (Xpert MTB® MTB Finger Stick) to detect a three gene host response signature in whole blood specimens using the

GeneXpert® system. The gene signature may be of value both for diagnosis and treatment monitoring. Several studies have demonstrated that the assay fulfils the most of the attributes of the WHO target product profile for a point of care triage test for TB. The assay can be performed on fingerstick blood [85–88]. A novel assay called MeMed BV that integrates measurements of blood-borne host-proteins (tumour necrosis factor-related apoptosis-inducing ligand, interferon  $\gamma$ -induced protein-10, and CRP) was developed and manufactured by MeMed Diagnostics (Tirat Carmel, Israel) to assist in differentiation between bacterial and viral disease [83,89]. A recent prospective, multicenter cohort study performed by Papan et al. validated the high diagnostic performance of the MeMed BV assay in a broad paediatric cohort, and supported its potential to reduce antibiotic over-use in children with viral infections [90].

It is now well-established that the gut microbiota plays a critical role in infection pathogenesis [91,92]. The importance of the human microbiome is becoming increasingly evident. While the contributions of individual microbes in the human health are still far from fully understood, the microbiome appears to play a key role in many vital functions, including synthesizing vitamins and amino acids, generating important metabolites, protecting against pathogens, utilizing non-human biochemical pathways and contributing to the immune system. A large body of evidence demonstrated, more than a decade ago, that



**Figure 2.** Polarized clinical microbiology practice in the near future with rapid, random-access tests done at point of care (left) and with batched, large volumes of tests done at central laboratory (right).

gut microbial alteration is a key factor in the pathogenesis of many local and systemic disorders, including infections [93]. Characterization of the composition of the gut microbiota as well as a dominant pathogen(s) in patients with microbial infections promises to open new avenues for the development of patient-centered personalized and precision diagnosis [94]. One example was to use a robust microbiota-based assay to enable simple diagnostics and disease activity monitoring for inflammatory bowel disease [95]. Laboratory-developed tests are commercially available for microbiome determination in several reference laboratories. For example, the Gut Intelligence Test (Viome Inc., Los Alamos, NM, USA) uses a robust and automated stool metatranscriptomic method offering a rapid and comprehensive taxonomic and functional readout of the gut microbiome [96]. An integrated diagnostic approach combining pathogen, host and microbiome would enhance the speed and accuracy for the laboratory diagnosis and monitoring of microbial infections (Figure 1). Among them, machine learning will gradually apply in clinical microbiology practice especially for unusual emerging pathogens including predicting drug targets or vaccine candidates, diagnosing microorganisms causing epidemics, predicting disease outbreaks and exploring microbial interactions [97–99]. Machine learning has been used in the clinical setting for classifying drug resistance against antimicrobial agents [100,101].

In the future, there will be a need for more rapid diagnoses, increased standardization of testing and greater adaptability to cope with new threats from emerging microbial pathogens. As early as 2004, Didier Raoult, the renowned French medical and clinical microbiologist, jointly with his team, proposed a bipolarization of future clinical microbiology services [102] (Figure 2). On one side, clinical microbiology practice will follow the general trend in the life sciences for large, centralized laboratories with the capacity to analyze large numbers of samples and to carry out a wide range of techniques. Total automation has been gradually achieved for bacterial culture, identification and antimicrobial susceptibility testing [23]. For molecular diagnosis, molecular platforms are increasingly designed with an emphasis on automation and sample-to-result capabilities. These include technologies by Roche (4800/6800/8800 platforms) [103], Abbott Laboratories (m2000 and Alinity m) [104], Hologic (Panther and Panther-Fusion) [105], Becton Dickinson (BDMax and BD COR) [106], and Cepheid (GeneXpert Infinity) [107].

On the other side, rapid and on-demand testing is performed at point of care (POC) based with relatively low throughput testing volumes. These include rapid screening for influenza in the emergency room and rapid mixed-sample screening for new coronaviruses

in the field [41,108]. The current widely accepted definition of POC testing includes testing that occurs at or near the point of patient care, such that the results drive patient care decisions made during that encounter [109,110]. POC tests can be performed in a variety of settings including physician offices, emergency department, urgent care facilities, school health clinics and pharmacies. Recently, the COVID-19 pandemic has shined a spotlight on Clinical and Laboratory Improvement Amendments (CLIA)-waived diagnostic testing. Some SARS-CoV-2 tests have received Emergency Use Authorization (EUA) from the U.S. Food and Drug Administration (FDA) for use in CLIA-waived testing sites [108]. Further optimization and validation, new technologies, as well as studies to determine clinical and epidemiological impact of SARS-CoV-2 POC tests are needed. Nevertheless, random-access, integrated devices available at the point of care with scalable capacities will increase its weight in the rapid and accurate diagnosis and monitoring of emerging pathogens in the near future.

In summary, the discipline of clinical microbiology, with a long and uneven past, has been thriving at present and is ready to embrace the future. There have been substantial changes in the role of clinical microbiology laboratories over the past decade. The ongoing technological revolution has rapidly transformed research, diagnostic and therapeutic tools. In the near future, clinical microbiology practice will be able to help clinicians implement real-time evidence-based treatments or significantly shorten the process from empirical to evidence-based treatments. While continuously strengthening its scientific attribution, clinical microbiology will be endowed with a more distinct, more profound, more ambitious medical landscape, social value and management significance. Full of challenges and uncertainties, though, the practice of clinical microbiology will keep abreast of the times, promising and never fading, for its demand-oriented, significance-driven, logic and science-based, and humanitarian characteristics.

## Acknowledgements

The authors thank Monica Wang, Yongzhong Ning and Jingwen Ai for their assistance and Charles Stratton and Sherry Dunbar for their constructive discussion and critical reviewing the manuscript. The Xpert MTB® MTB Finger Stick Prototype is a product in development, which is not for use in diagnostic procedures and not reviewed by any regulatory body.

## Disclosure statement

Y.-W.T. is an employee of Cepheid, the commercial manufacturer of the GeneXpert system and Xpert cartridges. No potential conflict of interest was reported by the rest authors.

## ORCID

Yi-Wei Tang  <http://orcid.org/0000-0003-4888-6771>

## References

- [1] Cintron M, Hauser JR, Otto C, et al. Diagnostic microbiology. In: Schmidt TM, editor. Encyclopedia of microbiology, 4th ed. Oxford: Elsevier Press; 2019. p. 1–17.
- [2] Isenberg HD. Clinical microbiology: past, present, and future. J Clin Microbiol. 2003;41:917–918. doi:10.1128/JCM.41.3.917-918.2003.
- [3] Hodgkin T, Lister JJ. Notice of some microscopic observations of the blood and animal tissues. Phil Mag. 1827;32:130–138.
- [4] Tang YW, Sussman M, Liu D, et al. Molecular medical microbiology, 2 ed. Boston (MA): Elsevier; 2014.
- [5] Hitchens AP, Leikind MC. The introduction of agar-agar into bacteriology. J Bacteriol. 1939;37:485–493. doi:10.1128/jb.37.5.485-493.1939.
- [6] Lagier JC, Edouard S, Pagnier I, et al. Current and past strategies for bacterial culture in clinical microbiology. Clin Microbiol Rev. 2015;28:208–236. doi:10.1128/CMR.00110-14.
- [7] Enders JF, Weller TH, Robbins FC. Cultivation of the Lansing strain of poliomyelitis virus in cultures of various human embryonic tissues. Science. 1949;109:85–87. doi:10.1126/science.109.2822.85.
- [8] Tang F, Zhang X, Huang Y, et al. Study on the pathogen of trachoma IV. Attempt to isolate the virus in the embryonated hens eggs. Acta Microbiol Sin. 1956;2:189–210.
- [9] Loeffelholz MJ, Tang YW. Laboratory diagnosis of emerging human coronavirus infections – the state of the art. Emerg Microbes Infect. 2020;9:747–756. doi:10.1080/22221751.2020.1745095.
- [10] Khalil N, Bernstein DI. Influenza vaccines: where we are, where we are going. Curr Opin Pediatr. 2022;34:119–125. doi:10.1097/MOP.0000000000001103.
- [11] Sternberg GM. Practical results of bacteriological researches. Tr A Am Physicians. 1892;7:68–86.
- [12] Ayres JC, Feemster RF. Serologic tests in the diagnosis of infectious diseases. Concluded. N Engl J Med. 1950;243:1034–1043. doi:10.1056/NEJM195012282432606.
- [13] Ayres JC, Feemster RF. Serologic tests in the diagnosis of infectious diseases. Part 1. N Engl J Med. 1950;243:996–1002. doi:10.1056/NEJM195012212432505.
- [14] Mullis KB, Faloona FA. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. Methods Enzymol. 1987;155:335–350. doi:10.1016/0076-6879(87)55023-6.
- [15] Schmitz JE, Stratton CW, Persing DH, et al. Forty years of molecular diagnostics for infectious diseases. J Clin Microbiol. 2022;19:02446–02421.
- [16] Baltimore D. RNA-dependent DNA polymerase in virions of RNA tumour viruses. Nature. 1970;226:1209–1211. doi:10.1038/2261209a0.
- [17] Temin HM, Mizutani S. RNA-dependent DNA polymerase in virions of Rous sarcoma virus. Nature. 1970;226:1211–1213. doi:10.1038/2261211a0.
- [18] Higuchi R, Fockler C, Dollinger G, et al. Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. Biotechnol (NY). 1993;11:1026–1030. doi:10.1038/nbt0993-1026.
- [19] Fleischmann RD, Adams MD, White O, et al. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science. 1995;269:496–512. doi:10.1126/science.7542800.
- [20] Fraser CM, Gocayne JD, White O, et al. The minimal gene complement of *Mycoplasma genitalium*. Science. 1995;270:397–403. doi:10.1126/science.270.5235.397.
- [21] Sanger F, Coulson AR. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J Mol Biol. 1975;94:441–448. doi:10.1016/0022-2836(75)90213-2.
- [22] Tang YW, Procop GW, Persing DH. Molecular diagnostics of infectious diseases. Clin Chem. 1997;43:2021–2038.
- [23] Bourbeau PP, Ledebor NA. Automation in clinical microbiology. J Clin Microbiol. 2013;51:1658–1665. doi:10.1128/JCM.00301-13. Epub 2013 Mar 20.
- [24] Gerber MA. Diagnosis of group A beta-hemolytic streptococcal pharyngitis. Use of antigen detection tests. Diagn Microbiol Infect Dis. 1986;4:5S–15S. doi:10.1016/s0732-8893(86)80038-4.
- [25] Smith TF, Wold AD, Espy MJ, et al. New developments in the diagnosis of viral diseases. Infect Dis Clin North Am. 1993;7:183–201.
- [26] Bellew S, Grijalva CG, Williams DJ, et al. Pneumococcal and legionella urinary antigen tests in community-acquired pneumonia: prospective evaluation of indications for testing. Clin Infect Dis. 2019;68:2026–2033. doi:10.1093/cid/ciy826.
- [27] Khalid MF, Selvam K, Jeffry AJN, et al. Performance of rapid antigen tests for COVID-19 diagnosis: A systematic review and meta-analysis. Diagnostics (Basel). 2022;12:110. doi:10.3390/diagnostics12010110.
- [28] Walsh KA, Broderick N, Ahern S, et al. Effectiveness of rapid antigen testing for screening of asymptomatic individuals to limit the transmission of SARS-CoV-2: A rapid review. Rev Med Virol. 2022;29:e2350.
- [29] Harcourt J, Tamin A, Lu X, et al. Severe acute respiratory syndrome coronavirus 2 from patient with coronavirus disease, United States. Emerg Infect Dis. 2020;26:1266–1273. doi:10.3201/eid2606.200516. Epub 2020 Jun 17.
- [30] Heinen N, Klöhn M, Steinmann E, et al. In vitro lung models and their application to study SARS-CoV-2 pathogenesis and disease. Viruses. 2021;13:792. doi:10.3390/v13050792.
- [31] Matsuyama S, Nao N, Shirato K, et al. Enhanced isolation of SARS-CoV-2 by TMPRSS2-expressing cells. Proc Natl Acad Sci U S A. 2020;117:7001–7003. doi:10.1073/pnas.2002589117. Epub 2020 Mar 12.
- [32] Wölfel R, Corman VM, Guggemos W, et al. Virological assessment of hospitalized patients with COVID-2019. Nature. 2020;581:465–469. doi:10.1038/s41586-020-2196-x. Epub 2020 Apr 1.
- [33] Bhat V, Chavan P, Khattry N, et al. Dynamics of viral RNA load, virus culture, seroconversion & infectivity in COVID-19 patients: implications on isolation policy. Indian J Med Res. 2021;153:585–590. doi:10.4103/ijmr.IJMR\_3564\_20.
- [34] Banga Ndzouboukou JL, Zhang YD, Fan XL. Recent developments in SARS-CoV-2 neutralizing antibody detection methods. Curr Med Sci. 2021;41:1052–1064. doi:10.1007/s11596-021-2470-7. Epub 2021 Dec 21.



- [35] Muruato AE, Fontes-Garfias CR, Ren P, et al. A high-throughput neutralizing antibody assay for COVID-19 diagnosis and vaccine evaluation. *Nat Commun.* 2020;11:4059. doi:10.1038/s41467-020-17892-0.
- [36] Manuel O, Husain S, Kumar D, et al. Assessment of cytomegalovirus-specific cell-mediated immunity for the prediction of cytomegalovirus disease in high-risk solid-organ transplant recipients: a multicenter cohort study. *Clin Infect Dis.* 2013;56:817–824. doi:10.1093/cid/cis993. Epub 2012 Nov 29.
- [37] Pai M, Denkinger CM, Kik SV, et al. Gamma interferon release assays for detection of *Mycobacterium tuberculosis* infection. *Clin Microbiol Rev.* 2014;27:3–20. doi:10.1128/CMR.00034-13.
- [38] Wang H, Ai J, Loeffelholz MJ, et al. Meta-analysis of diagnostic performance of serology tests for COVID-19: impact of assay design and post-symptom-onset intervals. *Emerg Microbes Infect.* 2020;9:2200–2211. doi:10.1080/22221751.2020.1826362.
- [39] Peeling RW, Wedderburn CJ, Garcia PJ, et al. Serology testing in the COVID-19 pandemic response. *Lancet Infect Dis.* 2020;20:e245–e249. doi:10.1016/S1473-3099(20)30517-X. Epub 2020 Jul 17.
- [40] Lichtenegger S, Saiger S, Hardt M, Kulnik S, Wagner GE, Kleinhappl B, Assig K, Zauner A, Ober M, Kimpel J, von Laer D, Zatloukal K, Steinmetz I. Development of a Rapid Live SARS-CoV-2 Neutralization Assay Based on a qPCR Readout. *J Clin Microbiol.* 2022;60:e0037622. doi:10.1128/jcm.00376-22.
- [41] Kozel TR, Burnham-Marusch AR. Point-of-Care testing for infectious diseases: past, present, and future. *J Clin Microbiol.* 2017;55:2313–2320. doi:10.1128/JCM.00476-17. Epub 2017 May 24.
- [42] Ling L, Kaplan SE, Lopez JC, et al. Parallel validation of three molecular devices for simultaneous detection and identification of influenza A and B and respiratory syncytial viruses. *J Clin Microbiol.* 2018;56:e01691–17. doi:10.1128/JCM.01691-17. Print 2018 Mar.
- [43] Abbott AN, Fang FC. Clinical impact of multiplex syndromic panels in the diagnosis of bloodstream, gastrointestinal, respiratory, and central nervous system infections. *Clin Microbiol Newsl.* 2017;39:133–142.
- [44] Gootenberg JS, Abudayyeh OO, Lee JW, et al. Nucleic acid detection with CRISPR-Cas13a/C2c2. *Science.* 2017;356:438–442. doi:10.1126/science.aam9321. Epub 2017 Apr 13.
- [45] Chiu CY, Miller SA. Clinical metagenomics. *Nat Rev Genet.* 2019;20:341–355. doi:10.1038/s41576-019-0113-7.
- [46] Bizzini A, Greub G. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry, a revolution in clinical microbial identification. *Clin Microbiol Infect.* 2010;16:1614–1619. doi:10.1111/j.1469-0691.2010.03311.x.
- [47] Torres-Sangiao E, Leal Rodriguez C, García-Riestra C. Application and perspectives of MALDI-TOF mass spectrometry in clinical microbiology laboratories. *Microorganisms.* 2021;9:1539. doi:10.3390/microorganisms9071539.
- [48] Schlebusch S, Price GR, Hinds S, et al. First outbreak of PVL-positive nonmultiresistant MRSA in a neonatal ICU in Australia: comparison of MALDI-TOF and SNP-plus-binary gene typing. *Eur J Clin Microbiol Infect Dis.* 2010;29:1311–1314.
- [49] Wolters M, Rohde H, Maier T, et al. MALDI-TOF MS fingerprinting allows for discrimination of major methicillin-resistant *staphylococcus aureus* lineages. *Int J Med Microbiol.* 2011;301:64–68.
- [50] Garrigos T, Dollat M, Magallon A, et al. Distribution of *Achromobacter* species in 12 French cystic fibrosis centers in 2020 by a retrospective MALDI-TOF MS spectrum analysis. *J Clin Microbiol.* 2022;60:e0242221. doi:10.1128/jcm.02422-21. Epub 2022 May 5.
- [51] Hrabak J, Walkova R, Studentova V, et al. Carbapenemase activity detection by matrix-assisted laser desorption ionization-time of flight mass spectrometry. *J Clin Microbiol.* 2011;49:3222–3227.
- [52] Yoon EJ, Jeong SH. MALDI-TOF mass spectrometry technology as a tool for the rapid diagnosis of antimicrobial resistance in bacteria. *Antibiotics (Basel).* 2021;10:982. doi:10.3390/antibiotics10080982.
- [53] Youn JH, Drake SK, Weingarten RA, et al. Clinical performance of a matrix-assisted laser desorption ionization-time of flight mass spectrometry method for detection of certain blaKPC-containing plasmids. *J Clin Microbiol.* 2016;54:35–42. doi:10.1128/JCM.01643-15. Epub 2015 Sep 2.
- [54] Sparbier K, Lange C, Jung J, et al. MALDI biotyper-based rapid resistance detection by stable-isotope labeling. *J Clin Microbiol.* 2013;51:3741–3748. doi:10.1128/JCM.01536-13. Epub 2013 Sep 4.
- [55] Correa-Martínez CL, Idelevich EA, Sparbier K, et al. Rapid detection of extended-spectrum  $\beta$ -lactamases (ESBL) and AmpC  $\beta$ -lactamases in enterobacterales: development of a screening panel using the MALDI-TOF MS-based direct-on-target microdroplet growth assay. *Front Microbiol.* 2019;10:13. doi:10.3389/fmicb.2019.00013. eCollection 2019.
- [56] Correa-Martínez CL, Idelevich EA, Sparbier K, et al. Development of a MALDI-TOF MS-based screening panel for accelerated differential detection of carbapenemases in enterobacterales using the direct-on-target microdroplet growth assay. *Sci Rep.* 2020;10:4988. doi:10.1038/s41598-020-61890-7.
- [57] Idelevich EA, Sparbier K, Kostrzewa M, et al. Rapid detection of antibiotic resistance by MALDI-TOF mass spectrometry using a novel direct-on-target microdroplet growth assay. *Clin Microbiol Infect.* 2018;24:738–743. doi:10.1016/j.cmi.2017.10.016. Epub 2017 Oct 24.
- [58] Furniss RCD, Dortet L, Bolland W, et al. Detection of colistin resistance in *Escherichia coli* by use of the MALDI biotyper sirius mass spectrometry system. *J Clin Microbiol.* 2019;57:e01427–19. doi:10.1128/JCM.01427-19. Print 2019 Dec.
- [59] Jeannot K, Hagart K, Dortet L, et al. Detection of colistin resistance in *Pseudomonas aeruginosa* using the MALDIxin test on the routine MALDI biotyper sirius mass spectrometer. *Front Microbiol.* 2021;12:725383. doi:10.3389/fmicb.2021.725383. eCollection 2021.
- [60] Suo T, Liu X, Feng J, et al. ddPCR: a more accurate tool for SARS-CoV-2 detection in low viral load specimens. *Emerg Microbes Infect.* 2020;9:1259–1268. doi:10.1080/22221751.2020.1772678.
- [61] Broughton JP, Deng X, Yu G, et al. CRISPR-Cas12-based detection of SARS-CoV-2. *Nat Biotechnol.* 2020;38:870–874. doi:10.1038/s41587-020-0513-4. Epub 2020 Apr 16.
- [62] Zhang H, Zhang Y, Wu J, et al. Risks and features of secondary infections in severe and critical ill



- COVID-19 patients. *Emerg Microbes Infect.* 2020;9:1958–1964. doi:10.1080/22221751.2020.1812437.
- [63] Loeffelholz MJ, Tang YW. Detection of SARS-CoV-2 at the point of care. *Bioanalysis.* 2021;13:1213–1223. doi:10.4155/bio-2021-0078. Epub 2021 Jul 22.
- [64] Hou H, Chen J, Wang Y, et al. Multicenter evaluation of the cepheid xpert xpress SARS-CoV-2 assay for the detection of SARS-CoV-2 in oropharyngeal swab specimens. *J Clin Microbiol.* 2020;58:e01288–20. doi:10.1128/JCM.01288-20. Print 2020 Jul 23.
- [65] Nichol ST, Spiropoulou CF, Morzunov S, et al. Genetic identification of a hantavirus associated with an outbreak of acute respiratory illness. *Science.* 1993;262:914–917. doi:10.1126/science.8235615.
- [66] Gao R, Cao B, Hu Y, et al. Human infection with a novel avian-origin influenza A (H7N9) virus. *N Engl J Med.* 2013;368:1888–1897. doi:10.1056/NEJMoa1304459. Epub 2013 Apr 11.
- [67] Zhang XA, Li H, Jiang FC, et al. A zoonotic henipavirus in febrile patients in China. *N Engl J Med.* 2022;387:470–472. doi:10.1056/NEJMc2202705.
- [68] Lederberg J, McCray AT. Ome sweet 'omics – a genealogical treasury of words. *The Scientist.* 2001;15:8.
- [69] Ko KKK, Chng KR, Nagarajan N. Metagenomics-enabled microbial surveillance. *Nat Microbiol.* 2022;7:486–496. doi:10.1038/s41564-022-01089-w. Epub 2022 Apr 1.
- [70] Carbo EC, Sidorov IA, Zevenhoven-Dobbe JC, et al. Coronavirus discovery by metagenomic sequencing: a tool for pandemic preparedness. *J Clin Virol.* 2020;131:104594. doi:10.1016/j.jcv.2020.104594. Epub 2020 Aug 21.
- [71] Wilson MR, Naccache SN, Samayoa E, et al. Actionable diagnosis of neuroleptospirosis by next-generation sequencing. *N Engl J Med.* 2014;370:2408–2417. doi:10.1056/NEJMoa1401268. Epub 2014 Jun 4.
- [72] Raymond S, Nicot F, Carcenac R, et al. HIV-1 genotypic resistance testing using the vela automated next-generation sequencing platform. *J Antimicrob Chemother.* 2018;73:1152–1157. doi:10.1093/jac/dky003.
- [73] Stratton CW, Tang YW. Diagnosing bacteremia in real time using next-generation sequencing-based technology. *J Mol Diagn.* 2020;22:301–303. doi:10.1016/j.jmoldx.2020.01.002. Epub 2020 Jan 21.
- [74] Wilson MR, Sample HA, Zorn KC, et al. Clinical metagenomic sequencing for diagnosis of meningitis and encephalitis. *N Engl J Med.* 2019;380:2327–2340. doi:10.1056/NEJMoa1803396.
- [75] Stratton CW, Schutzbank TE, Tang YW. Use of metagenomic next-generation sequencing in the clinical microbiology laboratory: A step forward, but not an end-all. *J Mol Diagn.* 2021;23:1415–1421. doi:10.1016/j.jmoldx.2021.09.003.
- [76] Kiechle FL, Holland-Staley CA. Genomics, transcriptomics, proteomics, and numbers. *Arch Pathol Lab Med.* 2003;127:1089–1097. doi:10.5858/2003-127-1089-GTPAN.
- [77] Mdivani N, Li H, Akhalaia M, et al. Monitoring therapeutic efficacy by real-time detection of mycobacterium tuberculosis mRNA in sputum. *Clin Chem.* 2009;55:1694–1700. doi:10.1373/clinchem.2009.124396. Epub 2009 Jul 2.
- [78] Koo S, Thomas HR, Daniels SD, et al. A breath fungal secondary metabolite signature to diagnose invasive aspergillosis. *Clin Infect Dis.* 2014;59:1733–1740. doi:10.1093/cid/ciu725. Epub 2014 Oct 22.
- [79] La Rosa R, Johansen HK, Molin S. Convergent metabolic specialization through distinct evolutionary paths in *Pseudomonas aeruginosa*. *mBio.* 2018;9:e00269–18. doi:10.1128/mBio.00269-18.
- [80] Horrocks V, Hind CK, Wand ME, et al. Nuclear magnetic resonance metabolomics of symbioses between bacterial vaginosis-associated bacteria. *mSphere.* 2022;7:e0016622. doi:10.1128/msphere.00166-22. Epub 2022 May 2.
- [81] Slupsky CM, Rankin KN, Fu H, et al. Pneumococcal pneumonia: potential for diagnosis through a urinary metabolic profile. *J Proteome Res.* 2009;8:5550–5558. doi:10.1021/pr9006427.
- [82] Sweeney TE, Braviak L, Tato CM, et al. Genome-wide expression for diagnosis of pulmonary tuberculosis: a multicohort analysis. *Lancet Respir Med.* 2016;4:213–224. doi:10.1016/S2213-2600(16)00048-5. Epub 2016 Feb 20.
- [83] van Houten CB, de Groot JAH, Klein A, et al. A host-protein based assay to differentiate between bacterial and viral infections in preschool children (OPPORTUNITY): a double-blind, multicentre, validation study. *Lancet Infect Dis.* 2017;17:431–440. doi:10.1016/S1473-3099(16)30519-9. Epub 2016 Dec 22.
- [84] Zhang H, Ai JW, Yang W, et al. Metatranscriptomic characterization of coronavirus disease 2019 identified a host transcriptional classifier associated with immune signaling. *Clin Infect Dis.* 2021;73:376–385. doi:10.1093/cid/ciaa663.
- [85] Moreira FMF, Verma R, Pereira Dos Santos PC, et al. Blood-based host biomarker diagnostics in active case finding for pulmonary tuberculosis: A diagnostic case-control study. *EClinicalMedicine.* 2021;33:100776. doi:10.1016/j.eclinm.2021.100776. eCollection 2021 Mar.
- [86] Södersten E, Ongarello S, Mantsoki A, et al. Diagnostic accuracy study of a novel blood-based assay for identification of tuberculosis in people living with HIV. *J Clin Microbiol.* 2021;59:e01643–20. doi:10.1128/JCM.01643-20. Print 2021 Feb 18.
- [87] Sutherland JS, van der Spuy G, Gindeh A, et al. Diagnostic accuracy of the cepheid 3-gene host response fingerstick blood test in a prospective, multi-site study: interim results. *Clin Infect Dis.* 2022;74:2136–2141. doi:10.1093/cid/ciab839.
- [88] Zimmer AJ, Schumacher SG, Södersten E, et al. A novel blood-based assay for treatment monitoring of tuberculosis. *BMC Res Notes.* 2021;14:247. doi:10.1186/s13104-021-05663-z.
- [89] Oved K, Cohen A, Boico O, et al. A novel host-proteome signature for distinguishing between acute bacterial and viral infections. *PLoS One.* 2015;10:e0120012. doi:10.1371/journal.pone.0120012. eCollection 2015.
- [90] Papan C, Argentiero A, Porwoll M, et al. A host signature based on TRAIL, IP-10, and CRP for reducing antibiotic overuse in children by differentiating bacterial from viral infections: a prospective, multicentre cohort study. *Clin Microbiol Infect.* 2021;10:00621–00622.
- [91] Boyd JH, Russell JA, Fjell CD. The meta-genome of sepsis: host genetics, pathogens and the acute

- immune response. *J Innate Immun.* 2014;6:272–283. doi:10.1159/000358835. Epub 2014 Feb 12.
- [92] Taur Y, Pamer EG. Microbiome mediation of infections in the cancer setting. *Genome Med.* 2016;8:40. doi:10.1186/s13073-016-0306-z.
- [93] Libertucci J, Young VB. The role of the microbiota in infectious diseases. *Nat Microbiol.* 2019;4:35–45. doi:10.1038/s41564-018-0278-4. Epub 2018 Dec 13.
- [94] Gebreyal P, Nicco C, Al Khodor S, et al. Microbiota medicine: towards clinical revolution. *J Transl Med.* 2022;20:111. doi:10.1186/s12967-022-03296-9.
- [95] Eck A, de Groot EFJ, de Meij TGJ, et al. Robust microbiota-based diagnostics for inflammatory bowel disease. *J Clin Microbiol.* 2017;55:1720–1732. doi:10.1128/JCM.00162-17. Epub 2017 Mar 22.
- [96] Hatch A, Horne J, Toma R, et al. A robust metatranscriptomic technology for population-scale studies of diet, gut microbiome, and human health. *Int J Genomics.* 2019;2019:1718741. doi:10.1155/2019/1718741. eCollection 2019.
- [97] Chiu HR, Hwang CK, Chen SY, et al. Machine learning for emerging infectious disease field responses. *Sci Rep.* 2022;12:328. doi:10.1038/s41598-021-03687-w.
- [98] Goodswen SJ, Barratt JLN, Kennedy PJ, et al. Machine learning and applications in microbiology. *FEMS Microbiol Rev.* 2021;45:fuab015. doi:10.1093/femsre/fuab015.
- [99] Wong OGW, Ng IFY, Tsun OKL, et al. Machine learning interpretation of extended human papillomavirus genotyping by onclarity in an Asian cervical cancer screening population. *J Clin Microbiol.* 2019;57:e00997–19. doi:10.1128/JCM.00997-19. Print 2019 Dec.
- [100] Anahtar MN, Yang JH, Kanjilal S. Applications of machine learning to the problem of antimicrobial resistance: an emerging model for translational research. *J Clin Microbiol.* 2021;59:e0126020. doi:10.1128/JCM.01260-20. Epub 2021 Jun 18.
- [101] Nguyen M, Long SW, McDermott PF, et al. Using machine learning To predict antimicrobial MICs and associated genomic features for nontyphoidal salmonella. *J Clin Microbiol.* 2019;57:e01260–18. doi:10.1128/JCM.01260-18. Print 2019 Feb.
- [102] Raoult D, Fournier PE, Drancourt M. What does the future hold for clinical microbiology? *Nat Rev Microbiol.* 2004;2:151–159. doi:10.1038/nrmicro820.
- [103] Cobb B, Simon CO, Stramer SL, et al. The cobas® 6800/8800 system: a new era of automation in molecular diagnostics. *Expert Rev Mol Diagn.* 2017;17:167–180. doi:10.1080/14737159.2017.1275962.
- [104] Hirschhorn JW, Kegl A, Dickerson T, et al. Verification and validation of SARS-CoV-2 assay performance on the Abbott m2000 and alinity m systems. *J Clin Microbiol.* 2021;59:e03119–20. doi:10.1128/JCM.03119-20. Print 2021 Apr 20.
- [105] Stellrecht KA, Cimino JL, Wilson LI, et al. Panther fusion® respiratory virus assays for the detection of influenza and other respiratory viruses. *J Clin Virol.* 2019;121:104204. doi:10.1016/j.jcv.2019.104204. Epub 2019 Nov 9.
- [106] Dalpke AH, Hofko M, Zorn M, et al. Evaluation of the fully automated BD MAX Cdiff and Xpert C. difficile assays for direct detection of clostridium difficile in stool specimens. *J Clin Microbiol.* 2013;51:1906–1908. doi:10.1128/JCM.00344-13. Epub 2013 Mar 20.
- [107] Jang D, Ratnam S, Gilchrist J, et al. Comparison of workflow, maintenance, and consumables in the GeneXpert infinity 80 and panther instruments while testing for *Chlamydia trachomatis* and *Neisseria gonorrhoeae*. *Sex Transm Dis.* 2016;43:377–381. doi:10.1097/OLQ.0000000000000444.
- [108] Loeffelholz MJ, Tang YW. Detection of SARS-CoV-2 infections at the point of care. *Bioanalysis.* 2021;13:1213–1223. doi:10.4155/bio-2021-0078. Epub 2021 Jul 22.
- [109] Chen H, Liu K, Li Z, et al. Point of care testing for infectious diseases. *Clin Chim Acta.* 2019;493:138–147. doi:10.1016/j.cca.2019.03.008. Epub 2019 Mar 8.
- [110] Drancourt M, Michel-Lepage A, Boyer S, et al. The point-of-care laboratory in clinical microbiology. *Clin Microbiol Rev.* 2016;29:429–447. doi:10.1128/CMR.00090-15.

22. C. J. Pedersen, U.S. Patent 3,232,914 (1 February 1966) to Du Pont.  
 23. ———, U.S. Patent 3,320,214 (16 May 1967) to Du Pont.  
 24. ———, U.S. Patent 3,361,778 (2 January 1968) to Du Pont.  
 25. ———, *J. Am. Chem. Soc.* **89**, 2495 (1967); *ibid.*, p. 7017.  
 26. ———, *Aldrichimica Acta* **4** (no. 1), 1 (1971).  
 27. M. R. Truter and C. J. Pedersen, *Endeavour* **30**, 142 (1971).  
 28. C. J. Pedersen, *Fed. Proc.* **27**, 1305 (1968).  
 29. ———, *J. Am. Chem. Soc.* **92**, 386 (1970).  
 30. ———, *ibid.*, p. 391.  
 31. ———, *J. Org. Chem.* **36**, 254 (1971).

32. ———, *ibid.*, p. 1690.  
 33. ——— and H. K. Frensdorff, *Angew. Chem.* **84**, 16 (1972); *Angew. Chem. Int. Ed. Engl.* **11**, 16 (1972).  
 34. C. J. Pedersen, *Org. Synth.* **52**, 66 (1972).  
 35. ———, U.S. Patents 3,562,295 (9 February 1971); 3,622,577 (23 November 1971); 3,686,225 (22 August 1972); 3,687,978 (29 August 1972); ——— and M. Bromels, U.S. Patent 3,847,949 (12 November 1974); C. J. Pedersen, U.S. Patent 3,856,813 (24 December 1974); 3,873,569 (25 March 1975); 3,987,061 (19 October 1976); 3,998,838 (21 December 1976); ——— and M. Bromels, U.S. Patent 4,031,111 (21 June 1977), all to Du Pont.

# Social Relationships and Health

JAMES S. HOUSE, KARL R. LANDIS, DEBRA UMBERSON

Recent scientific work has established both a theoretical basis and strong empirical evidence for a causal impact of social relationships on health. Prospective studies, which control for baseline health status, consistently show increased risk of death among persons with a low quantity, and sometimes low quality, of social relationships. Experimental and quasi-experimental studies of humans and animals also suggest that social isolation is a major risk factor for mortality from widely varying causes. The mechanisms through which social relationships affect health and the factors that promote or inhibit the development and maintenance of social relationships remain to be explored.

... my father told me of a careful observer, who certainly had heart-disease and died from it, and who positively stated that his pulse was habitually irregular to an extreme degree; yet to his great disappointment it invariably became regular as soon as my father entered the room.—Charles Darwin (1)

SCIENTISTS HAVE LONG NOTED AN ASSOCIATION BETWEEN social relationships and health. More socially isolated or less socially integrated individuals are less healthy, psychologically and physically, and more likely to die. The first major work of empirical sociology found that less socially integrated people were more likely to commit suicide than the most integrated (2). In subsequent epidemiologic research age-adjusted mortality rates from all causes of death are consistently higher among the unmarried than the married (3–5). Unmarried and more socially isolated people have also manifested higher rates of tuberculosis (6), accidents (7), and psychiatric disorders such as schizophrenia (8, 9). And as the above quote from Darwin suggests, clinicians have also observed potentially health-enhancing qualities of social relationships and contacts.

The causal interpretation and explanation of these associations

J. S. House is professor and chair of sociology and a research scientist in the Survey Research Center of the Institute for Social Research, Institute of Gerontology, and Department of Epidemiology at the University of Michigan, Ann Arbor, MI 48109. K. R. Landis is a doctoral candidate in the Department of Sociology and research assistant in the Survey Research Center. D. Umberson is a postdoctoral fellow in the Survey Research Center at the University of Michigan and assistant professor-designate of sociology at the University of Texas, Austin.

has, however, been less clear. Does a lack of social relationships cause people to become ill or die? Or are unhealthy people less likely to establish and maintain social relationships? Or is there some other factor, such as a misanthropic personality, which predisposes people both to have a lower quantity or quality of social relationships and to become ill or die?

Such questions have been largely unanswerable before the last decade for two reasons. First, there was little theoretical basis for causal explanation. Durkheim (2) proposed a theory of how social relationships affected suicide, but this theory did not generalize to morbidity and mortality from other causes. Second, evidence of the association between social relationships and health, especially in general human populations, was almost entirely retrospective or cross-sectional before the late 1970s. Retrospective studies from death certificates or hospital records ascertained the nature of a person's social relationships after they had become ill or died, and cross-sectional surveys of general populations determined whether people who reported ill health also reported a lower quality or quantity of relationships. Such studies used statistical control of potential confounding variables to rule out third factors that might produce the association between social relationships and health, but could do this only partially. They could not determine whether poor social relationships preceded or followed ill health.

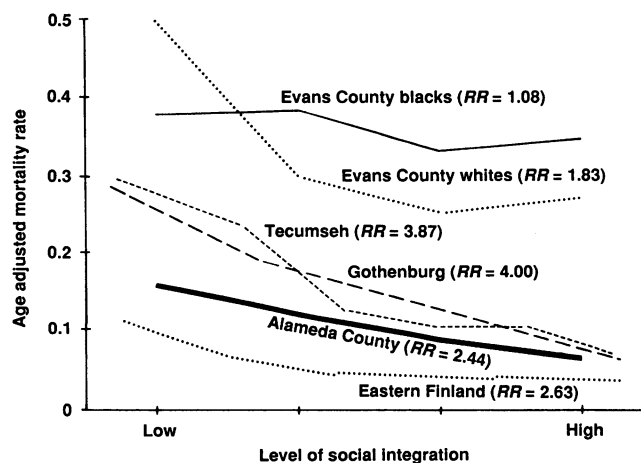


Fig. 1. Level of social integration and age-adjusted mortality for males in five prospective studies. RR, the relative risk ratio of mortality at the lowest versus highest level of social integration.

In this article, we review recent developments that have altered this state of affairs dramatically: (i) emergence of theoretical models for a causal effect of social relationships on health in humans and animals; (ii) cumulation of empirical evidence that social relationships are a consequential predictor of mortality in human populations; and (iii) increasing evidence for the causal impact of social relationships on psychological and physiological functioning in quasi-experimental and experimental studies of humans and animals. These developments suggest that social relationships, or the relative lack thereof, constitute a major risk factor for health—rivaling the effects of well-established health risk factors such as cigarette smoking, blood pressure, blood lipids, obesity, and physical activity. Indeed, the theory and evidence on social relationships and health increasingly approximate that available at the time of the U.S. Surgeon General's 1964 report on smoking and health (10), with similar implications for future research and public policy.

## The Emergence of “Social Support” Theory and Research

The study of social relationships and health was revitalized in the middle 1970s by the emergence of a seemingly new field of scientific research on “social support.” This concept was first used in the mental health literature (11, 12), and was linked to physical health in separate seminal articles by physician-epidemiologists Cassel (13) and Cobb (14). These articles grew out of a rapidly developing literature on stress and psychosocial factors in the etiology of health and illness (15). Chronic diseases have increasingly replaced acute infectious diseases as the major causes of disability and death, at least in industrialized countries. Consequently, theories of disease etiology have shifted from ones in which a single factor (usually a microbe) caused a single disease, to ones in which multiple behavioral and environmental as well as biologic and genetic factors combine, often over extended periods, to produce any single disease, with a given factor often playing an etiologic role in multiple diseases.

Cassel (13) and Cobb (14) reviewed more than 30 human and animal studies that found social relationships protective of health. Recognizing that any one study was open to alternative interpretations, they argued that the variety of study designs (ranging from retrospective to experimental), of life stages studied (from birth to death), and of health outcomes involved (including low birth weight, complications of pregnancy, self-reported symptoms, blood pressure, arthritis, tuberculosis, depression, alcoholism, and mortality) suggested a robust, putatively causal, association. Cassel and Cobb indicated that social relationships might promote health in several ways, but emphasized the role of social relationships in moderating or buffering potentially deleterious health effects of psychosocial stress or other health hazards. This idea of “social support,” or something that maintains or sustains the organism by promoting adaptive behavior or neuroendocrine responses in the face of stress or other health hazards, provided a general, albeit simple, theory of how and why social relationships should causally affect health (16).

Publications on “social support” increased almost geometrically from 1976 to 1981. By the late 1970s, however, serious questions emerged about the empirical evidence cited by Cassel and Cobb and the evidence generated in subsequent research. Concerns were expressed about causal priorities between social support and health (since the great majority of studies remained cross-sectional or retrospective and based on self-reported data), about whether social relationships and supports buffered the impact of stress on health or had more direct effects, and about how consequential the effects of

social relationships on health really were (17–19). These concerns have been addressed by a continuing cumulation of two types of empirical data: (i) a new series of prospective mortality studies in human populations and (ii) a broadening base of laboratory and field experimental studies of animals and humans.

## Prospective Mortality Studies of Human Populations

Just as concerns began to surface about the nature and strength of the impact of social relationships on health, data from long-term, prospective studies of community populations provided compelling evidence that lack of social relationships constitutes a major risk factor for mortality. Berkman and Syme (20) analyzed a probability sample of 4775 adults in Alameda County, California, who were between 30 and 69 in 1965 when they completed a survey that assessed the presence or extent of four types of social ties—marriage, contacts with extended family and friends, church membership, and other formal and informal group affiliations. Each type of social relationship predicted mortality through the succeeding 9 years. A combined “social network” index remained a significant predictor of mortality (with a relative risk ratio for mortality of about 2.0, indicating that persons low on the index were twice as likely to die as persons high on the index) in multivariate analyses that controlled for self-reports in 1965 of physical health, socioeconomic status, smoking, alcohol consumption, physical activity, obesity, race, life satisfaction, and use of preventive health services. Such adjustment or control for baseline health and other risk factors provides a conservative estimate of the predictive power of social relationships, since some of their impact may be mediated through effects on these risk factors.

The major limitation of the Berkman and Syme study was the lack of other than self-reported data on health at baseline. Thus, House *et al.* (21) sought to replicate and extend the Alameda County results in a study of 2754 adults between 35 and 69 at their initial interview and physical examinations in 1967 through 1969 by the Tecumseh (Michigan) Community Health Study. Composite indices of social relationships and activities (as well as a number of the individual components) were inversely associated with mortality during the succeeding 10- to 12-year follow-up period, with relative risks of 2.0 to 3.0 for men and 1.5 to 2.0 for women, after adjustment for the effects of age and a wide range of biomedically assessed (blood pressure, cholesterol, respiratory function, and electrocardiograms) as well as self-reported risk factors of mortality. Analyzing data on 2059 adults in the Evans County (Georgia) Cardiovascular Epidemiologic Study, Schoenbach *et al.* (22) also found that a social network index similar to that of Berkman and Syme (20) predicted mortality for an 11- to 13-year follow-up period, after adjustment for age and baseline measures of biomedical as well as self-reported risk factors of mortality. The Evans County associations were somewhat weaker than those in Tecumseh and Alameda County, and as in Tecumseh were stronger for males than females.

Studies in Sweden and Finland have described similar results. Tibblin, Welin, and associates (23, 24) studied two cohorts of men born in 1913 and 1923, respectively, and living in 1973 in Gothenberg, Sweden's second largest city. After adjustments for age, baseline levels of systolic blood pressure, serum cholesterol, smoking habits, and perceived health status, mortality in both cohorts through 1982 was inversely related to the number of persons in the household and the men's level of social and outside home activities in 1973. Orth-Gomer *et al.* (25) analyzed the mortality experience through 1981 of a random sample of 17,433 Swedish adults aged 29 to 74 at the time of their 1976 or 1977



baseline interviews. Frequency of contact with family, friends, neighbors, and co-workers in 1976–77 was predictive of mortality through 1981, after adjustment for age, sex, education, employment status, immigrant status, physical exercise, and self-reports of chronic conditions. The effects were stronger among males than among females, and were somewhat nonlinear, with the greatest increase in mortality risk occurring in the most socially isolated third of the sample. In a prospective study of 13,301 adults in predominantly rural eastern Finland, Kaplan *et al.* (26) found a measure of “social connections” similar to those used in Alameda County, Tecumseh, and Evans County to be a significant predictor of male mortality from all causes during 5 years, again after adjustments for other biomedical and self-reported risk factors. Female mortality showed similar, but weaker and statistically nonsignificant, effects.

These studies manifest a consistent pattern of results, as shown in Figs. 1 and 2, which show age-adjusted mortality rates plotted for the five prospective studies from which we could extract parallel data. The report of the sixth study (25) is consistent with these trends. The relative risks (*RR*) in Figs. 1 and 2 are higher than those reported above because they are only adjusted for age. The levels of mortality in Figs. 1 and 2 vary greatly across studies depending on the follow-up period and composition of the population by age, race, and ethnicity, and geographic locale, but the patterns of prospective association between social integration (that is, the number and frequency of social relationships and contacts) and mortality are remarkably similar, with some variations by race, sex, and geographic locale.

Only the Evans County study reported data for blacks. The predictive association of social integration with mortality among Evans County black males is weaker than among white males in Evans County or elsewhere (Fig. 1), and the relative risk ratio for black females in Evans County, although greater than for Evans County white females, is smaller than the risk ratios for white females in all other studies (Fig. 2). More research on blacks and other minority populations is necessary to determine whether these differences are more generally characteristic of blacks compared to whites.

Modest differences emerge by sex and rural as opposed to urban locale. Results for men and women are strong, linear, and similar in the urban populations of Alameda County (that is, Oakland and environs) and Gothenberg, Sweden (only men were studied in Gothenberg). In the predominantly small-town and rural populations of Tecumseh, Evans County, and eastern Finland, however, two notable deviations from the urban results appear: (i) female risk ratios are consistently weaker than those for men in the same rural populations (Figs. 1 and 2), and (ii) the results for men in more rural populations, although rivaling those in urban populations in terms of risk ratios, assume a distinctly nonlinear, or threshold, form. That is, in Tecumseh, Evans County, and eastern Finland, mortality is clearly elevated among the most socially isolated, but declines only modestly, if at all, between moderate and high levels of social integration.

Explanation of these sex and urban-rural variations awaits research on broader regional or national populations in which the same measures are applied to males and females across the full rural-urban continuum. The current results may have both substantive and methodological explanations. Most of the studies reviewed here, as well as others (27–29), suggest that being married is more beneficial to health, and becoming widowed more detrimental, for men than for women. Women, however, seem to benefit as much or more than men from relationships with friends and relatives, which tend to run along same-sex lines (20, 30). On balance, men may benefit more from social relationships than women, especially in cross-gender relationships. Small communities may also provide a

broader context of social integration and support that benefits most people, except for a relatively small group of socially isolated males.

These results may, however, have methodological rather than substantive explanations. Measures of social relationships or integration used in the existing prospective studies may be less valid or have less variance in rural and small town environments, and for women, thus muting their relationship with mortality. For example, the data for women in Fig. 2 are similar to the data on men if we assume that women have higher quality relationships and hence that their true level of social integration is moderate even at low levels of quantity. The social context of small communities may similarly provide a moderate level of social integration for everyone except quite isolated males. Thus measures of frequency of social contact may be poorer indices of social integration for women and more rural populations than for men and urban dwellers.

Variations in the results in Figs. 1 and 2 should not, however, detract from the remarkable consistency of the overall finding that social relationships do predict mortality for men and women in a wide range of populations, even after adjustment for biomedical risk factors for mortality. Additional prospective studies have shown that social relationships are similarly predictive of all-cause and cardiovascular mortality in studies of people who are elderly (31–33) or have serious illnesses (34, 35).

## Experimental and Quasi-Experimental Research

The prospective mortality data are made more compelling by their congruence with growing evidence from experimental and clinical research on animals and humans that variations in exposure to social contacts produce psychological or physiological effects that could, if prolonged, produce serious morbidity and even mortality. Cassel (13) reviewed evidence that the presence of a familiar member of the same species could buffer the impact of experimentally induced stress on ulcers, hypertension, and neurosis in rats, mice, and goats, respectively; and the presence of familiar others has also been shown to reduce anxiety and physiological arousal (specifically secretion of free fatty acids) in humans in potentially stressful laboratory situations (36, 37). Clinical and laboratory data indicate that the presence of or physical contact with another person can modulate human cardiovascular activity and reactivity in general, and in stressful contexts such as intensive care units (38, pp. 122–141). Research also points to the operation of such processes across species. Affectionate petting by humans, or even their mere presence, can reduce the cardiovascular sequelae of stressful situations among dogs, cats, horses, and rabbits (38, pp. 163–180). Nerem *et al.* (39) found that human handling also reduced the arteriosclerotic impact of a high fat diet in rabbits. Recent interest in the potential health benefits of pets for humans, especially the isolated aged, is based on similar notions, although the evidence for such efforts is only suggestive (40).

Bovard (41) has proposed a psychophysiological theory to explain how social relationships and contacts can promote health and protect against disease. He reviews a wide range of human and animal studies suggesting that social relationships and contacts, mediated through the amygdala, activate the anterior hypothalamic zone (stimulating release of human growth hormone) and inhibit the posterior hypothalamic zone (and hence secretion of adrenocorticotropic hormone, cortisol, catecholamines, and associated sympathetic autonomic activity). These mechanisms are consistent with the impact of social relationships on mortality from a wide range of causes and with studies of the adverse effects of lack of adequate social relationships on the development of human and animal infants (42). This theory is also consistent with sociobiological processes

which, due to the survival benefit of social relationships and collective activity, would promote genetic selection of organisms who find social contact and relatedness rewarding and the lack of such contact and relatedness aversive (43).

The epidemiologic evidence linking social relationships and supports to morbidity in humans is limited and not fully consistent. For example, although laboratory studies show short-term effects of social relationships on cardiovascular functioning that would, over time, produce cardiovascular disease, and prospective studies show impacts of social relationships on mortality from cardiovascular disease, the link between social relationships and the incidence of cardiovascular morbidity has yet to be firmly demonstrated (19, 44). Overall, however, the theory and evidence for the impact of social relationships on health are building steadily (45, 46).

## Social Relationships as a Risk Factor for Health: Research and Policy Issues

The theory and data reviewed above meet reasonable criteria for considering social relationships a cause or risk factor of mortality, and probably morbidity, from a wide range of diseases (10; 46; 47, pp. 289–321). These criteria include strength and consistency of statistical associations across a wide range of studies, temporal ordering or prediction from cause to effect, a gradient of response (which may in this case be nonlinear), experimental data on animals and humans consistent with nonexperimental human data, and a plausible theory (41) of biopsychosocial mechanisms explaining the observed associations.

The evidence on social relationships is probably stronger, especially in terms of prospective studies, than the evidence which led to the certification of the Type A behavior pattern as a risk factor for coronary heart disease (48). The evidence regarding social relationships and health increasingly approximates the evidence in the 1964 Surgeon General's report (10) that established cigarette smoking as a cause or risk factor for mortality and morbidity from a range of diseases. The age-adjusted relative risk ratios shown in Figs. 1 and 2 are stronger than the relative risks for all cause mortality reported for cigarette smoking (10). There is, however, less specificity in the associations of social relationships with mortality than has been observed for smoking, which is strongly linked to cancers of the lung and respiratory tract (with age-adjusted risk ratios between 3.0 and 11.0). Better theory and data are needed on the links between social relationships and major specific causes of morbidity and mortality.

Although a lack of social relationships has been established as a risk factor for mortality, and probably morbidity, three areas need further investigation: (i) mechanisms and processes linking social relationships to health, (ii) determinants of levels of "exposure" to social relationships, and (iii) the means to lower the prevalence of relative social isolation in the population or to lessen its deleterious effects on health.

## Mechanisms and Processes Linking Social Relationships to Health

Although grounded in the literature on social relationships and health, investigators on social support in the last decade leaped almost immediately to the interpretation that what was consequential for health about social relationships was their supportive quality, especially their capacity to buffer or moderate the deleterious effects of stress or other health hazards (13, 14). Many recent studies have reported either a general positive association between social support

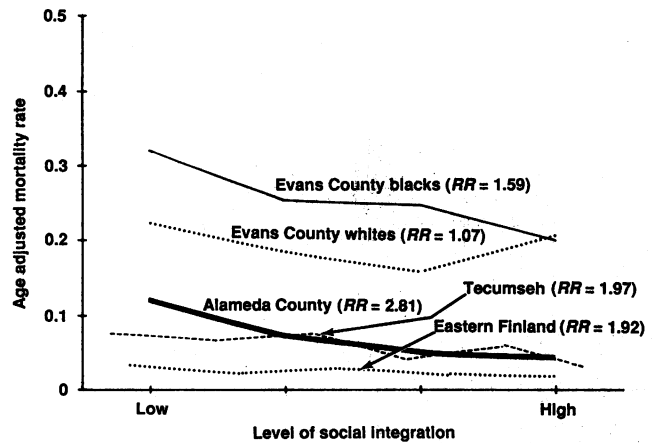


Fig. 2. Level of social integration and age-adjusted mortality for females in five prospective studies. RR, the relative risk ratio of mortality at the lowest versus highest level of social integration.

and health or a buffering effect in the presence of stress (49), but these studies are problematic because the designs are largely cross-sectional or retrospective and the data usually self-reported. The most compelling evidence of the causal significance of social relationships on health has come from the experimental studies of animals and humans and the prospective mortality studies reviewed above—studies in which the measures of social relationships are merely the presence or absence of familiar other organisms, or relative frequency of contact with them, and which often do not distinguish between buffering and main effects. Thus, social relationships appear to have generally beneficial effects on health, not solely or even primarily attributable to their buffering effects, and there may be aspects of social relationships other than their supportive quality that account for these effects.

We now need a broader theory of the biopsychosocial mechanisms and processes linking social relationships to health than can be provided by extant concepts or theories of social support. That broader theory must do several things. First, it must clearly distinguish between (i) the existence or quantity of social relationships, (ii) their formal structure (such as their density or reciprocity), and (iii) the actual content of these relationships such as social support. Only by testing the effects on health of these different aspects of social relationships in the same study can we understand what it is about social relationships that is consequential for health.

Second, we need better understanding of the social, psychological, and biological processes that link the existence, quantity, structure, or content of social relationships to health. Social support—whether in the form of practical help, emotional sustenance, or provision of information—is only one of the social processes involved here. Not only may social relationships affect health because they are or are not supportive, they may also regulate or control human thought, feeling and behavior in ways that promote health, as in Durkheim's (2) theory relating social integration to suicide. Current views based on this perspective suggest that social relationships affect health either by fostering a sense of meaning or coherence that promotes health (50) or by facilitating health-promoting behaviors such as proper sleep, diet, or exercise, appropriate use of alcohol, cigarettes, and drugs, adherence to medical regimens, or seeking appropriate medical care (51). The negative or conflictive aspects of social relationships need also to be considered, since they may be detrimental to the maintenance of health and of social relationships (52).

We must further understand the psychological and biological processes or mechanisms linking social relationships to health, either

as extensions of the social processes just discussed [for example, processes of cognitive appraisal and coping (53)] or as independent mechanisms. In the latter regard, psychological and sociobiological theories suggest that the mere presence of, or sense of relatedness with, another organism may have relatively direct motivational, emotional, or neuroendocrinal effects that promote health either directly or in the face of stress or other health hazards but that operate independently of cognitive appraisal or behavioral coping and adaptation (38, pp. 87–180; 42, 43, 54).

## Determinants of Social Relationships: Scientific and Policy Issues

Although social relationships have been extensively studied during the past decade as independent, intervening, and moderating variables affecting stress or health or the relations between them, almost no attention has been paid to social relationships as dependent variables. The determinants of social relationships, as well as their consequences, are crucial to the theoretical and causal status of social relationships in relation to health. If exogenous biological, psychological, or social variables determine both health and the nature of social relationships, then the observed association of social relationships to health may be totally or partially spurious. More practically, Cassel (13), Cobb (14), and others became interested in social support as a means of improving health. This, in turn, requires understanding of the broader social, as well as psychological or biological, structures and processes that determine the quantity and quality of social relationships and support in society.

It is clear that biology and personality must and do affect both people's health and the quantity and quality of their social relationships. Research has established that such factors do not, however, explain away the experimental, cross-sectional, and prospective evidence linking social relationships to health (55). In none of the prospective studies have controls for biological or health variables been able to explain away the predictive association between social relationships and mortality. Efforts to explain away the association of social relationships and supports with health by controls for personality variables have similarly failed (56, 57). Social relationships have a predictive, arguably causal, association with health in their own right.

The extent and quality of social relationships experienced by individuals is also a function of broader social forces. Whether people are employed, married, attend church, belong to organizations, or have frequent contact with friends and relatives, and the nature and quality of those relationships, are all determined in part by their positions in a larger social structure that is stratified by age, race, sex, and socioeconomic status and is organized in terms of residential communities, work organizations, and larger political and economic structures. Older people, blacks, and the poor are generally less socially integrated (58), and differences in social relationships by sex and place of residence have been discussed in relation to Figs. 1 and 2. Changing patterns of fertility, mortality, and migration in society affect opportunities for work, marriage, living and working in different settings, and having relationships with friends and relatives, and can even affect the nature and quality of these relations (59). These demographic patterns are themselves subject to influence by both planned and unplanned economic and political change, which can also affect individuals' social relationships more directly—witness the massive increase in divorce during the last few decades in response to the women's movement, growth in women's labor force participation, and changing divorce law (60, 61).

In contrast with the 1950s, adults in the United States in the

1970s were less likely to be married, more likely to be living alone, less likely to belong to voluntary organizations, and less likely to visit informally with others (62). Changes in marital and childbearing patterns and in the age structure of our society will produce in the 21st century a steady increase of the number of older people who lack spouses or children—the people to whom older people most often turn for relatedness and support (59). Thus, just as we discover the importance of social relationships for health, and see an increasing need for them, their prevalence and availability may be declining. Changes in other risk factors (for example, the decline of smoking) and improvements in medical technology are still producing overall improvements on health and longevity, but the improvements might be even greater if the quantity and quality of social relationships were also improving.

### REFERENCES AND NOTES

1. C. Darwin, *Expression of the Emotions in Man and Animals* (Univ. of Chicago Press, Chicago, 1965 [1872]).
2. E. Durkheim, *Suicide* (Free Press, New York, 1951 [1897]).
3. A. S. Kraus and A. N. Lilienfeld, *J. Chronic Dis.* **10**, 207 (1959).
4. H. Carter and P. C. Glick, *Marriage and Divorce: A Social and Economic Study* (Harvard Univ. Press, Cambridge, MA, 1970).
5. E. M. Kitigawa and P. M. Hauser, *Differential Mortality in the United States: A Study in Socio-Economic Epidemiology* (Harvard Univ. Press, Cambridge, MA, 1973).
6. T. H. Holmes, in *Personality, Stress and Tuberculosis*, P. J. Sparer, Ed. (International Univ. Press, New York, 1956).
7. W. A. Tillman and G. E. Hobbs, *Am. J. Psychiatr.* **106**, 321 (1949).
8. R. E. L. Faris, *Am. J. Sociol.* **39**, 155 (1934).
9. M. L. Kohn and J. A. Clausen, *Am. Sociol. Rev.* **20**, 268 (1955).
10. U.S. Surgeon General's Advisory Committee on Smoking and Health, *Smoking and Health* (U.S. Public Health Service, Washington, DC, 1964).
11. G. Caplan, *Support Systems and Community Mental Health* (Behavioral Publications, New York, 1974).
12. President's Commission on Mental Health, *Report to the President* (Government Printing Office, Washington, DC, 1978), vols. 1 to 5.
13. J. Cassel, *Am. J. Epidemiol.* **104**, 107 (1976).
14. S. Cobb, *Psychosomatic Med.* **38**, 300 (1976).
15. J. Cassel, in *Social Stress*, S. Levine and N. A. Scotch, Eds. (Aldine, Chicago, 1970), pp. 189–209.
16. J. S. House, *Work Stress and Social Support* (Addison-Wesley, Reading, MA, 1981).
17. K. Heller, in *Maximizing Treatment Gains: Transfer Enhancement in Psychotherapy*, A. P. Goldstein and F. H. Kanter, Eds. (Academic Press, New York, 1979), pp. 353–382.
18. P. A. Thoits, *J. Health Soc. Behav.* **23**, 145 (1982).
19. D. Reed *et al.*, *Am. J. Epidemiol.* **117**, 384 (1983).
20. L. F. Berkman and S. L. Syme, *ibid.* **109**, 186 (1979).
21. J. S. House, C. Robbins, H. M. Metzner, *ibid.* **116**, 123 (1982).
22. V. J. Schoenbach *et al.*, *ibid.* **123**, 577 (1986).
23. G. Tibblin *et al.*, in *Social Support: Health and Disease*, S. O. Isacsson and L. Janzon, Eds. (Almqvist & Wiksell, Stockholm, 1986), pp. 11–19.
24. L. Welin *et al.*, *Lancet* **i**, 915 (1985).
25. K. Orth-Gomer and J. Johnson, *J. Chron. Dis.* **40**, 949 (1987).
26. G. A. Kaplan *et al.*, *Am. J. Epidemiol.*, in press.
27. M. Stroebe and W. Stroebe, *Psychol. Bull.* **93**, 279 (1983).
28. W. R. Gove, *Soc. Forces* **51**, 34 (1972).
29. K. J. Helsing and M. Szklo, *Am. J. Epidemiol.* **114**, 41 (1981).
30. L. Wheeler, H. Reis, J. Nezlek, *J. Pers. Soc. Psychol.* **45**, 943 (1983).
31. D. Blazer, *Am. J. Epidemiol.* **115**, 684 (1982).
32. D. M. Zuckerman, S. V. Kasl, A. M. Ostfeld, *ibid.* **119**, 410 (1984).
33. T. E. Seeman *et al.*, *ibid.* **126**, 714 (1987).
34. W. E. Ruberman *et al.*, *N. Engl. J. Med.* **311**, 552 (1984).
35. K. Orth-Gomer *et al.*, in *Social Support: Health and Disease*, S. O. Isacsson and L. Janzon, Eds. (Almqvist & Wiksell, Stockholm, 1986), pp. 21–31.
36. L. S. Wrightsman, Jr., *J. Abnorm. Soc. Psychol.* **61**, 216 (1960).
37. K. W. Back and M. D. Bogdonoff, *Behav. Sci.* **12**, 384 (1967).
38. J. J. Lynch, *The Broken Heart* (Basic Books, New York, 1979).
39. R. M. Nerem, M. J. Levesque, J. F. Cornhill, *Science* **208**, 1475 (1980).
40. J. Goldmeier, *Gerontologist* **26**, 203 (1986).
41. E. W. Boward, in *Perspectives on Behavioral Medicine*, R. B. Williams (Academic Press, New York, 1985), vol. 2.
42. J. Bowlby, in *Loneliness: The Experience of Emotional and Social Isolation*, R. S. Weiss, Ed. (MIT Press, Cambridge, MA, 1973).
43. S. P. Mendoza, in *Social Cohesion: Essays Toward a Sociophysiological Perspective*, P. R. Barchas and S. P. Mendoza, Eds. (Greenwood Press, Westport, CT, 1984).
44. S. Cohen, *Health Psychol.* **7**, 269 (1988).
45. L. F. Berkman, in *Social Support and Health*, S. Cohen and S. L. Syme, Eds. (Academic Press, New York, 1985), pp. 241–262.
46. W. E. Broadhead *et al.*, *Am. J. Epidemiol.* **117**, 521 (1983).
47. A. M. Lilienfeld and D. E. Lilienfeld, *Foundations of Epidemiology* (Oxford Univ. Press, New York, 1980).
48. National Heart, Lung and Blood Institute, *Circulation* **63**, 1199 (1982).
49. S. Cohen and S. L. Syme, *Social Support and Health* (Academic Press, New York,

- 1985).
50. A. Antonovsky, *Health, Stress and Coping* (Jossey-Bass, San Francisco, 1979).
  51. D. Umberson, *J. Health Soc. Behav.* **28**, 306 (1987).
  52. K. Rook, *J. Pers. Soc. Psychol.* **46**, 1097 (1984).
  53. R. S. Lazarus and S. Folkman, *Stress, Appraisal, and Coping* (Springer, New York, 1984).
  54. R. B. Zajonc, *Science* **149**, 269 (1965).
  55. J. S. House, D. Umberson, K. Landis, *Annu. Rev. Sociol.*, in press.
  56. S. Cohen, D. R. Sherrod, M. S. Clark, *J. Pers. Soc. Psychol.* **50**, 963 (1986).
  57. R. Schultz and S. Decker, *ibid.* **48**, 1162 (1985).
  58. J. S. House, *Socio. Forum* **2**, 135 (1987).
  59. S. C. Watkins, J. A. Menken, J. Bongaarts, *Am. Sociol. Rev.* **52**, 346 (1987).
  60. A. Cherlin, *Marriage, Divorce, Remarriage* (Harvard Univ. Press, Cambridge, MA, 1981).
  61. L. J. Weitzman, *The Divorce Revolution* (Free Press, New York, 1985).
  62. J. Veroff, E. Douvan, R. A. Kulka, *The Inner American: A Self-Portrait from 1957 to 1976* (Basic Books, New York, 1981).
  63. Supported by a John Simon Guggenheim Memorial Foundation Fellowship and NIA grant 1-PO1-AG05561 (to J.S.H.), NIMH training grant 5-T32-MH16806-06 traineeship (to K.R.L.), NIMH training grant 5-T32-MH16806-05 and NIA 1-F32-AG05440-01 postdoctoral fellowships (to D.U.). We are indebted to D. Buss, P. Converse, G. Duncan, R. Kahn, R. Kessler, H. Schuman, L. Syme, and R. Zajonc for comments on previous drafts, to many other colleagues who have contributed to this field, and to M. Klatt for preparing the manuscript.

# The Coupling of Neurotransmitter Receptors to Ion Channels in the Brain

ROGER A. NICOLL

**Recent studies on the action of neurotransmitters on hippocampal pyramidal cells indicate that different neurotransmitter receptors that use either the same or different coupling mechanisms converge onto the same ion channel. Conversely, virtually all of the neurotransmitters act on at least two distinct receptor subtypes coupled to different ion channels on the same cell. The existence of both convergence and divergence in the action of neurotransmitters results in a remarkable diversity in neuronal signaling.**

**I**N THE 1950S THE USE OF INTRACELLULAR RECORDING ESTABLISHED the chemical nature of synaptic transmission in the mammalian central nervous system (1). Specifically, neurotransmitters produce their excitatory and inhibitory effects on neurons by rapidly and briefly increasing membrane ion permeability. Over the past 25 years, the use of iontophoretic techniques, in which drugs can be directly applied onto single neurons *in vivo*, allowed such compounds as  $\gamma$ -aminobutyric acid (GABA), glycine, glutamate, norepinephrine, acetylcholine (ACh), and serotonin to be identified as neurotransmitters in the mammalian brain (2). However, investigators had to turn to simpler invertebrate preparations for more detailed studies of the mechanisms involved in the action of neurotransmitters. These preparations, as well as preparations of peripheral ganglia from vertebrates, avoided the numerous constraints of *in vivo* mammalian preparations, such as the complexity of the neuronal circuitry, barriers to drug delivery, and lack of stability for intracellular recording. Investigators were thus able to demonstrate an extraordinary diversity in neurotransmitter action (3). The results to be discussed in this review will demonstrate that studies that could originally only be done in these invertebrate *in vitro* preparations can now be performed in *in vitro* slice preparations of the mammalian brain. The slice preparation has the advantage that fully differentiated adult brain tissue, in which the local neuronal circuitry is often intact and functioning, is used.

Although biochemical studies on brain slices have been carried out for over half a century, electrical recordings were first made in

the mid-1960s by Yamamoto (4). The development of the transverse hippocampal slice in which all of the local circuits can be functionally preserved (5), coupled with the demonstration that high quality intracellular recording could routinely be obtained from these slices (6), led to a rapid explosion in the use of brain slice preparations for electrophysiological analysis (7).

Previous results by neurochemists suggested a functional role for a variety of putative neurotransmitters. For instance, in the hippocampus, neurochemical and immunohistochemical studies have found close to 20 different putative neurotransmitters in nerve cell bodies or fibers or both. Specific receptor binding sites for many of these same compounds have been identified and characterized (8). In addition, the brain has long provided the richest source of tissue for biochemists studying various components of second messenger systems (9), which mediate the effects of many neurotransmitters.

I have focused my attention on the hippocampal slice preparation for a number of reasons. There are many neurotransmitter candidates in this cortical structure (Table 1). In addition, the precise laminar organization of cellular components permits much of the local circuitry to remain intact and greatly facilitates intracellular recording and selective stimulation of afferent neuronal pathways. Finally, the hippocampus is of considerable interest in its own right, because it plays a key role in memory (10) and epilepsy (11).

Neurotransmitters exert their effects by opening or closing ion channels in the neuronal membrane, either directly or through the activation of intermediate proteins. Until recently, it was believed that most neurotransmitters produced a simple hyperpolarization or depolarization of brief duration in the postsynaptic membrane. However, it is now known that neurotransmitters can alter, over a longer time span, the properties of voltage-dependent ionic conductances that are involved in the control of cell excitability (3). In particular, a wide range of  $K^+$  currents and  $Ca^{2+}$  currents are modulated by neurotransmitters, resulting in such changes as increased action potential duration, changes in firing frequency and firing pattern, and increased  $Ca^{2+}$  entry during an action potential. Thus, the modulation of voltage-dependent ion channels allows for a finer control via multiple mechanisms of various aspects of cell excitability.

The author is a professor in the Departments of Pharmacology and Physiology, School of Medicine, University of California, San Francisco, CA 94143.



## Review

# The neural circuitry of social homeostasis: Consequences of acute versus chronic social isolation

Christopher R. Lee,<sup>1,2</sup> Alon Chen,<sup>3</sup> and Kay M. Tye<sup>1,\*</sup>

<sup>1</sup>The Salk Institute for Biological Studies, La Jolla, CA 92037, USA

<sup>2</sup>Neurosciences Graduate Program, University of California, San Diego, La Jolla, CA 92093, USA

<sup>3</sup>Department of Neurobiology, Weizmann Institute of Science, Rehovot 76100, Israel

\*Correspondence: [tye@salk.edu](mailto:tye@salk.edu)

<https://doi.org/10.1016/j.cell.2021.02.028>

## SUMMARY

Social homeostasis is the ability of individuals to detect the quantity and quality of social contact, compare it to an established set-point in a command center, and adjust the effort expended to seek the optimal social contact expressed via an effector system. Social contact becomes a positive or negative valence stimulus when it is deficient or in excess, respectively. Chronic deficits lead to set-point adaptations such that reintroduction to the previous optimum is experienced as a surplus. Here, we build upon previous models for social homeostasis to include adaptations to lasting changes in environmental conditions, such as with chronic isolation.

## INTRODUCTION

How do we adapt when opportunities for social engagement change in a long-lasting manner? Social homeostasis is an adaptive function to regulate behaviors that govern social connection to an optimal level, to avoid a surplus of social contact (wherein competition for resources, space, and mates becomes too fierce) as well as a deficit of social contact (wherein safety, warmth, observational learning, and play are lacking) (Matthews and Tye, 2019).

Drawing from the conceptual framework established by Abraham Maslow for a hierarchy of needs (Maslow, 1943a, 1943b), there are two classes of needs: “being needs” or “self-actualization” and “deficiency needs,” which are reminiscent of homeostatic need states wherein the motivational drive to meet the needs increases when the needs are continually unmet. Even within “deficiency needs,” Maslow conceptualized several tiers: (1) *physiological needs* that are essential for survival (air, food, water, shelter, clothing, sleep, sex); (2) *safety needs* that allow an individual a sense of order, predictability, and control (family, police, schools, medical care, legal system, governance structure); (3) *love and belongingness needs* (love, intimacy, friendship, acceptance); and (4) *self-esteem needs* that are subcategorized into esteem for oneself and esteem from others (dignity, status, prestige). Importantly, social needs are actually described in all 4 of these tiers. Within the social dimension, Maslow framed sexual contact as the most basic or essential, with a social structure to provide safety, protection, and stability on the next tier, intimate relationships on the next, and finally social rank and influence in the fourth tier. However, Maslow’s evolving framework assumes modern human society, rather than the environments in which our brains have evolved.

A deficit or surplus of social contact can be interpreted as a type of stressor that induces efforts to correct the deviation from the optimum that may trigger both adaptive and pathological responses. A social surplus such as overcrowding may induce psychological stress, social conflict, and aggression (Loo and Ong, 1984), with robustly consistent results in animal studies (for review, see Christian, 1961, 1970). However, human studies on the effects of density and crowding on social behavior have produced inconsistent results, perhaps due to the contribution of variables specific to humans (self-report, social programming, cultural norms) as well as individual variability in social status, gender, culture, and context (Evans et al., 1998; Regoeczi, 2003).

A social deficit can occur with objective or subjective social isolation, social exclusion, or subordination. Social isolation is universally aversive to social species, with behavioral and neuroendocrine changes associated with mental and physical health consequences including depression, shortened lifespan, and increased rates of cancer (Hermes et al., 2009; Ma et al., 2011; Steptoe et al., 2013; Weiss, 1973). Chronic social isolation, particularly during rearing, is well known to be used as a rodent model for schizophrenia (Geyer et al., 1993; Kohn and Clausen, 1955).

In humans, perceived deficits in the objective quantity, or subjective quality, of social contact (“loneliness” [Weiss, 1973]), are correlated with deficits in mental (Cacioppo et al., 2006a) and physical (Hawkley and Cacioppo, 2010; Hawkley et al., 2006) health and higher mortality rates (Berkman and Syme, 1979; Holt-Lunstad et al., 2010; Holwerda et al., 2012; Perissinotto et al., 2012; Steptoe et al., 2013). Perceived loneliness correlates with increased morbidity and mortality with cancer and cardiovascular disease (Hawkley and Cacioppo, 2003). Perceived

loneliness is also correlated with the symptom severity in response to viral immune challenges (LeRoy et al., 2017) and inflammatory responses (Balter et al., 2019).

Solitary confinement has been deemed a form of torture (Hreško, 2006; Thoenig, 1972). Yet, the negative valence serves an adaptive purpose—representing a homeostatic need state. We hypothesize that the unpleasant state of loneliness reflects evolutionarily advantageous neuroadaptations that increase motivation to seek social contact (Matthews and Tye, 2019).

With acute social isolation, animals from rodents to humans perform prosocial behaviors (a rebound of social interaction, increased affiliative behaviors) following social isolation for 24 h (Matthews and Tye, 2019; Matthews et al., 2016; Panksepp and Beatty, 1980; Tomova et al., 2019, 2020a). However, with chronic social isolation, flies, rodents, and humans display antisocial behaviors (aggression, avoidance, social anxiety) (Ma et al., 2011; Sciolino et al., 2010; Zelikowsky et al., 2018) that may be manifested in the form of mental health disorders in our modern-day society.

A number of studies have demonstrated that individual differences both across and within species with respect to patterns of social engagement are predetermined by genetic factors (Forkosh et al., 2019; Hoekstra and Coyne, 2007; Lim et al., 2004; Wang et al., 2008); yet, relatively little is known about the neural circuit mechanisms underlying how a given individual adapts and responds to a changing social environment.

With the unprecedented global pandemic of 2020 that induced a government-mandated lockdown in many cities and countries, as well as prolonged “social-distancing,” the urgent need to quarantine dramatically changed the social landscape of almost every human being on the planet. We were pointedly faced with the competing needs of slowing the spread of coronavirus disease 2019 (COVID-19) and maintaining our social needs. Indeed, the uncertainty in how to evaluate these two competing needs has been a challenge that has sparked controversy, has been heavily politicized, and has manifested a wide range of individualized assessments.

Here, we present a conceptual framework toward understanding the dynamic responses and adaptations made in the face of changes in an individual’s social environment. This conceptual framework provides a mechanistic explanation for how the valence of a social stimulus (whether it is evaluated as positive or negative) can change to the same stimulus (reintroduction to a social group) from acute isolation to chronic isolation. We propose that the emotional valence (Tye, 2018) of a social stimulus can be modulated by the internal need state of an individual. Overcrowding changes the perceived valence of social contact, and an isolation-induced reduction in the social homeostatic set-point may make a social stimulus previously perceived as optimal now be interpreted as a surplus. This may explain why chronic social isolation can produce antisocial behaviors, whereas acute social isolation produces prosocial behaviors.

## COMPONENTS OF HOMEOSTATIC SYSTEMS

Homeostasis refers to the physiological processes wherein stable states are maintained through compensatory mechanisms and is schematized with receptors, a control center, and effec-

tors. Homeostatic systems are known to exist for a number of physiological needs (Cannon, 1929). For example, in thermoregulation, when an imbalance is detected by temperature-sensing receptors in the skin and brain, signals are sent to a control center which coordinates a response, and signals will be sent to effector systems such as sweat glands to promote cooling in the case of overheating to maintain homeostasis (Tan et al., 2016; Wendt et al., 2007). Motivated behaviors associated with aversive “drive” states are governed by the physiological needs of the individual (hunger or thirst) when in a deficit (Bai et al., 2019; Betley et al., 2013; Oka et al., 2015; Zimmerman et al., 2016). The neural circuit bases of many homeostatic systems (i.e., thermoregulation, osmoregulation, sleep homeostasis, and energy balance) are supported by a rich literature, yet social homeostasis remains uncharted territory ripe for investigation.

## NEURAL COMPONENTS OF SOCIAL HOMEOSTASIS

Social homeostasis refers to the ability of an individual to detect the quantity and quality of social contact, compare it to an established *set-point* in a *command center*, and adjust the effort expended to seek social contact through an *effector* system (Figure 1; Table 1; Matthews and Tye, 2019).

### Detector

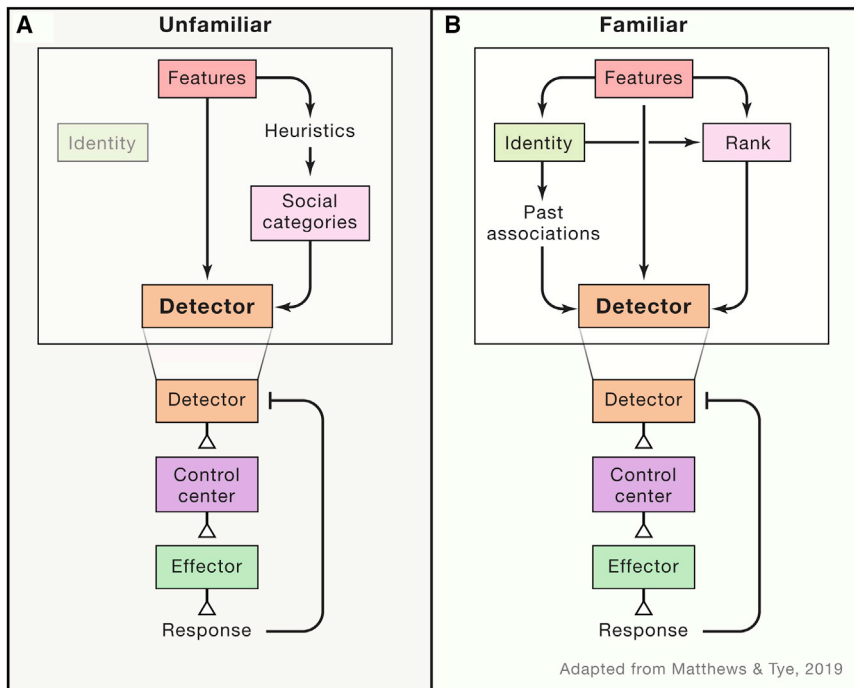
The first step within a social homeostatic system is detecting the social environment. The detector node is responsible for sensing deviations in the *quality* and *quantity* of social contact (Matthews and Tye, 2019). While the quantity of social interaction is an objective measure, the quality of a social interaction is subjective and is highly dependent on the context of the social interaction—factoring in relative dominance, social hierarchical structure, relationship history, etc. (Shemesh et al., 2013, 2016). We hypothesize that, especially in unfamiliar individuals, the information about the characteristics of the *social stimulus* (i.e., sex, age, etc.) is weighted more heavily in the assessment of a social interaction and acts as a heuristic (Figure 1A). With familiarization, the learned identity of the individual carries more weight in determining the quality of a specific interaction (Figure 1B).

We conceptualize the overall detected social interaction—which we operationally define as *social utility*—as the product of both social quantity and quality, where the preferred *quality* of social interactions monotonically increases while the preferred *quantity* increases to an optimal point and then declines when there is a surplus (Figure 2). For example, if the quality of social contact is very low, even if there is the optimal or very high quantity the social utility remains low. Additionally, if the quality of social contact is very high, but the quantity is low, the social utility remains low. If the quantity is very high, even if the quality is high, there is a surplus that reduces the social utility. However, if the quality is high and the quantity is optimal, this offers the maximum social utility.

In the following sections, we examine the factors, as well as the underlying neural circuits and mechanisms, that feed into a detector system.

### Features and agent characteristics

An individual’s social environment is often diverse and complex, with features such as age and sex of a conspecific determining



**Figure 1. Integration of features, rank, and identity at the level of social detection**

(A and B) Schematic of social information detection from (A) unfamiliar individuals and (B) familiar individuals. The first step of a social homeostatic system is detection, which integrates many social variables such as social features, rank, and identity to determine the overall quality of a social interaction. (A) When an individual first interacts with an unfamiliar conspecific, the individual primarily relies on the social features of the unfamiliar individual, such as age, sex, etc. These social features provide a heuristic to assess a social agent and determine an appropriate behavioral response. (B) When interacting with a familiar individual, information from social features and the learned identity of the other individual feed into the detector node as well as provide information on the relative rank of the individual. The detector integrates all of this information in evaluating social interactions, which is then fed forward to subsequent nodes in the social homeostatic system.

the type of social interaction and the behavioral output of that individual (for review, see [Chen and Hong, 2018](#)). For example, an individual's behavior will likely be different for different social stimuli (e.g., a newborn baby, attractive individual of the same age, or an elderly relative). Especially for unfamiliar social stimuli, individuals may rely on sensory cues—such as olfactory, auditory, or visual—as heuristics in determining an appropriate behavioral output in response to a social stimulus (i.e., more affiliative, supportive behavior toward the newborn, and more confident behavior toward the attractive individual). These complex social behaviors and decisions extend to other mammals. For instance, rats will immediately liberate cagemates who are trapped in restrainers, perhaps a demonstration of empathy ([Ben-Ami Bartal et al., 2011](#)), and rats are more inclined to help rats that are of the same strain as rats with whom the subject is most familiar ([Ben-Ami Bartal et al., 2014](#)), suggesting that rats rely on feature-based heuristics. Additionally, for observational learning, socially derived information about the environment requires the anterior cingulate cortex (ACC)-basolateral amygdala (BLA) circuit ([Allsop et al., 2018](#)).

Pheromonal detection occurs in the vomeronasal organ (VNO) and to a lesser extent in the main olfactory epithelium and ultimately leads to innate behaviors owing to hardwired circuitry (for review, see [Stowers and Kuo, 2015](#)). For example, chemosensory detection of the major urinary proteins (MUPs) MUP3 and MUP20 in the VNO of male mice induces intermale aggressive behaviors through projections to the accessory olfactory bulb (AOB), which activates hardwired neural circuitry driving aggression (such as the “core aggression circuit,” including the medial amygdala [MeA] and the ventrolateral part of the ventromedial hypothalamus [VMHvl], as described in [Lischinsky and Lin, 2020](#); [Hashikawa et al., 2016](#)). By contrast, detection of

MUP20 in the VNO of female mice induces attraction toward males ([Kaur et al., 2014](#); [Roberts et al., 2010](#)). Meanwhile, sulfated estrogens found in female mouse urine result in male attraction when sensed in the male VNO ([Haga-Yamanaka et al., 2014](#)). Additionally, exocrine-gland-secreting peptide 22 (ESP22), a pheromone secreted in the tears of juvenile (2–3 week old) mice, strongly inhibits sexual behavior when detected in the VNO of male mice, whereas abolishing ESP22 in juveniles increases sexual behavior of adult males with juveniles, suggesting that this pheromone is critical for mice to discriminate adults from juveniles ([Ferrero et al., 2013](#)). Removal of the VNO in mice, and specifically, removal of transient receptor potential cation channel, subfamily C, member 2 (TRPC2)—a protein integral for chemosensory transduction in the VNO—results in social behavior deficits, including a decrease in intermale aggression, dysfunction in discriminating males and females, and reduction in sexual behavior between males and females ([Clancy et al., 1984](#); [Leypold et al., 2002](#); [Stowers et al., 2002](#)). Interestingly, oxytocin (Oxt) signaling in the MeA (downstream of VNO) is necessary for sex discrimination in mice, suggesting further downstream processing is required in the orchestration of detecting agent characteristics ([Yao et al., 2017](#)). Altogether, chemosensory cues are necessary for detecting the type of social stimulus presented and determining the appropriate response to that social agent; in a social homeostatic system, detection of social features is critical for determining the quality of a social interaction.

In addition to chemosensory cues, auditory cues offer another sensory modality to social individuals to determine the characteristics of another social agent. Marmosets, for instance, use phee calls to re-unite isolated animals back to their social group ([Eliades and Miller, 2017](#); [Moynihan, 1970](#)). Phee calls in male marmosets are markedly different from female phee calls in that they tend to be higher frequency and have greater variability ([Norcross and Newman, 1993](#)), providing information about the

**Table 1. Glossary of terms in the context of social behavior**

Term	Definition
Social homeostasis	The ability of individuals to regulate the quantity and quality of social contact and maintain stability within a social structure.
Detector	A neural system that senses changes in the quantity of social interactions and quality of an individual's social environment, integrating factors such as social agent characteristics, relative rank, and identity.
Control center	A neural system that compares deviations in social utility to an encoded homeostatic set-point to calculate deficits or surpluses of social interaction.
Effector	A neural system that coordinates motivated behavior to resolve deficits and surpluses in social utility (e.g., prosocial affiliative behavior or antisocial aggressive behavior).
Social utility	The product of the detected social quantity and quality. The preferred <i>quality</i> of social interactions monotonically increases while the preferred <i>quantity</i> increases to an optimal point and then declines when there is a surplus.
Homeostatic set-point	An individual's ideal level of social utility.
Despotic Hierarchy	Social structure in which one animal is dominant over all other individuals, and there is a large resource/power disparity between the dominant and others.
Linear Hierarchy	Social structure in which each individual has a ranking, and the linear rankings of the "pecking order" obey the laws of transitivity
Egalitarian Hierarchy	Social structure in which resources and power are distributed equally amongst individuals.
Prosocial (affiliative) behaviors	Social behaviors that promote group cohesion (friendly/positive gestures), e.g., grooming, touching, hugging.
Antisocial behaviors	Social behaviors that hinder group cohesion, e.g., aggression, intimidation, fighting.
Loneliness	State of distress or discomfort that results when one perceives a gap between one's desires for social connection and actual experiences of it.
Valence	Positive or negative motivational significance.
Acute versus chronic	For the purpose of this Review, acute and chronic social isolation refer to relative timescales, as species with different lifespans, reproductive cycles, metabolisms, etc., will likely also have differing thresholds for acute versus chronic social isolation.

individual's sex. Indeed, social context influences the cortical activity in marmosets, as the state of frontal cortex neurons before a vocalization was heard predicted whether the

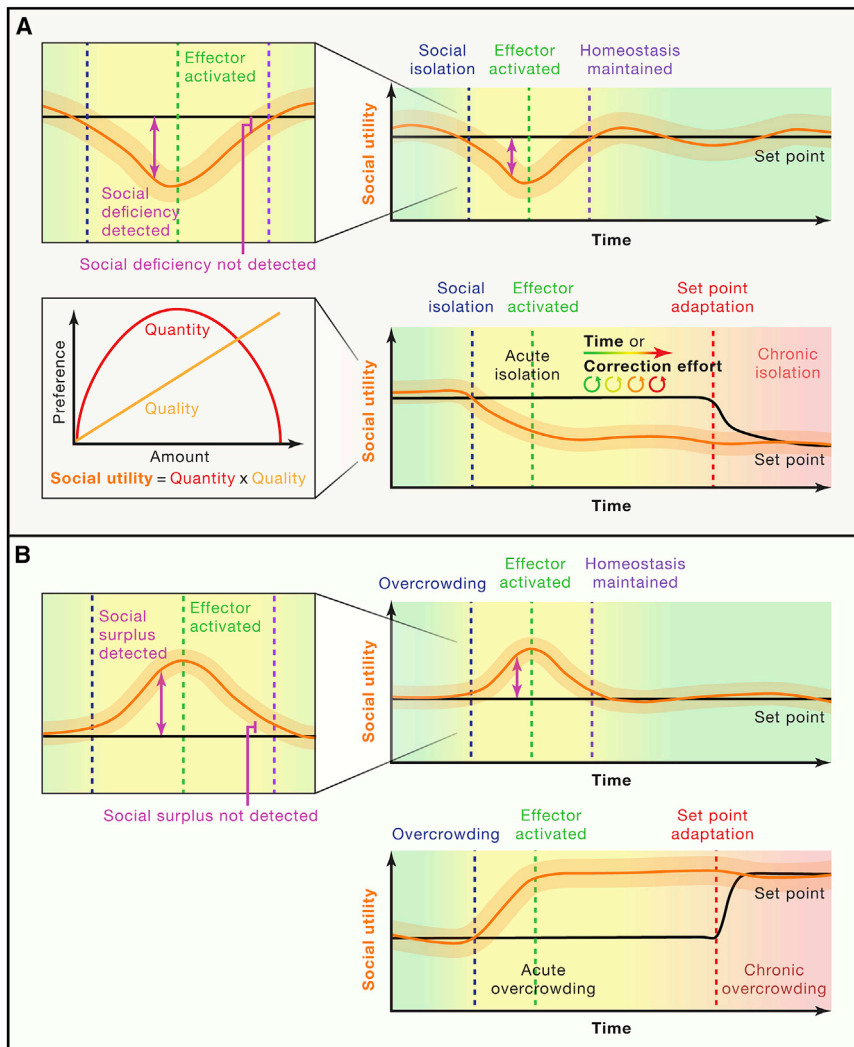
marmoset would respond to a conspecific (Nummela et al., 2017). Free-ranging male baboons use vocalizations—a distinctive "wahoo" call—to indicate their social rank (discussed below), with more frequent and more drawn-out calls indicative of dominance (Fischer et al., 2004). In mice, ultrasonic vocalizations (USVs) provide a mode of communication between individuals. This has been well studied in the context of communication between mothers and pups, where pups—especially those outside the nest—will emit USVs based on changes in their body temperature or in their environment (Ehret, 2005). The acoustic properties of the USVs reflect the pup's degree of arousal and emotional state, and the USVs elicit a phonotactic response in the mother to retrieve her pup (Dulac et al., 2014; Kohl and Dulac, 2018; Schiavo et al., 2020). Interestingly, optogenetic activation of oxytocin neurons from the paraventricular nucleus of the hypothalamus (PVN) in the left auditory cortex of naive virgin mice, who initially do not retrieve pups, accelerates pup retrieval behavior (Marlin et al., 2015).

Humans and non-human primates—and to a lesser extent, rodents (Chen and Hong, 2018)—also rely on vision to perceive agent characteristics. The human fusiform face area has long been known to respond to images of faces, showing increased fMRI signals when a subject is exposed to photos of faces and not inanimate objects or other body parts (Kanwisher et al., 1997). Through fMRI in conjunction with single-unit recordings, studies in the inferotemporal cortex of macaques show that these neurons respond selectively to faces and can differentiate faces by tuning to facial geometrical features (Freiwald et al., 2009; Tsao et al., 2006). Additionally, female macaques can discriminate between the photos of male and female monkeys, with neurons in the orbitofrontal cortex responding to macaque photos in a sex-specific manner, suggesting that macaques can perform sex discrimination based solely on visual information (Mizuno et al., 2007). Conspecific sex information is also represented in the dorsomedial prefrontal cortex (dmPFC) of mice, and specific re-activation of "male" and "female" neuronal ensembles modulate social preference toward male and female conspecific mice, respectively (Kingsbury et al., 2020). Interestingly, humans with extensive bilateral amygdala damage experience dysfunction in recognizing emotions signaled by facial expression, such as fear and anger, indicating that social and emotional processing of social stimuli may be downstream of sensory regions such as visual cortex (Adolphs et al., 1999).

#### Identity and rank

From children recognizing their parents when they come to pick them up to emperor penguins finding their partners and young after months of separation (Jouventin et al., 1999), the ability to recognize an individual's social identity is paramount to evolutionary fitness. Familiarity of a social agent provides context for an interaction, and while some species may prefer interacting with familiar individuals, others, such as mice, prefer to interact with unfamiliar individuals over familiar individuals (Moy et al., 2004). Studies in rodents, non-human primates, and humans show converging evidence that the hippocampus plays a significant role in social memory (for review, see Miller et al., 2016). Neural recordings from mice show stable representations of familiar mice in ventral hippocampal CA1 (vCA1) pyramidal neurons, and optogenetic stimulation of the vCA1 neuronal ensemble





**Figure 2. Hypothesized progression of acute and chronic social isolation and crowding**

The detector node of the social homeostatic system determines the social utility of a social interaction. The detected social utility is a multiplicative function that integrates both the quantity and quality of social interactions (callout).

(A) When a social utility deficit is detected due to social isolation, the effector system activates and drives motivated behavior in an individual to seek social contact. If the effector system is successful in bringing the detected social utility to the set-point encoded by the control center, homeostasis is ultimately maintained and the effector system inactivates (top). In the case where the effector system fails to bring the detected social utility to the encoded set-point, the individual will experience a transition from acute to chronic social isolation (bottom), indicated by a compensatory set-point adjustment that we hypothesize to occur due to time or repeated correction effort by the effector system.

(B) When a social surplus is detected due to overcrowding, the effector system drives antisocial behavior to rescue the discrepancy. If the effector system is successful in bringing the detected social utility to the set-point encoded by the control center, homeostasis is ultimately maintained and the effector system inactivates (top). In the case where the effector system fails to bring the detected social utility to the encoded set-point, the individual will experience a transition from acute to chronic social overcrowding (bottom).

The ability to recognize social rank within a hierarchy is essential to minimize conflict and maintain order, and an individual must be able to know its social rank to emit appropriate behaviors, and to evaluate the quality of social interactions from others. Many highly successful species

live in social hierarchies, benefitting from the advantages they provide. However, the structure of hierarchies is diverse, spanning parameters from large to small, inherited to earned, linear to flat, and despotic to egalitarian. Social structures can be as small as a pair or as large as a civilization. Hierarchies can be maintained for life (such as in female rhesus monkeys who are born into their rank) (for review, see Sapolsky, 2005) or they can be dynamic, such as mice, who upon removal of the alpha will detect a “power vacuum” within minutes (Williamson et al., 2017). Linear social hierarchies were first observed in domestic fowl who demonstrate a pecking order (Schjelderup-Ebbe, 1922) and are also observed in CD1 mice (So et al., 2015). By contrast, honeybees organize themselves into a flat hierarchy and work democratically to make collective decisions, such as finding a new home (Seeley, 2010). Some species, such as Asian elephants, exhibit an egalitarian hierarchical structure owing to a year-round productive environment that favors dispersal of resources to all individuals, such that resources cannot be monopolized (de Silva et al., 2017). Other species with despotic hierarchies, such as African wild dogs and dwarf mongooses, maintain dominance

encoding a familiar mouse can sufficiently induce social discrimination, with the vCA1-nucleus accumbens (NAc) circuit being critical for social memory (Okuyama et al., 2016). Additionally, dorsal hippocampal CA2 (dCA2) has been implicated in social recognition memory in rodents, non-human primates, and humans (for review, see Dudek et al., 2016). Specifically, silencing dCA2 pyramidal neuronal output (Hitti and Siegelbaum, 2014), and in particular, input to vCA1, disrupts social recognition memory, but not other hippocampal-dependent tasks (Meira et al., 2018). dCA2 receives strong input from the supramammillary nucleus (SuM) (Cui et al., 2013), and optogenetic stimulation of SuM inputs into dCA2 via feedforward inhibition induces a social novelty response (Chen et al., 2020). Moreover, dCA2 selectively expresses vasopressin receptors (Avpr1b) in the hippocampus (Young et al., 2006), and optogenetic activation of vasopressinergic neurons from the PVN enhances social memory and is reversible with pharmacological Avpr1b antagonism (Smith et al., 2016). Taken together, these hypothalamic-hippocampal circuits are critically important in the identification of conspecifics, a necessary step in determining the quality of a social interaction.

through aggression over subordinates; in this case, dominant animals show chronically elevated glucocorticoid levels (a proxy for stress) (Creel et al., 1996). Conversely, subordinate baboons and rats (who establish rank in their despotic social hierarchy through intimidation and subordination gestures), show markedly increased glucocorticoid levels (Popova and Naumenko, 1972; Sapolsky, 1990). The heterogeneity in a species' environment and social structure determines the quality of conspecific interactions and the quantity an individual prefers and is therefore a major factor in detecting one's social environment.

How is social rank encoded in the central nervous system? In mice, potentiating and depressing glutamatergic signaling in the medial prefrontal cortex (mPFC) increases and decreases relative hierarchical rank, respectively (Wang et al., 2011). Specifically, optogenetically increasing the synaptic strength of the mediodorsal thalamus (MDT)-to-dmPFC circuit results in sustained winning in the tube test, a common assay used to determine dominance between pairs of mice (Zhou et al., 2017). Additionally, the magnitude of dmPFC activity correlation between a pair of mice during a competitive tube test assay predicts their relative dominance relationship (Kingsbury et al., 2019). Furthermore, a recent study reveals that population-level activity in the mPFC predicts social rank and success in a social competition assay (Padilla-Coreano et al., 2020). In particular, mPFC neurons projecting to the lateral hypothalamus (LH) encode social rank, and optogenetic activation of mPFC-LH neurons promotes social dominance. Although firing rate in BLA neurons correlates with the social rank of conspecifics in rhesus macaques (Munuera et al., 2018), BLA-projecting neurons in the mPFC of mice do not encode social rank (Padilla-Coreano et al., 2020). Remarkably, the encoded relative rank of a mouse can determine the magnitude of sociability and place avoidance when stimulating dorsal raphe nucleus (DRN) dopamine neurons (Matthews et al., 2016) and also the number and size of urine marks a mouse makes, an effect reversible through inhibition of GABAergic medial preoptic area neurons (Hou et al., 2016).

## Control center

### Computing deficits or surpluses

To understand any homeostatic system, we need to be able to measure the signal that is being regulated by the system—and in the case of social homeostasis this signal is not only dynamic but also high-dimensional. For thermoregulation, the signal is body temperature, whereas for osmoregulation, it would be blood osmolality. In these cases, both of these regulated signals are one-dimensional, where an objective measurement can be taken. For social homeostasis, it would be the perceived social environment—which is a high-dimensional state (including factors such as group size, relative rank, hierarchical structure, individual pairwise relationships, and observational learning), where subjective factors including previous experience and preference are paramount. While we postulate certain aspects of the social environment to be objective (e.g., social rank), the perception of other parameters of the social environment may be influenced by individual experiences, innate individual differences, or other contextual factors.

The requisite functions of any homeostatic control center are that (1) it receives input from detector systems, (2) it stores information about the homeostatic set-point, (3) it computes the delta between the received input and the stored set-point, and (4) it sends information about any deviation from the set-point to a downstream effector system (Cannon, 1929).

The social homeostatic control center integrates information about the current state of social engagement with the desired quality/quantity of social contact (homeostatic set-point) to compute deficits or surpluses in social contact. If a deficit or surplus is detected, this deviation from the desired set-point will trigger engagement of the “effector” system to correct the change.

### Surpluses of social contact

Maintenance of social homeostasis is critical for health and survival; deviations (in the form of both deficits and surpluses) from the social homeostatic set-point can result in health consequences. For instance, both objective (persons/room) and subjective (perceived excessive social demands and lack of privacy) human overcrowding in the home, even when controlled for socioeconomic variables, is correlated with poor mental health and social relationships within the home (Gove et al., 1979). Baboons demonstrate an increase in salivary cortisol concentrations following both acute (4-day) and chronic (months-long) overcrowding (Pearson et al., 2015). In mice, social crowding (8 mice/cage) increases anxiety-like behavior and corticosterone levels, as well as hypothalamic corticotropin-releasing factor (CRF), agouti-related peptide (AgRP), and neuropeptide Y (NPY) expression, indicating that crowding is stress inducing (Lin et al., 2015).

Aggression also increases with group size in mice (Van Loo et al., 2001), perhaps indicative of a behavioral mechanism to restore social homeostasis (but see Flanigan et al., 2020; Golden et al., 2016). Interestingly, adult mice who are placed in an overcrowded environment (8 mice/standard laboratory mouse cage) show an increase in plasma corticosterone concentration compared with non-crowded control mice both 1 and 7 days, but not 14 days, after crowding, suggesting an acute stress response to social crowding that diminishes with chronic exposure (Peng et al., 1989).

### Deficits in social contact

With acute social isolation, animals from rodents to humans have been shown to display prosocial behaviors (a rebound of social interaction, increased affiliative behaviors) following social isolation for 24 h (Matthews and Tye, 2019; Matthews et al., 2016; Panksepp and Beatty, 1980; Tomova et al., 2019, 2020b). With acute isolation, social homeostatic systems are engaged, and animals will increase vocalizations (Ehret, 2005; Norcross and Newman, 1993), experience a surge in blood glucocorticoids/cortisol (Rukstalis and French, 2005; Taylor et al., 2014), and expend energy to seek social contact (Panksepp and Beatty, 1980). However, with chronic social isolation, flies, rodents, and humans have been shown to perform antisocial behaviors (aggression, avoidance, social anxiety) (Ma et al., 2011; Sciolino et al., 2010; Zelikowsky et al., 2018) that may be manifested in the form of mental health disorders in our modern-day society.

We hypothesize that, with sustained effort toward correcting this deficit, animals will eventually switch from an active coping

strategy to a passive coping strategy and that the homeostatic set-point will be adjusted to the new baseline of social contact (Figure 2).

### **Mechanisms orchestration/need competition**

The ability to orchestrate competing homeostatic needs and to select one behavioral output that is most adaptive, taking conflicting mechanisms into account, is essential for survival. Importantly, we already know that while separate homeostatic systems may exist for different needs, they may still be interdependent. For example, although separate homeostatic systems exist for hunger and thirst (Augustine et al., 2018; Burton et al., 1976; Sternson, 2013; Zimmerman et al., 2016, 2019), even in a severe energy balance deficit, if an animal is too dehydrated, it will not be able to eat (Bolles, 1961). Similarly, both food deprivation and artificially activating arcuate AgRP neurons—a regulator for appetite—in singly housed male mice shift preference from social interaction with male juvenile and receptive female mice to food consumption in a preference assay (Burnett et al., 2016). Additionally, the intensity of food deprivation determines the magnitude of preference of food over social interaction (Burnett et al., 2019).

Across evolution, social needs exist for any sexually reproductive animal, which includes all mammals. In many cases, social needs can be overshadowed by more urgent threats to our survival. For example, a predatory threat would take precedence over feeding, sleeping, and sex due to the simple calculation that failure to escape the predator would be a greater threat to survival than delaying feeding, sleeping, or sex. Although many social animals can survive in the absence of social contact, the necessity of socialization for reproduction has embedded social needs as essential for species survival, even if not an individual's survival. Thus, it is adaptive for animals to be able to prioritize social needs over basic survival needs until those deficits become life-threatening. For example, bears hibernating will stay nurturing their offspring until internal physiological needs become potential threats on their own survival (Evans et al., 2016). Similarly, penguins alternate between extended periods of foraging or incubating offspring, wherein higher corticosterone levels drove foraging while higher prolactin levels drove incubation (Spée et al., 2010).

### **Effector**

The effector system in any homeostatic system is primarily responsible for driving motivated behavior and physiological adaptations to maintain homeostasis (Cannon, 1929). For instance, in the case of thermoregulation, when an individual's internal temperature is higher than the encoded set-point, the effector in this homeostatic system may motivate the individual to head to a shaded area or it may stimulate sweat glands to cool down the body. Requisites for an effector system include that (1) it must receive input from a control center and (2) activation of the effector system must be able to drive behavior or physiological adaptation independent of the individual's homeostatic need state. The fairly recent advent of neural recording and manipulation technologies has prompted unprecedented insight into the circuits underlying social behavior, such as the bidirectional modulation of social behaviors by BLA-mPFC and BLA-vCA1 (Felix-Ortiz and Tye, 2014; Felix-Ortiz et al., 2016),

the influence of neuromodulation on social reward and social behavior (Dölen et al., 2013; Gunaydin et al., 2014), and the neural circuits underlying aggression (Hong et al., 2014; Lin et al., 2011; Lischinsky and Lin, 2020).

For the purpose of this review, we operationally define the “effector system” as any mechanism that produces a behavior for the purpose of rescuing a social deficit or surplus of an individual and maintains social homeostasis.

### **Prosocial behaviors**

When a social connection deficit is computed, the effector system will attempt to resolve the social inadequacy by driving motivated behavior to interact with others. In doing so, the effector must engage the social motivational circuitry. Acute social isolation generally results in an increase in affiliative, prosocial behaviors across many species. Humans, when experiencing social connection deficits, show increased social memory and heightened social attention toward socially relevant cues, suggesting that a “loneliness” state can activate a social hypervigilant state (Gardner et al., 2005; Pickett et al., 2004). Similarly, following acute (3 day) social isolation, rats demonstrate an increase in social interaction and playful behavior (Panksepp and Beatty, 1980).

Many brain regions are recruited in the social motivation circuitry, including the ventral tegmental area (VTA), which sits at the nexus of the social reward circuitry. Specifically, social interaction during a juvenile intruder task in mice increases activity of VTA dopamine neurons, specifically those neurons that project to D1-receptor-expressing medium spiny neurons in the NAc (Gunaydin et al., 2014). Optogenetic activation of the VTA-NAc projection promotes affiliative social behavior. Additionally, optogenetic stimulation of serotonin terminals (projecting from the DRN) in the NAc also increases social interaction time during a juvenile intruder task (Walsh et al., 2018). In rats, social interaction results in dopamine release in the NAc (Robinson et al., 2002), and antagonizing either D1 or D2 receptors in the NAc decreases time spent playing socially (Manduca et al., 2016). Oxytocin neurons—canonically known for their role in affiliative social behavior, pair bonding, and maternal behaviors (Shamay-Tsoory and Abu-Akel, 2016)—in the PVN project to VTA dopamine neurons, and oxytocin release in the VTA gates social reward (Hung et al., 2017). PVN oxytocin neurons show increased activity during social interactions in mice, and optogenetically and chemogenetically exciting and inhibiting PVN oxytocin neurons enhances and dampens social preference, respectively (Anpilov et al., 2020; Resendez et al., 2020).

While VTA dopamine neurons are typically first considered when investigating dopaminergic involvement in social motivation, another midbrain population of dopamine neurons in the DRN exhibits relevance in producing the “negative” (aversive) drive to motivate social behavior (Matthews et al., 2016). Although historically regarded as a caudal extension of VTA dopamine neurons, DRN dopamine neurons project to distinct downstream regions and are functionally different from VTA dopamine neurons (Cho et al., 2017; Dugalis et al., 2012; Groessl et al., 2018; Hasue and Shammah-Lagnado, 2002; Matthews et al., 2016). In addition, the DRN receives input from hypothalamic areas, such as LH and PVN<sup>Oxt</sup> neurons (Roeling et al., 1993, 1994), regions typically associated with integrating

information about an individual's need state. After 24 h of social isolation, glutamatergic input onto DRN dopamine neurons strengthens in adult male mice, and optogenetically activating these neurons drives both sociability and aversion (Matthews et al., 2016). Additionally, DRN dopamine neurons exhibit generally increased activity during social contact with a novel mouse, as demonstrated by bulk neuronal signals recorded through fiber photometry. Optogenetically inhibiting these neurons decreases sociability only when followed by 24 h of social isolation. Taken together, these results suggest that DRN dopamine neurons are recruited to provide social motivation induced by aversive state. This contrasts with the VTA dopamine-NAc circuit, which seems to elicit the “positive” (rewarding) drive to seek social contact. A recent finding in acutely isolated (10 h) humans shows an increase in social craving alongside increased midbrain dopamine neuronal activity when presented with social cues (Tomova et al., 2020b), consistent with the hypothesized role of DRN dopamine neurons as observed in mice. Altogether, it is tempting to speculate that DRN dopamine neurons may be an element of the effector system of the larger social homeostatic system.

#### Antisocial behaviors

In a manner reminiscent of a social surplus, chronic social isolation has been shown to increase antisocial behaviors upon reintroduction to a social group, namely territorial aggression, across many species, including rodents (Wiberg and Grice, 1963; Malkesman et al., 2006; Wongwitdecha and Marsden, 1996), fish (Gómez-Laplaza and Morgan, 2000; Clayton and Hinde, 1968), and *Drosophila* (Agrawal et al., 2020; Liu et al., 2011). While aggression can be prosocial in certain circumstances, such as in the case of establishing and maintaining social hierarchy (largely mediated by dopamine-transporter-expressing neurons in the ventral premammillary nucleus [Stagkourakis et al., 2018]), aggressive behaviors used to ward off other social agents may be considered antisocial.

Aggression can be understood as an adaptive strategy to access and secure resources. Once a threat is detected in one's environment, key neural populations are recruited to drive an aggressive response (for review, see Lischinsky and Lin, 2020). In particular, the posterior dorsal subdivision of the MeA (MeApd) is a central node in controlling antisocial behaviors in rodents; activation of the GABAergic population drives aggressive and attacking behaviors, while activation of the glutamatergic population drives asocial and self-grooming behaviors (Hong et al., 2014). Additionally, activation of the ventrolateral portion of the VMHvl, which receives projections from the MeA (Pardo-Bellver et al., 2012), produces attacking behavior in both male and female mice (Hashikawa et al., 2017; Lee et al., 2014; Lin et al., 2011). Estrogen receptor alpha (Esr1)-expressing neurons in the posterior amygdala (PA) projecting to the VMHvl are active in intermale aggression, and chemogenetic inhibition of this projection reduces intermale attack duration (Yamaguchi et al., 2020). Glutamatergic VMHvl neurons also project to glutamatergic neurons in the lateral periaqueductal gray (IPAG), which then project to musculature in the jaw to initiate biting; attack and social investigation modulate the activity of VMHvl glutamatergic neurons, while only attack modulates IPAG glutamatergic activity, and inactivation of the VMHvl<sup>VGlut2</sup>-IPAG<sup>VGlut2</sup> projection results in reduced aggression (Falkner et al., 2020).

Neuromodulation may be responsible for producing chronic isolation-induced aggression, as evidenced by conserved mechanisms in different species. Cholecystokinin (CCK) is a neuropeptide that is implicated in negative affective states and disease, such as anxiety-like states and panic disorders, in both rodents and in humans (Rehfeld, 2000; Singh et al., 1991). Singly housed male rats that display aggressive behaviors during a resident-intruder task demonstrate increased CCK concentration in the posterior cortex and tegmentum compared with isolated rats who did not display aggressive behaviors (Panksepp et al., 2004). Interestingly, isolated (4 days) *Drosophila* display downregulation of the neuropeptide drosulfakinin (*Dsk*), the homolog of the mammalian CCK (Agrawal et al., 2020). Both upregulation and downregulation of *Dsk* in *Drosophila* increase isolation-driven aggressive behavior, indicating that *Dsk* balance in the *Drosophila* nervous system plays a critical role in the production of isolation-induced antisocial behavior. Intriguingly, overexpression of tachykinin, but not *Dsk* activation, is sufficient in producing aggressive behaviors in group-housed *Drosophila* (Asahina et al., 2014), suggesting that separate mechanisms may exist in the *Drosophila* central nervous system in the generation of aggressive behaviors depending on the social context.

#### TEMPORAL DYNAMICS OF SOCIAL HOMEOSTASIS (ACUTE VERSUS CHRONIC ISOLATION)

Although we use the terms “acute” and “chronic” to discretize the qualitative phenomenological divergence of consequences from social isolation of different durations, the reference to temporal dynamics is relative rather than absolute. Indeed, the phenomena described related to the diametrically opposed behavioral responses of acute versus chronic social isolation may be mechanistically linked to a parameter other than time, including effort expended, number of rejected attempts, and social consequences to effector-system-activated behaviors.

Because of the increased relevance of the societal impact of social isolation, we focus our discussion on the health consequences of acute and chronic social isolation below.

#### Consequences of acute isolation

##### Physiological health—hypervigilance/arousal, stress, immune response

Isolation from the safety of a social group necessitates modifications in strategy to promote survival. While these strategical changes may prove adaptive in the short term, to the detriment of the isolated individual, the unintended consequences of acute social isolation are often maladaptive.

The immune system is a target for functional adaptation that results from social isolation. Upon social isolation, pro-inflammatory interleukin (IL) genes (such as IL-1B and IL-8) are upregulated, while type I interferon genes associated with antiviral response (such as interferon-stimulated genes and interferon  $\gamma$ -inducible protein family genes) along with antibody production genes are downregulated (Cole, 2014; Cole et al., 2007, 2011). It is hypothesized that this transition occurs to protect the individual from the dangers of being alone (i.e., bacterial infections from physical injury) while reducing the need for protection against viral infections that are normally socially transmitted.



Another major target of physiological adaptations in response to isolation is the hypothalamic-pituitary-adrenal (HPA) axis. The HPA axis acts as a central stress response system, such that when a stressor is detected, a neuroendocrine response is activated whereby the PVN releases CRF into the bloodstream, which then binds to CRF receptors on the anterior pituitary gland and results in adrenocorticotropic hormone (ACTH) release (Chrousos, 2009; Deussing and Chen, 2018; Henckens et al., 2016). ACTH then binds to receptors on the adrenal gland, which results in the release of cortisol. The HPA axis is regulated through a negative feedback loop, such that sufficient cortisol concentrations in the bloodstream will inhibit the release of CRF from the hypothalamus and ACTH from the pituitary gland (Deussing and Chen, 2018; Ramot et al., 2017). While acute HPA axis activation may seem adaptive in that it provides a short-term physiological reaction to an environmental stressor, accumulation of cortisol spikes is maladaptive and can cause physiological “wear and tear” (see *Allostatic load and alternative models*), much less will it assist the individual in returning to social homeostasis. Pair-bonded prairie voles, male mice, and marmosets all exhibit increased cortisol levels following 12 h to 5 days of social isolation, suggesting that the HPA axis is sensitive to acute social isolation (Bosch et al., 2009; McNeal et al., 2014; Rukstalis and French, 2005; Sun et al., 2014; Takatsu-Coleman et al., 2013; Taylor et al., 2014). Additionally, acutely isolated (<24 h) female, but not male, mice demonstrate changes in intrinsic properties of PVN-CRF neurons, including an increase in first spike latency and decreased excitability (Senst et al., 2016). Notably, acute (3 h) crowding in adult male rats also results in an increase of plasma corticosterone (Djordjevi et al., 2003), suggesting that there may be an interaction between deviations in social homeostasis and the central stress response.

In addition to the physiological responses to acute social isolation, changes occur in the central nervous system potentially to guide defensive behavior and protect the isolated individual from environmental threats. Perceived loneliness and objective social isolation produce a state of hypervigilance and heightened arousal, likely an evolutionary feature to detect and protect oneself from environmental threats that may arise from isolation (Cacioppo et al., 2006b). While meaningful for the purposes of threat detection, hypervigilance is not adaptive for social homeostasis, as it does not assist in the maintenance of social needs. Interestingly, individuals who identify as lonely attend more quickly to social threats than those who do not (Cacioppo et al., 2016). Increase in hypervigilance may in part be due to isolation-induced HPA axis activation. Additionally, acute social isolation may modify the noradrenergic system. Noradrenaline is broadly known for its role in increasing arousal and vigilance and is produced in the locus coeruleus (Berridge and Waterhouse, 2003; Sara and Bouret, 2012). In adult male rats, 24 h of social isolation upregulates tyrosine hydroxylase—the rate-limiting enzyme in noradrenaline production—transcription in the locus coeruleus, and antagonizing angiotensin II receptor type 1 (AT<sub>1</sub>) receptors negates the increase in tyrosine hydroxylase (TH) in the locus coeruleus (Saavedra et al., 2006).

### Consequences of chronic isolation

Chronic social isolation (which refers to a relative timescale that may vary with species, context, individual, and experience) results in a number of changes, both arguably adaptive and maladaptive, in an individual's biology and behavior. In this section, we discuss the physiological, mental, and behavioral consequences of chronic social isolation.

#### Physiological

Dysregulation of the HPA axis due to chronic activation results in allostatic load (McEwen, 1998), which describes the cumulative physiological effects of chronic cortisol elevation. Allostatic load is known to result in elevated risk of cardiovascular disease, immune system dysregulation, and cognitive decline through corticosterone binding in the hippocampus (McEwen, 1998). The direction of HPA axis activity resulting from chronic social isolation is highly variable in social organisms. In humans, young and older adults with small social circles and who identify as lonely generally exhibit more chronic activation of HPA axis than those who do not (Arnetz et al., 1983; Pressman et al., 2005). While transient activation of HPA axis can be acutely adaptive to meet the energy demands necessary for overcoming a stressor in an environment, chronic activation, resulting in allostatic load, can cause many adverse physiological effects (Deussing and Chen, 2018; Lee et al., 2015). Chronic social isolation also increases the risk of developing obesity and type 2 diabetes in mice (Nonogaki et al., 2007) and increases the risk of spontaneous, malignant mammary tumor development in rats (Hermes et al., 2009).

#### Mental health

Chronic social isolation in both youth and adulthood can result in changes in biological changes in the central nervous system as well as deleterious changes in behavior, often to the detriment of mental health. Although much of this review is focused on the effects of social isolation on adults, there exists extensive literature on the effects of early-life social deprivation on the learning of social reward (Nardou et al., 2019) and also as developmental models for early-life, stress-induced social dysfunction (Haller et al., 2014; Shin et al., 2018) and schizophrenia (Fone and Porkess, 2008; Lapiz et al., 2003) in rodents. Perhaps one of the most well-known studies on early-life social deprivation is that of Harry Harlow, who showed that early social deprivation resulted in social dysfunction later in life (Harlow and Suomi, 1971). A recent study in mice shows that social reward learning occurs through oxytocin-mediated plasticity in the NAC during a critical window of development (peaking at postnatal day 48) (Nardou et al., 2019), suggesting a developmental role in determining the rewarding value of social interactions. Isolation during the post-weaning period also impairs cognitive function, resulting in long-term social memory dysfunction due to downregulation of ephrin type-B receptor 2 (EphB2) in hippocampal CA1 neurons (Wu et al., 2020). Altogether, the mental health consequences of chronic isolation during youth prove severely detrimental and are unique from those observed in chronically isolated adults.

Chronic social isolation in adulthood also results in many adverse effects in the brain and in behavior. Social isolation

(both 1 and 7 days) impairs long-term social recognition memory in mice through elevated Ras-related C3 botulinum toxin substrate 1 (Rac1) activity in the hippocampus (Liu et al., 2018). Rac1 is a GTPase that has been shown to play a role in active forgetting in *Drosophila*, mice, and rats (Jiang et al., 2016; Liu et al., 2016; Shuai et al., 2010), suggesting a conserved function across species. Additionally, in adult male rats, prolonged social isolation (10–14 weeks) results in sexual behavior deficits, increased anxiety-like behavior, and reduction of transcription factor cyclic AMP (cAMP) receptor element-binding protein (CREB) expression in the shell region of the NAc (Barrot et al., 2005). Notably, these behavioral deficits are rescued through overexpression of CREB in NAc shell. Chronic social isolation also results in anhedonic behaviors, as demonstrated by less sucrose intake and preference, suggesting that chronic social isolation impacts reward processing (Wallace et al., 2009). While overexpression of CREB rescues anxiety-like behavior, it does not rescue anhedonia-related behaviors. Interestingly, chronically isolated (4 weeks) female, but not male, prairie voles also display anhedonia, suggesting that the behavioral effects of chronic social isolation show variability across species and sex (Grippo et al., 2007). Increased resting heart rate, anhedonia, and increased immobility time during forced swim test (indicative of despair observed in depressive-like behaviors) observed in socially isolated female prairie voles are all rescued by subcutaneous injection of oxytocin (Grippo et al., 2009).

### Consequences of perceived loneliness (humans)

Loneliness is a ubiquitous condition that most humans have either indirectly or directly experienced, such as through interactions with an isolated elderly relative or through the more recent global lockdowns to contain the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) outbreak. In humans, perceived deficits in the objective quantity, or subjective quality of social contact (referred to as “loneliness” in psychology [Weiss, 1973]), is correlated with deficits in mental (Hawkey et al., 2006) and physical (Hawkey and Cacioppo, 2010; Hawkey et al., 2006) health and higher mortality rates (Berkman and Syme, 1979; Holt-Lunstad et al., 2010; Holwerda et al., 2012; Perissinotto et al., 2012; Steptoe et al., 2013). Interestingly, perceived loneliness reportedly increases functional communication within the “default network” (regions including the prefrontal cortex that are typically active during wakeful rest), suggesting that rumination may be used to fill a given social void (Spreng et al., 2020). Perceived loneliness correlates with increased morbidity and mortality with diseases such as cancer and cardiovascular disease (Hawkey and Cacioppo, 2003), and most noteworthy perceived loneliness is correlated with the severity of symptoms reported in response to a viral immune challenge (LeRoy et al., 2017), as well as inflammatory responses (Balter et al., 2019). Moreover, on average, humans perceive lonely individuals as having lower achievement and lower social skills and as being less liked and attractive, with lonely male individuals being more stigmatized than lonely female individuals (Lau and Gruen, 1992). The stigmatization surrounding loneliness

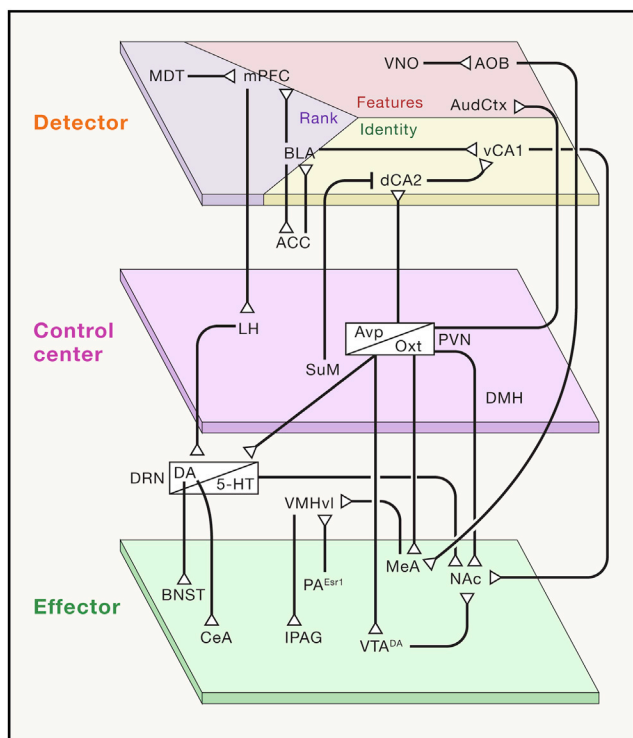
may preclude support to those who need it, further reinforcing loneliness.

### Triggers/parameters that govern the shift

The neurobiological triggers and parameters that regulate the transition that occurs between acute social isolation (which induces prosocial behavior) and chronic social isolation (which induces antisocial behavior) are largely shrouded in mystery, but we do have a few potential clues. Notably, a major shift occurs such that an active coping strategy (prosocial, affiliative behaviors) turns into a passive coping strategy (antisocial behaviors and reduced effort to seek social contact) prior to reintroduction to a social group. In adult mice, there is an increase in aggressive behaviors that occurs after 48 h, but not 24 h, of social isolation (Lister and Hilakivi, 1988), and aggressive behaviors in socially isolated adult mice gradually increase for the first few weeks and plateau at ~4 weeks of social isolation (Matsumoto et al., 2005). However, in adult rats, there is an increase in affiliative behavior that lasts up to 7 days following social isolation (Niesink and van Ree, 1982), suggesting that the time course for the transition from prosocial to antisocial behavior is heterogeneous and depends on the species of the animal. Siamese fighting fish, for instance, demonstrate a much quicker transition, displaying aggressive behaviors soon after social isolation (as early as 15 min) (Gómez-Laplaza and Morgan, 2000; Clayton and Hinde, 1968).

What mechanism underlies the transition of behaviors that occurs from acute to chronic social isolation? A recent study shows that, in mice, chronic (2 week) social isolation results in brain-wide upregulation of tachykinin 2 (Tac2) and neurokinin B (NkB), which endogenously binds to the G-protein-coupled neurokinin 3 (NK3) receptor (NK3R) (Zelikowsky et al., 2018). Tachykinin activation and silencing in *Drosophila* have previously been shown to modulate intermale aggressive behaviors (Asahina et al., 2014). Targeted NK3R antagonism specifically in the dorsomedial hypothalamus (DMH) mitigates chronic isolation-induced aggression in mice, whereas NK3R antagonism in central nucleus of the amygdala (CeA) and bed nucleus of the stria terminalis (BNST) mitigates other aspects of chronic isolation-induced behaviors, such as fear expression (Zelikowsky et al., 2018). Interestingly, knocking out *Avpr1b* in mice reduces chronic (2 week) isolation-induced aggression and is rescued by expressing *Avpr1b* into *dCA2*, suggesting that vasopressinergic signaling in *dCA2* may be involved in the transition from affiliative to antisocial behaviors resulting from chronic social isolation as well (Pagani et al., 2015).

We postulate that the transition occurs due to either (1) time or (2) correction effort by the social homeostatic effector system (Figure 2A). Perhaps there is a time-keeping component of the neural circuitry that approximates the duration of time to identify that a transient environmental change is actually long-lasting enough to trigger a neuroadaptation. Alternatively, rather than the trigger being dictated by external signals, the transition from acute to chronic isolation behaviors may be triggered by a threshold of effort exerted to correct the surplus or deficit. Future experiments are needed to differentiate between these possibilities.



**Figure 3. Neural circuits underlying homeostatic nodes**

Proposed neural circuits involved in each social homeostatic node. Note: we acknowledge that there are dynamic states that allow flexibility in the positioning of each of the regions and circuits that we speculate to be functionally representing each node of the social homeostatic system. We also acknowledge that there will likely be functional heterogeneity among neurons in any brain region and that many functions are indeed distributed across both local and long-range circuits. ACC, anterior cingulate cortex; AOB, accessory olfactory bulb; AudCtx, auditory cortex; BLA, basolateral amygdala; BNST, bed nucleus of the stria terminalis; CeA, central nucleus of the amygdala; dCA2, dorsal hippocampal CA2; DMH, dorsomedial hypothalamus; DRN, dorsal raphe nucleus (DA: dopamine; 5-HT: serotonin); LH, lateral hypothalamus; IPAG, lateral periaqueductal gray; MDT, mediadorsal thalamus; MeA, medial amygdala; mPFC, medial prefrontal cortex; NAc, nucleus accumbens; PA, posterior amygdala (Esr1: estrogen 1 receptor); PVN, paraventricular nucleus of the hypothalamus (Oxt: oxytocin; Avp: vasopressin); SuM, supramammillary nucleus; vCA1, ventral hippocampal CA1; VMHvl, ventrolateral portion of the ventromedial hypothalamus; VNO, vomeronasal organ; VTA<sup>DA</sup>, ventral tegmental area (DA).

Considering that many of the systems that drive motivated social behavior are mechanistically performed through neuromodulation, it is plausible that chronic activation of neuromodulator systems to correct for social deficits may ultimately trigger the transition from acute to chronic isolation. Many examples exist of neuromodulator plasticity induced by aversive experiences. For instance, adult male mice who undergo a social defeat paradigm and are susceptible to social subordination exhibit changes in mesolimbic dopamine system, including increased firing rate of VTA dopamine neurons and increased brain-derived neurotrophic factor (BDNF) signaling in the NAc, inducing plasticity in the VTA-NAc circuit that is involved in emotion and reward processing (Krishnan et al., 2007). Additionally, while stimulation of VTA-NAc in adult female mice promotes social behavior (Gunaydin et al., 2014), stimulation of the VTA-NAc circuit following

recent or remote stress produces an antisocial effect and increases optically induced dopamine release in the NAc (Wichmann et al., 2017), suggesting that stressors can induce circuit plasticity that ultimately changes behavioral output. Given that acute social isolation increases the activity of DRN dopamine neurons in response to social stimuli (Matthews et al., 2016), it is tempting to speculate that the chronic homeostatic correction effort expended by the DRN dopamine neurons may indeed induce downstream plasticity that is ultimately responsible for the shift from prosocial to antisocial behaviors.

### Allostatic load and alternative models

In addition to considering the immediate social challenge (a deficit or surplus), it may be important to consider the history of challenges as well as the duration of the challenge. An important model that may extend a social homeostatic model is the contribution of increasing allostatic load as being the “trigger” to the negative health consequences of social isolation that resemble those of chronic stress-induced disease states linked to allostatic load (e.g., increased depression, increased morbidity, and mortality of cancer and heart disease) (Cacioppo and Hawley, 2003; Juster et al., 2010; Karamihalev et al., 2020). An allostatic load model may also be relevant when considering the microstructure of homeostatic set-point adaptation (Lee et al., 2015; Sandi and Haller, 2015).

An alternative model would be to consider social challenges as a subset of stressors, and to contribute to general HPA axis, as discussed above. It is yet unclear to what degree stressors accumulate along specific dimensions (where social challenges would be detected separately from other stress challenges) or whether they are aggregated irrespective of the modality or dimension.

### OUTLOOK

Amidst a global pandemic with unpredictable lockdowns and varying quarantine guidelines around the world, no other time in recent history has compelled the urgent need to interrogate the neural mechanisms of social homeostasis more than now. Efforts to contain the SARS-CoV-2 outbreak demand urgent need in parallel with further research on the unintended health consequences of social isolation. At present, existing frameworks from other homeostatic systems and recent findings in socially isolated animals allow us an entry point to conceptualize the neural circuits and mechanisms underlying social homeostasis (Figure 3), although much work is left in fully understanding this system. Considering pandemics and other socially disruptive events are likely to occur again, proactive research into the costs of social isolation can mitigate the effects of an uncertain future.

### ACKNOWLEDGMENTS

K.M.T. is the Wylie Vale Chair at Salk Institute for Biological Studies, a New York Stem Cell Foundation-Robertson Investigator, and McKnight Scholar. This work was supported by funding from the JPB Foundation, the PIIF, PNDRF, JFDP, Alfred P. Sloan Foundation, New York Stem Cell Foundation, Klingenstein Foundation, McKnight Foundation, Clayton Foundation, Dolby Family Fund, R01-MH102441 (NIMH), R01-MH115920 (NIMH), RF1-

AG047661 (NIA), the NIH Director's New Innovator Award DP2-DK102256 (NIDDK), and Pioneer Award DP1-AT009925 (NCCIH). C.R.L. is supported by a fellowship from the University of California, San Diego, Neurosciences Graduate Program.

## REFERENCES

- Adolphs, R., Tranel, D., Hamann, S., Young, A.W., Calder, A.J., Phelps, E.A., Anderson, A., Lee, G.P., and Damasio, A.R. (1999). Recognition of facial emotion in nine individuals with bilateral amygdala damage. *Neuropsychologia* *37*, 1111–1117.
- Agrawal, P., Kao, D., Chung, P., and Looger, L.L. (2020). The neuropeptide Drosulfakinin regulates social isolation-induced aggression in *Drosophila*. *J. Exp. Biol.* *223*, jeb207407.
- Allsop, S.A., Wichmann, R., Mills, F., Burgos-Robles, A., Chang, C.-J., Felix-Ortiz, A.C., Vienne, A., Beyeler, A., Izadmehr, E.M., Glober, G., et al. (2018). Corticoamygdala Transfer of Socially Derived Information Gates Observational Learning. *Cell* *173*, 1329–1342.e18.
- Anpilov, S., Shemesh, Y., Eren, N., Harony-Nicolas, H., Benjamin, A., Dine, J., Oliveira, V.E.M., Forkosh, O., Karamihalev, S., Hüttl, R.-E., et al. (2020). Wireless Optogenetic Stimulation of Oxytocin Neurons in a Semi-natural Setup Dynamically Elevates Both Pro-social and Agonistic Behaviors. *Neuron* *107*, 644–655.e7.
- Arnetz, B.B., Theorell, T., Levi, L., Kallner, A., and Eneroth, P. (1983). An experimental study of social isolation of elderly people: psychoendocrine and metabolic effects. *Psychosom. Med.* *45*, 395–406.
- Asahina, K., Watanabe, K., Duistermars, B.J., Hoopfer, E., González, C.R., Eyjólfsson, E.A., Perona, P., and Anderson, D.J. (2014). Tachykinin-expressing neurons control male-specific aggressive arousal in *Drosophila*. *Cell* *156*, 221–235.
- Augustine, V., Gokce, S.K., Lee, S., Wang, B., Davidson, T.J., Reimann, F., Gribble, F., Deisseroth, K., Lois, C., and Oka, Y. (2018). Hierarchical neural architecture underlying thirst regulation. *Nature* *555*, 204–209.
- Bai, L., Mesgarzadeh, S., Ramesh, K.S., Huey, E.L., Liu, Y., Gray, L.A., Aitken, T.J., Chen, Y., Beutler, L.R., Ahn, J.S., et al. (2019). Genetic Identification of Vagal Sensory Neurons That Control Feeding. *Cell* *179*, 1129–1143.e23.
- Balter, L.J.T., Raymond, J.E., Aldred, S., Drayson, M.T., Veldhuijzen van Zanten, J.J.C.S., Higgs, S., and Bosch, J.A. (2019). Loneliness in healthy young adults predicts inflammatory responsiveness to a mild immune challenge in vivo. *Brain Behav. Immun.* *82*, 298–301.
- Barrot, M., Wallace, D.L., Bolaños, C.A., Graham, D.L., Perrotti, L.I., Neve, R.L., Chambliss, H., Yin, J.C., and Nestler, E.J. (2005). Regulation of anxiety and initiation of sexual behavior by CREB in the nucleus accumbens. *Proc. Natl. Acad. Sci. USA* *102*, 8357–8362.
- Ben-Ami Bartal, I., Decety, J., and Mason, P. (2011). Empathy and pro-social behavior in rats. *Science* *334*, 1427–1430.
- Ben-Ami Bartal, I., Rodgers, D.A., Bernardes Sarria, M.S., Decety, J., and Mason, P. (2014). Pro-social behavior in rats is modulated by social experience. *eLife* *3*, e01385.
- Berkman, L.F., and Syme, S.L. (1979). Social networks, host resistance, and mortality: a nine-year follow-up study of Alameda County residents. *Am. J. Epidemiol.* *109*, 186–204.
- Berridge, C.W., and Waterhouse, B.D. (2003). The locus coeruleus-noradrenergic system: modulation of behavioral state and state-dependent cognitive processes. *Brain Res. Brain Res. Rev.* *42*, 33–84.
- Betley, J.N., Cao, Z.F.H., Ritola, K.D., and Sternson, S.M. (2013). Parallel, redundant circuit organization for homeostatic control of feeding behavior. *Cell* *155*, 1337–1350.
- Bolles, R.C. (1961). The interaction of hunger and thirst in the rat. *J. Comp. Physiol. Psychol.* *54*, 580–584.
- Bosch, O.J., Nair, H.P., Ahern, T.H., Neumann, I.D., and Young, L.J. (2009). The CRF system mediates increased passive stress-coping behavior following the loss of a bonded partner in a monogamous rodent. *Neuropsychopharmacology* *34*, 1406–1415.
- Burnett, C.J., Li, C., Webber, E., Tsaousidou, E., Xue, S.Y., Brüning, J.C., and Krashes, M.J. (2016). Hunger-Driven Motivational State Competition. *Neuron* *92*, 187–201.
- Burnett, C.J., Funderburk, S.C., Navarrete, J., Sabol, A., Liang-Guallpa, J., Desrochers, T.M., and Krashes, M.J. (2019). Need-based prioritization of behavior. *eLife* *8*, e44527.
- Burton, M.J., Rolls, E.T., and Mora, F. (1976). Effects of hunger on the responses of neurons in the lateral hypothalamus to the sight and taste of food. *Exp. Neurol.* *51*, 668–677.
- Cacioppo, J.T., and Hawkley, L.C. (2003). Social isolation and health, with an emphasis on underlying mechanisms. *Perspect. Biol. Med.* *46* (3, Suppl), S39–S52.
- Cacioppo, J.T., Hughes, M.E., Waite, L.J., Hawkley, L.C., and Thisted, R.A. (2006a). Loneliness as a specific risk factor for depressive symptoms: cross-sectional and longitudinal analyses. *Psychol. Aging* *21*, 140–151.
- Cacioppo, J.T., Hawkley, L.C., Ernst, J.M., Burleson, M., Berntson, G.G., Nouriani, B., and Spiegel, D. (2006b). Loneliness within a nomological net: An evolutionary perspective. *J. Res. Pers.* *40*, 1054–1085.
- Cacioppo, S., Bangee, M., Balogh, S., Cardenas-Iniguez, C., Qualter, P., and Cacioppo, J.T. (2016). Loneliness and implicit attention to social threat: A high-performance electrical neuroimaging study. *Cogn. Neurosci.* *7*, 138–159.
- Cannon, W.B. (1929). Organization for Physiological Homeostasis. *Physiol. Rev.* *9*, 399–431.
- Chen, P., and Hong, W. (2018). Neural Circuit Mechanisms of Social Behavior. *Neuron* *98*, 16–30.
- Chen, S., He, L., Huang, A.J.Y., Boehringer, R., Robert, V., Wintzer, M.E., Polygalov, D., Weitemier, A.Z., Tao, Y., Gu, M., et al. (2020). A hypothalamic novelty signal modulates hippocampal memory. *Nature* *586*, 270–274.
- Cho, J.R., Treweek, J.B., Robinson, J.E., Xiao, C., Bremner, L.R., Greenbaum, A., and Gradinaru, V. (2017). Dorsal Raphe Dopamine Neurons Modulate Arousal and Promote Wakefulness by Salient Stimuli. *Neuron* *94*, 1205–1219.e8.
- Christian, J.J. (1961). Phenomena associated with population density. *Proc. Natl. Acad. Sci. USA* *47*, 428–449.
- Christian, J.J. (1970). Social subordination, population density, and mammalian evolution. *Science* *168*, 84–90.
- Chrousos, G.P. (2009). Stress and disorders of the stress system. *Nat. Rev. Endocrinol.* *5*, 374–381.
- Clancy, A.N., Coquelin, A., Macrides, F., Gorski, R.A., and Noble, E.P. (1984). Sexual behavior and aggression in male mice: involvement of the vomeronasal system. *J. Neurosci.* *4*, 2222–2229.
- Clayton, F.L., and Hinde, R.A. (1968). The habituation and recovery of aggressive display in *Betta splendens*. *Behaviour* *30*, 96–106.
- Cole, S.W. (2014). Human Social Genomics. *PLOS Genetics* *10*, e1004601.
- Cole, S.W., Hawkley, L.C., Arevalo, J.M.G., and Cacioppo, J.T. (2011). Transcript origin analysis identifies antigen-presenting cells as primary targets of socially regulated gene expression in leukocytes. *PNAS* *108*, 3080–3085.
- Cole, S.W., Hawkley, L.C., Arevalo, J.M., Sung, C.Y., Rose, R.M., and Cacioppo, J.T. (2007). Social regulation of gene expression in human leukocytes. *Genome Biol.* *8*, R189.
- Creel, S., MarushaCreel, N., and Monfort, S.L. (1996). Social stress and dominance. *Nature* *379*, 212.
- Cui, Z., Gerfen, C.R., and Young, W.S., 3rd. (2013). Hypothalamic and other connections with dorsal CA2 area of the mouse hippocampus. *J. Comp. Neurol.* *521*, 1844–1866.
- de Silva, S., Schmid, V., and Wittemyer, G. (2017). Fission fusion processes weaken dominance networks of female Asian elephants in a productive habitat. *Behav. Ecol.* *28*, 243–252.
- Deussing, J.M., and Chen, A. (2018). The Corticotropin-Releasing Factor Family: Physiology of the Stress Response. *Physiol. Rev.* *98*, 2225–2286.



- Djordjevi, J., Cviji, G., and Davidovi, V. (2003). Different Activation of ACTH and Corticosterone Release in Response to Various Stressors in Rats. *Physiol. Res.* 52, 67–72.
- Dölen, G., Darvishzadeh, A., Huang, K.W., and Malenka, R.C. (2013). Social reward requires coordinated activity of nucleus accumbens oxytocin and serotonin. *Nature* 501, 179–184.
- Dougalis, A.G., Matthews, G.A.C., Bishop, M.W., Brischoux, F., Kobayashi, K., and Ungless, M.A. (2012). Functional properties of dopamine neurons and co-expression of vasoactive intestinal polypeptide in the dorsal raphe nucleus and ventro-lateral periaqueductal grey. *Eur. J. Neurosci.* 36, 3322–3332.
- Dudek, S.M., Alexander, G.M., and Farris, S. (2016). Rediscovering area CA2: unique properties and functions. *Nat. Rev. Neurosci.* 17, 89–102.
- Dulac, C., O'Connell, L.A., and Wu, Z. (2014). Neural control of maternal and paternal behaviors. *Science* 345, 765–770.
- Ehret, G. (2005). Infant rodent ultrasounds – a gate to the understanding of sound communication. *Behav. Genet.* 35, 19–29.
- Eliades, S.J., and Miller, C.T. (2017). Marmoset vocal communication: Behavior and neurobiology. *Dev. Neurobiol.* 77, 286–299.
- Evans, G.W., Lepore, S.J., Shejwal, B.R., and Palsane, M.N. (1998). Chronic residential crowding and children's well-being: an ecological perspective. *Child Dev.* 69, 1514–1523.
- Evans, A.L., Singh, N.J., Friebe, A., Amemo, J.M., Laske, T.G., Fröbert, O., Swenson, J.E., and Blanc, S. (2016). Drivers of hibernation in the brown bear. *Front. Zool.* 13, 7.
- Falkner, A.L., Wei, D., Song, A., Watsek, L.W., Chen, I., Chen, P., Feng, J.E., and Lin, D. (2020). Hierarchical Representations of Aggression in a Hypothalamic-Midbrain Circuit. *Neuron* 106, 637–648.e6.
- Felix-Ortiz, A.C., and Tye, K.M. (2014). Amygdala inputs to the ventral hippocampus bidirectionally modulate social behavior. *J. Neurosci.* 34, 586–595.
- Felix-Ortiz, A.C., Burgos-Robles, A., Bhagat, N.D., Leppla, C.A., and Tye, K.M. (2016). Bidirectional modulation of anxiety-related and social behaviors by amygdala projections to the medial prefrontal cortex. *Neuroscience* 327, 197–209.
- Ferrero, D.M., Moeller, L.M., Osakada, T., Horio, N., Li, Q., Roy, D.S., Cichy, A., Spehr, M., Touhara, K., and Liberles, S.D. (2013). A juvenile mouse pheromone inhibits sexual behaviour through the vomeronasal system. *Nature* 502, 368–371.
- Fischer, J., Kitchen, D.M., Seyfarth, R.M., and Cheney, D.L. (2004). Baboon loud calls advertise male quality: acoustic features and their relation to rank, age, and exhaustion. *Behav. Ecol. Sociobiol.* 56, 140–148.
- Flanigan, M.E., Aleyasin, H., Li, L., Burnett, C.J., Chan, K.L., LeClair, K.B., Lucas, E.K., Matikainen-Ankney, B., Durand-de Cuttoli, R., Takahashi, A., et al. (2020). Orexin signaling in GABAergic lateral habenula neurons modulates aggressive behavior in male mice. *Nature Neuroscience* 23, 638–650.
- Fone, K.C.F., and Porkess, M.V. (2008). Behavioural and neurochemical effects of post-weaning social isolation in rodents—relevance to developmental neuropsychiatric disorders. *Neurosci. Biobehav. Rev.* 32, 1087–1102.
- Forkosh, O., Karamihalev, S., Roeh, S., Alon, U., Anpilov, S., Touma, C., Nussbaumer, M., Flachskamm, C., Kaplick, P.M., Shemesh, Y., and Chen, A. (2019). Identity domains capture individual differences from across the behavioral repertoire. *Nat. Neurosci.* 22, 2023–2028.
- Freiwald, W.A., Tsao, D.Y., and Livingstone, M.S. (2009). A face feature space in the macaque temporal lobe. *Nat. Neurosci.* 12, 1187–1196.
- Gardner, W.L., Pickett, C.L., Jefferis, V., and Knowles, M. (2005). On the outside looking in: Loneliness and social monitoring. *Pers. Soc. Psychol. Bull.* 31, 1549–1560.
- Geyer, M.A., Wilkinson, L.S., Humby, T., and Robbins, T.W. (1993). Isolation rearing of rats produces a deficit in prepulse inhibition of acoustic startle similar to that in schizophrenia. *Biol. Psychiatry* 34, 361–372.
- Golden, S.A., Heshmati, M., Flanigan, M., Christoffel, D.J., Guise, K., Pfau, M.L., Aleyasin, H., Menard, C., Zhang, H., Hodes, G.E., et al. (2016). Basal forebrain projections to the lateral habenula modulate aggression reward. *Nature* 534, 688–692.
- Gómez-Laplaza, L.M., and Morgan, E. (2000). Laboratory studies of the effects of short-term isolation on aggressive behaviour in fish. *Mar. Freshwat. Behav. Physiol.* 33, 63–102.
- Gove, W.R., Hughes, M., and Galle, O.R. (1979). Overcrowding in the home: an empirical investigation of its possible pathological consequences. *Am. Sociol. Rev.* 44, 59–80.
- Grippe, A.J., Gerena, D., Huang, J., Kumar, N., Shah, M., Ughreja, R., and Carter, C.S. (2007). Social isolation induces behavioral and neuroendocrine disturbances relevant to depression in female and male prairie voles. *Psychoneuroendocrinology* 32, 966–980.
- Grippe, A.J., Trahanas, D.M., Zimmerman, R.R., 2nd, Porges, S.W., and Carter, C.S. (2009). Oxytocin protects against negative behavioral and autonomic consequences of long-term social isolation. *Psychoneuroendocrinology* 34, 1542–1553.
- Groessl, F., Munsch, T., Meis, S., Griessner, J., Kaczanowska, J., Pliota, P., Kargl, D., Badurek, S., Kraitsy, K., Rassouli, A., et al. (2018). Dorsal tegmental dopamine neurons gate associative learning of fear. *Nat. Neurosci.* 21, 952–962.
- Gunaydin, L.A., Grosenick, L., Finkelstein, J.C., Kauvar, I.V., Fenno, L.E., Adhikari, A., Lammel, S., Mirzabekov, J.J., Airan, R.D., Zalocusky, K.A., et al. (2014). Natural neural projection dynamics underlying social behavior. *Cell* 157, 1535–1551.
- Haga-Yamanaka, S., Ma, L., He, J., Qiu, Q., Lavis, L.D., Looger, L.L., and Yu, C.R. (2014). Integrated action of pheromone signals in promoting courtship behavior in male mice. *eLife* 3, e03025.
- Haller, J., Harold, G., Sandi, C., and Neumann, I.D. (2014). Effects of Adverse Early-Life Events on Aggression and Anti-Social Behaviours in Animals and Humans. *Journal of Neuroendocrinology* 26, 724–738.
- Harlow, H.F., and Suomi, S.J. (1971). Social recovery by isolation-reared monkeys. *Proc. Natl. Acad. Sci. USA* 68, 1534–1538.
- Hashikawa, K., Hashikawa, Y., Falkner, A., and Lin, D. (2016). The neural circuits of mating and fighting in male mice. *Curr. Opin. Neurobiol.* 38, 27–37.
- Hashikawa, K., Hashikawa, Y., Tremblay, R., Zhang, J., Feng, J.E., Sabol, A., Piper, W.T., Lee, H., Rudy, B., and Lin, D. (2017). Esr1<sup>+</sup> cells in the ventromedial hypothalamus control female aggression. *Nat. Neurosci.* 20, 1580–1590.
- Hasue, R.H., and Shammah-Lagnado, S.J. (2002). Origin of the dopaminergic innervation of the central extended amygdala and accumbens shell: a combined retrograde tracing and immunohistochemical study in the rat. *J. Comp. Neurol.* 454, 15–33.
- Hawkey, L.C., and Cacioppo, J.T. (2003). Loneliness and pathways to disease. *Brain Behav. Immun.* 17 (Suppl 1), S98–S105.
- Hawkey, L.C., and Cacioppo, J.T. (2010). Loneliness matters: a theoretical and empirical review of consequences and mechanisms. *Ann. Behav. Med.* 40, 218–227.
- Hawkey, L.C., Masi, C.M., Berry, J.D., and Cacioppo, J.T. (2006). Loneliness is a unique predictor of age-related differences in systolic blood pressure. *Psychol. Aging* 21, 152–164.
- Henckens, M.J., Deussing, J.M., and Chen, A. (2016). Region-specific roles of the corticotropin-releasing factor-urocortin system in stress. *Nat. Rev. Neurosci.* 17, 636–651.
- Hermes, G.L., Delgado, B., Tretiakova, M., Cavigelli, S.A., Krausz, T., Conzen, S.D., and McClintock, M.K. (2009). Social isolation dysregulates endocrine and behavioral stress while increasing malignant burden of spontaneous mammary tumors. *Proc. Natl. Acad. Sci. USA* 106, 22393–22398.
- Hitti, F.L., and Siegelbaum, S.A. (2014). The hippocampal CA2 region is essential for social memory. *Nature* 508, 88–92.
- Hoekstra, H.E., and Coyne, J.A. (2007). The locus of evolution: evo devo and the genetics of adaptation. *Evolution* 61, 995–1016.
- Holt-Lunstad, J., Smith, T.B., and Layton, J.B. (2010). Social relationships and mortality risk: a meta-analytic review. *PLoS Med.* 7, e1000316.

- Holwerda, T.J., Beekman, A.T.F., Deeg, D.J.H., Stek, M.L., van Tilburg, T.G., Visser, P.J., Schmand, B., Jonker, C., and Schoevers, R.A. (2012). Increased risk of mortality associated with social isolation in older men: only when feeling lonely? Results from the Amsterdam Study of the Elderly (AMSTEL). *Psychol. Med.* **42**, 843–853.
- Hong, W., Kim, D.-W., and Anderson, D.J. (2014). Antagonistic control of social versus repetitive self-grooming behaviors by separable amygdala neuronal subsets. *Cell* **158**, 1348–1361.
- Hou, X.H., Hyun, M., Taranda, J., Huang, K.W., Todd, E., Feng, D., Atwater, E., Croney, D., Zeidel, M.L., Osten, P., and Sabatini, B.L. (2016). Central Control Circuit for Context-Dependent Micturition. *Cell* **167**, 73–86.e12.
- Hresko, T. (2006). In the Cellars of the Hollow Men: Use of Solitary Confinement in U.S. Prisons and Its Implications under International Laws against Torture. *Pace Int. Law Rev.* **18**, 1.
- Hung, L.W., Neuner, S., Polepalli, J.S., Beier, K.T., Wright, M., Walsh, J.J., Lewis, E.M., Luo, L., Deisseroth, K., Dölen, G., and Malenka, R.C. (2017). Gating of social reward by oxytocin in the ventral tegmental area. *Science* **357**, 1406–1411.
- Jiang, L., Mao, R., Zhou, Q., Yang, Y., Cao, J., Ding, Y., Yang, Y., Zhang, X., Li, L., and Xu, L. (2016). Inhibition of Rac1 Activity in the Hippocampus Impairs the Forgetting of Contextual Fear Memory. *Mol. Neurobiol.* **53**, 1247–1253.
- Jouventin, P., Lengagne, T., and Aubin, T. (1999). Finding One's Mate in a King Penguin Colony: Efficiency of Acoustic Communication. *Behaviour* **136**, 833–846.
- Juster, R.-P., McEwen, B.S., and Lupien, S.J. (2010). Allostatic load biomarkers of chronic stress and impact on health and cognition. *Neurosci. Biobehav. Rev.* **35**, 2–16.
- Kanwisher, N., McDermott, J., and Chun, M.M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* **17**, 4302–4311.
- Karamihalev, S., Brivio, E., Flachskamm, C., Stoffel, R., Schmidt, M.V., and Chen, A. (2020). Social dominance mediates behavioral adaptation to chronic stress in a sex-specific manner. *eLife* **9**, e58723.
- Kaur, A.W., Ackels, T., Kuo, T.-H., Cichy, A., Dey, S., Hays, C., Kateri, M., Logan, D.W., Marton, T.F., Spehr, M., and Stowers, L. (2014). Murine pheromone proteins constitute a context-dependent combinatorial code governing multiple social behaviors. *Cell* **157**, 676–688.
- Kingsbury, L., Huang, S., Wang, J., Gu, K., Golshani, P., Wu, Y.E., and Hong, W. (2019). Correlated Neural Activity and Encoding of Behavior across Brains of Socially Interacting Animals. *Cell* **178**, 429–446.e16.
- Kingsbury, L., Huang, S., Raam, T., Ye, L.S., Wei, D., Hu, R.K., Ye, L., and Hong, W. (2020). Cortical Representations of Conspecific Sex Shape Social Behavior. *Neuron* **107**, 941–953.e7.
- Kohl, J., and Dulac, C. (2018). Neural control of parental behaviors. *Curr. Opin. Neurobiol.* **49**, 116–122.
- Kohn, M.L., and Clausen, J.A. (1955). Social Isolation and Schizophrenia. *Am. Sociol. Rev.* **20**, 265–273.
- Krishnan, V., Han, M.-H., Graham, D.L., Berton, O., Renthal, W., Russo, S.J., Laplant, Q., Graham, A., Lutter, M., Lagace, D.C., et al. (2007). Molecular adaptations underlying susceptibility and resistance to social defeat in brain reward regions. *Cell* **131**, 391–404.
- Lapiz, M.D.S., Fulford, A., Muchimapura, S., Mason, R., Parker, T., and Marsden, C.A. (2003). Influence of postweaning social isolation in the rat on brain development, conditioned behavior, and neurotransmission. *Neurosci. Behav. Physiol.* **33**, 13–29.
- Lau, S., and Gruen, G.E. (1992). The Social Stigma of Loneliness: Effect of Target Person's and Perceiver's Sex. *Pers. Soc. Psychol. Bull.* **18**, 182–189.
- Lee, H., Kim, D.-W., Remedios, R., Anthony, T.E., Chang, A., Madisen, L., Zeng, H., and Anderson, D.J. (2014). Scalable control of mounting and attack by Esr1+ neurons in the ventromedial hypothalamus. *Nature* **509**, 627–632.
- Lee, D.Y., Kim, E., and Choi, M.H. (2015). Technical and clinical aspects of cortisol as a biochemical marker of chronic stress. *BMB Rep.* **48**, 209–216.
- LeRoy, A.S., Murdock, K.W., Jaremka, L.M., Loya, A., and Fagundes, C.P. (2017). Loneliness predicts self-reported cold symptoms after a viral challenge. *Health Psychol.* **36**, 512–520.
- Leypold, B.G., Yu, C.R., Leinders-Zufall, T., Kim, M.M., Zufall, F., and Axel, R. (2002). Altered sexual and social behaviors in *trp2* mutant mice. *Proc. Natl. Acad. Sci. USA* **99**, 6376–6381.
- Lim, M.M., Wang, Z., Olazábal, D.E., Ren, X., Terwilliger, E.F., and Young, L.J. (2004). Enhanced partner preference in a promiscuous species by manipulating the expression of a single gene. *Nature* **429**, 754–757.
- Lin, D., Boyle, M.P., Dollar, P., Lee, H., Lein, E.S., Perona, P., and Anderson, D.J. (2011). Functional identification of an aggression locus in the mouse hypothalamus. *Nature* **470**, 221–226.
- Lin, E.-J.D., Sun, M., Choi, E.Y., Magee, D., Stets, C.W., and During, M.J. (2015). Social overcrowding as a chronic stress model that increases adiposity in mice. *Psychoneuroendocrinology* **57**, 318–330.
- Lischinsky, J.E., and Lin, D. (2020). Neural mechanisms of aggression across species. *Nat. Neurosci.* **23**, 1317–1328.
- Lister, R.G., and Hilakivi, L.A. (1988). The effects of novelty, isolation, light and ethanol on the social behavior of mice. *Psychopharmacology (Berl.)* **96**, 181–187.
- Liu, W., Liang, X., Gong, J., Yang, Z., Zhang, Y.-H., Zhang, J.-X., and Rao, Y. (2011). Social regulation of aggression by pheromonal activation of Or65a olfactory neurons in *Drosophila*. *Nat. Neurosci.* **14**, 896–902.
- Liu, Y., Du, S., Lv, L., Lei, B., Shi, W., Tang, Y., Wang, L., and Zhong, Y. (2016). Hippocampal Activation of Rac1 Regulates the Forgetting of Object Recognition Memory. *Curr. Biol.* **26**, 2351–2357.
- Liu, Y., Lv, L., Wang, L., and Zhong, Y. (2018). Social Isolation Induces Rac1-Dependent Forgetting of Social Memory. *Cell Rep.* **25**, 288–295.e3.
- Loo, C., and Ong, P. (1984). Crowding Perceptions, Attitudes, and Consequences among the Chinese. *Environ. Behav.* **16**, 55–87.
- Ma, X.C., Jiang, D., Jiang, W.H., Wang, F., Jia, M., Wu, J., Hashimoto, K., Dang, Y.H., and Gao, C.G. (2011). Social isolation-induced aggression potentiates anxiety and depressive-like behavior in male mice subjected to unpredictable chronic mild stress. *PLoS ONE* **6**, e20955.
- Malkesman, O., Maayan, R., Weizman, A., and Weller, A. (2006). Aggressive behavior and HPA axis hormones after social isolation in adult rats of two different genetic animal models for depression. *Behav. Brain Res.* **175**, 408–414.
- Manduca, A., Servadio, M., Damsteegt, R., Campolongo, P., Vanderschuren, L.J., and Trezza, V. (2016). Dopaminergic Neurotransmission in the Nucleus Accumbens Modulates Social Play Behavior in Rats. *Neuropsychopharmacology* **41**, 2215–2223.
- Marlin, B.J., Mitre, M., D'amour, J.A., Chao, M.V., and Froemke, R.C. (2015). Oxytocin enables maternal behaviour by balancing cortical inhibition. *Nature* **520**, 499–504.
- Maslow, A.H. (1943a). A Theory of Human Motivation. *Psychol. Rev.* **50**, 370–396.
- Maslow, A.H. (1943b). Preface to motivation theory. *Psychosom. Med.* **5**, 85–92.
- Matsumoto, K., Pinna, G., Puia, G., Guidotti, A., and Costa, E. (2005). Social isolation stress-induced aggression in mice: a model to study the pharmacology of neurosteroidogenesis. *Stress* **8**, 85–93.
- Matthews, G.A., and Tye, K.M. (2019). Neural mechanisms of social homeostasis. *Ann. N Y Acad. Sci.* **1457**, 5–25.
- Matthews, G.A., Nieh, E.H., Vander Weele, C.M., Halbert, S.A., Pradhan, R.V., Yosafat, A.S., Gieber, G.F., Izadmehr, E.M., Thomas, R.E., Lacy, G.D., et al. (2016). Dorsal Raphe Dopamine Neurons Represent the Experience of Social Isolation. *Cell* **164**, 617–631.
- McEwen, B.S. (1998). Protective and damaging effects of stress mediators. *N. Engl. J. Med.* **338**, 171–179.
- McNeal, N., Scotti, M.-A.L., Wardwell, J., Chandler, D.L., Bates, S.L., Larocca, M., Trahanas, D.M., and Grippo, A.J. (2014). Disruption of social bonds

- induces behavioral and physiological dysregulation in male and female prairie voles. *Auton. Neurosci.* 180, 9–16.
- Meira, T., Leroy, F., Buss, E.W., Oliva, A., Park, J., and Siegelbaum, S.A. (2018). A hippocampal circuit linking dorsal CA2 to ventral CA1 critical for social memory dynamics. *Nat. Commun.* 9, 4163.
- Miller, C.T., Freiwald, W.A., Leopold, D.A., Mitchell, J.F., Silva, A.C., and Wang, X. (2016). Marmosets: A Neuroscientific Model of Human Social Behavior. *Neuron* 90, 219–233.
- Mizuno, M., Inoue, T., Lukáts, B., Sakai, K., Sasaki, A., and Aou, S. (2007). Behavioral analysis of visual discrimination of sex in female macaque monkeys. *Int. Congr. Ser.* 1301, 222–225.
- Moy, S.S., Nadler, J.J., Perez, A., Barbaro, R.P., Johns, J.M., Magnuson, T.R., Piven, J., and Crawley, J.N. (2004). Sociability and preference for social novelty in five inbred strains: an approach to assess autistic-like behavior in mice. *Genes Brain Behav.* 3, 287–302.
- Moynihan, M.H. (1970). Some Behavior Patterns of Platyrrhine Monkeys: II. *Saguinus Geoffroyi* and Some Other Tamarins, *Volume 2* (Smithsonian Institution Press).
- Munuera, J., Rigotti, M., and Salzman, C.D. (2018). Shared neural coding for social hierarchy and reward value in primate amygdala. *Nat. Neurosci.* 21, 415–423.
- Nardou, R., Lewis, E.M., Rothhaas, R., Xu, R., Yang, A., Boyden, E., and Dölen, G. (2019). Oxytocin-dependent reopening of a social reward learning critical period with MDMA. *Nature* 569, 116–120.
- Niesink, R.J., and van Ree, J.M. (1982). Short-term isolation increases social interactions of male rats: a parametric analysis. *Physiol. Behav.* 29, 819–825.
- Nonogaki, K., Nozue, K., and Oka, Y. (2007). Social isolation affects the development of obesity and type 2 diabetes in mice. *Endocrinology* 148, 4658–4666.
- Norcross, J.L., and Newman, J.D. (1993). Context and gender-specific differences in the acoustic structure of common marmoset (*Callithrix jacchus*) phee calls. *Am. J. Primatol.* 30, 37–54.
- Nummela, S.U., Jovanovic, V., de la Mothe, L., and Miller, C.T. (2017). Social Context-Dependent Activity in Marmoset Frontal Cortex Populations during Natural Conversations. *J. Neurosci.* 37, 7036–7047.
- Oka, Y., Ye, M., and Zuker, C.S. (2015). Thirst driving and suppressing signals encoded by distinct neural populations in the brain. *Nature* 520, 349–352.
- Okuyama, T., Kitamura, T., Roy, D.S., Itohara, S., and Tonegawa, S. (2016). Ventral CA1 neurons store social memory. *Science* 353, 1536–1541.
- Padilla-Coreano, N., Batra, K., Patarino, M., Chen, Z., Rock, R., Zhang, R., Hausmann, S., Weddington, J., Patel, R., Zhang, Y., et al. (2020). A cortical-hypothalamic circuit decodes social rank and promotes dominance behavior (In Review) <https://doi.org/10.21203/rs.3.rs-94115/v1>.
- Pagani, J.H., Zhao, M., Cui, Z., Avram, S.K., Caruana, D.A., Dudek, S.M., and Young, W.S. (2015). Role of the vasopressin 1b receptor in rodent aggressive behavior and synaptic plasticity in hippocampal area CA2. *Mol. Psychiatry* 20, 490–499.
- Panksepp, J., and Beatty, W.W. (1980). Social deprivation and play in rats. *Behav. Neural Biol.* 30, 197–206.
- Panksepp, J., Burgdorf, J., Beinfeld, M.C., Kroes, R.A., and Moskal, J.R. (2004). Regional brain cholecystokinin changes as a function of friendly and aggressive social interactions in rats. *Brain Res.* 1025, 75–84.
- Pardo-Bellver, C., Cádiz-Moretti, B., Novejarque, A., Martínez-García, F., and Lanuza, E. (2012). Differential efferent projections of the anterior, posterovenral, and posterodorsal subdivisions of the medial amygdala in mice. *Front. Neuroanat.* 6, 33.
- Pearson, B.L., Reeder, D.M., and Judge, P.G. (2015). Crowding increases salivary cortisol but not self-directed behavior in captive baboons. *Am. J. Primatol.* 77, 462–467.
- Peng, X., Lang, C.M., Drozdowicz, C.K., and Ohlsson-Wilhelm, B.M. (1989). Effect of cage population density on plasma corticosterone and peripheral lymphocyte populations of laboratory mice. *Lab. Anim.* 23, 302–306.
- Perissinotto, C.M., Stijacic Cenzer, I., and Covinsky, K.E. (2012). Loneliness in older persons: a predictor of functional decline and death. *Arch. Intern. Med.* 172, 1078–1083.
- Pickett, C.L., Gardner, W.L., and Knowles, M. (2004). Getting a cue: the need to belong and enhanced sensitivity to social cues. *Pers. Soc. Psychol. Bull.* 30, 1095–1107.
- Popova, N.K., and Naumenko, E.V. (1972). Dominance relations and the pituitary-adrenal system in rats. *Anim. Behav.* 20, 108–111.
- Pressman, S.D., Cohen, S., Miller, G.E., Barkin, A., Rabin, B.S., and Treanor, J.J. (2005). “Loneliness, social network size, and immune response to influenza vaccination in college freshman”: Correction to Pressman et al. (2005). *Health Psychol.* 24, 348.
- Ramot, A., Jiang, Z., Tian, J.B., Nahum, T., Kuperman, Y., Justice, N., and Chen, A. (2017). Hypothalamic CRFR1 is essential for HPA axis regulation following chronic stress. *Nat. Neurosci.* 20, 385–388.
- Regoeczi, W.C. (2003). When context matters: a multilevel analysis of household and neighbourhood crowding on aggression and withdrawal. *J. Environ. Psychol.* 23, 457–470.
- Rehfeld, J.F. (2000). Cholecystokinin and panic disorder—three unsettled questions. *Regul. Pept.* 93, 79–83.
- Resendez, S.L., Namboodiri, V.M.K., Otis, J.M., Eckman, L.E.H., Rodriguez-Romaguera, J., Ung, R.L., Basiri, M.L., Kosyk, O., Rossi, M.A., Dichter, G.S., and Stuber, G.D. (2020). Social Stimuli Induce Activation of Oxytocin Neurons Within the Paraventricular Nucleus of the Hypothalamus to Promote Social Behavior in Male Mice. *J. Neurosci.* 40, 2282–2295.
- Roberts, S.A., Simpson, D.M., Armstrong, S.D., Davidson, A.J., Robertson, D.H., McLean, L., Beynon, R.J., and Hurst, J.L. (2010). Darcin: a male pheromone that stimulates female memory and sexual attraction to an individual male’s odour. *BMC Biol.* 8, 75.
- Robinson, D.L., Heien, M.L.A.V., and Wightman, R.M. (2002). Frequency of dopamine concentration transients increases in dorsal and ventral striatum of male rats during introduction of conspecifics. *J. Neurosci.* 22, 10477–10486.
- Roeling, T.A.P., Veening, J.G., Peters, J.P.W., Vermelis, M.E.J., and Nieuwenhuys, R. (1993). Efferent connections of the hypothalamic “grooming area” in the rat. *Neuroscience* 56, 199–225.
- Roeling, T.A.P., Veening, J.G., Kruk, M.R., Peters, J.P.W., Vermelis, M.E.J., and Nieuwenhuys, R. (1994). Efferent connections of the hypothalamic “aggression area” in the rat. *Neuroscience* 59, 1001–1024.
- Rukstalis, M., and French, J.A. (2005). Vocal buffering of the stress response: exposure to conspecific vocalizations moderates urinary cortisol excretion in isolated marmosets. *Horm. Behav.* 47, 1–7.
- Saavedra, J.M., Armando, I., Bregonzio, C., Juorio, A., Macova, M., Pavel, J., and Sanchez-Lemus, E. (2006). A centrally acting, anxiolytic angiotensin II AT1 receptor antagonist prevents the isolation stress-induced decrease in cortical CRF1 receptor and benzodiazepine binding. *Neuropsychopharmacology* 31, 1123–1134.
- Sandi, C., and Haller, J. (2015). Stress and the social brain: behavioural effects and neurobiological mechanisms. *Nature Reviews Neuroscience* 16, 290–304.
- Sapolsky, R.M. (1990). A. E. Bennett Award paper. Adrenocortical function, social rank, and personality among wild baboons. *Biol. Psychiatry* 28, 862–878.
- Sapolsky, R.M. (2005). The influence of social hierarchy on primate health. *Science* 308, 648–652.
- Sara, S.J., and Bouret, S. (2012). Orienting and reorienting: the locus coeruleus mediates cognition through arousal. *Neuron* 76, 130–141.
- Schiavo, J.K., Valtcheva, S., Bair-Marshall, C.J., Song, S.C., Martin, K.A., and Froemke, R.C. (2020). Innate and plastic mechanisms for maternal behaviour in auditory cortex. *Nature* 587, 426–431.
- Schjelderup-Ebbe, T. (1922). Beiträge zur Sozialpsychologie des Haushuhns. [Observation on the social psychology of domestic fowls.] *Zeitschrift Für Psychologie Und Physiologie Der Sinnesorgane. Abt. 1. Z. Psychol. Z. Angew. Psychol.* 88, 225–252.

- Sciolino, N.R., Bortolato, M., Eisenstein, S.A., Fu, J., Oveysi, F., Hohmann, A.G., and Piomelli, D. (2010). Social isolation and chronic handling alter endocannabinoid signaling and behavioral reactivity to context in adult rats. *Neuroscience* 168, 371–386.
- Seeley, T.D. (2010). *Honeybee Democracy* (Princeton University Press).
- Senst, L., Baimoukhametova, D., Sterley, T.-L., and Bains, J.S. (2016). Sexually dimorphic neuronal responses to social isolation. *eLife* 5, e18726.
- Shamay-Tsoory, S.G., and Abu-Akel, A. (2016). The Social Salience Hypothesis of Oxytocin. *Biol. Psychiatry* 79, 194–202.
- Shemesh, Y., Sztainberg, Y., Forkosh, O., Shlapobersky, T., Chen, A., and Schneidman, E. (2013). High-order social interactions in groups of mice. *eLife* 2, e00759.
- Shemesh, Y., Forkosh, O., Mahn, M., Anpilov, S., Sztainberg, Y., Manashirov, S., Shlapobersky, T., Elliott, E., Tabouy, L., Ezra, G., et al. (2016). Ucn3 and CRF-R2 in the medial amygdala regulate complex social dynamics. *Nat. Neurosci.* 19, 1489–1496.
- Shin, S., Pribiag, H., Lilascharoen, V., Knowland, D., Wang, X.-Y., and Lim, B.K. (2018). Drd3 Signaling in the Lateral Septum Mediates Early Life Stress-Induced Social Dysfunction. *Neuron* 97, 195–208.e6.
- Shuai, Y., Lu, B., Hu, Y., Wang, L., Sun, K., and Zhong, Y. (2010). Forgetting is regulated through Rac activity in *Drosophila*. *Cell* 140, 579–589.
- Singh, L., Lewis, A.S., Field, M.J., Hughes, J., and Woodruff, G.N. (1991). Evidence for an involvement of the brain cholecystokinin B receptor in anxiety. *Proc. Natl. Acad. Sci. USA* 88, 1130–1133.
- Smith, A.S., Williams Avram, S.K., Cymerblit-Sabba, A., Song, J., and Young, W.S. (2016). Targeted activation of the hippocampal CA2 area strongly enhances social memory. *Mol. Psychiatry* 21, 1137–1144.
- So, N., Franks, B., Lim, S., and Curley, J.P. (2015). A Social Network Approach Reveals Associations between Mouse Social Dominance and Brain Gene Expression. *PLoS ONE* 10, e0134509.
- Spée, M., Beaulieu, M., Dervaux, A., Chastel, O., Le Maho, Y., and Raclot, T. (2010). Should I stay or should I go? Hormonal control of nest abandonment in a long-lived bird, the Adélie penguin. *Horm. Behav.* 58, 762–768.
- Spreng, R.N., Dimas, E., Mwilambwe-Tshilobo, L., Dagher, A., Koellinger, P., Nave, G., Ong, A., Kernbach, J.M., Wiecki, T.V., Ge, T., et al. (2020). The default network of the human brain is associated with perceived social isolation. *Nat. Commun.* 11, 6393.
- Stagkourakis, S., Spigolon, G., Williams, P., Protzmann, J., Fisone, G., and Broberger, C. (2018). A neural network for intermale aggression to establish social hierarchy. *Nat. Neurosci.* 21, 834–842.
- Steptoe, A., Shankar, A., Demakakos, P., and Wardle, J. (2013). Social isolation, loneliness, and all-cause mortality in older men and women. *Proc. Natl. Acad. Sci. USA* 110, 5797–5801.
- Stenson, S.M. (2013). Hypothalamic survival circuits: blueprints for purposive behaviors. *Neuron* 77, 810–824.
- Stowers, L., and Kuo, T.-H. (2015). Mammalian pheromones: emerging properties and mechanisms of detection. *Curr. Opin. Neurobiol.* 34, 103–109.
- Stowers, L., Holy, T.E., Meister, M., Dulac, C., and Koentges, G. (2002). Loss of sex discrimination and male-male aggression in mice deficient for TRP2. *Science* 295, 1493–1500.
- Sun, P., Smith, A.S., Lei, K., Liu, Y., and Wang, Z. (2014). Breaking bonds in male prairie vole: long-term effects on emotional and social behavior, physiology, and neurochemistry. *Behav. Brain Res.* 265, 22–31.
- Takatsu-Coleman, A.L., Patti, C.L., Zanin, K.A., Zager, A., Carvalho, R.C., Borçoi, A.R., Ceccon, L.M.B., Berro, L.F., Tufik, S., Andersen, M.L., and Frussa-Filho, R. (2013). Short-term social isolation induces depressive-like behaviour and reinstates the retrieval of an aversive task: mood-congruent memory in male mice? *J. Psychiatry Neurosci.* 38, 259–268.
- Tan, C.L., Cooke, E.K., Leib, D.E., Lin, Y.-C., Daly, G.E., Zimmerman, C.A., and Knight, Z.A. (2016). Warm-Sensitive Neurons that Control Body Temperature. *Cell* 167, 47–59.e15.
- Taylor, J.H., Mustoe, A.C., and French, J.A. (2014). Behavioral responses to social separation stressor change across development and are dynamically related to HPA activity in marmosets. *Am. J. Primatol.* 76, 239–248.
- Thoening, R.H. (1972). Solitary Confinement - Punishment within the Letter of the Law, or Psychological Torture. *Wis. L. Rev.* 1972, 223.
- Tomova, L., Tye, K., and Saxe, R. (2019). The neuroscience of unmet social needs. *Soc. Neurosci.*, 10.1080/17470919.2019.1694580.
- Tomova, L., Wang, K., Thompson, T., Matthews, G., Takahashi, A., Tye, K., and Saxe, R. (2020a). The need to connect: Acute social isolation causes neural craving responses similar to hunger. *BioRxiv*. <https://doi.org/10.1101/2020.03.25.006643>.
- Tomova, L., Wang, K.L., Thompson, T., Matthews, G.A., Takahashi, A., Tye, K.M., and Saxe, R. (2020b). The need to connect: Acute social isolation evokes midbrain craving responses similar to hunger. *Nat. Neurosci.* 23, 1597–1605.
- Tsao, D.Y., Freiwald, W.A., Tootell, R.B.H., and Livingstone, M.S. (2006). A cortical region consisting entirely of face-selective cells. *Science* 311, 670–674.
- Tye, K.M. (2018). Neural Circuit Motifs in Valence Processing. *Neuron* 100, 436–452.
- Van Loo, P.L., Mol, J.A., Koolhaas, J.M., Van Zutphen, B.F., and Baumans, V. (2001). Modulation of aggression in male mice: influence of group size and cage size. *Physiol. Behav.* 72, 675–683.
- Wallace, D.L., Han, M.-H., Graham, D.L., Green, T.A., Vialou, V., Iñiguez, S.D., Cao, J.-L., Kirk, A., Chakravarty, S., Kumar, A., et al. (2009). CREB regulation of nucleus accumbens excitability mediates social isolation-induced behavioral deficits. *Nat. Neurosci.* 12, 200–209.
- Walsh, J.J., Christoffel, D.J., Heifets, B.D., Ben-Dor, G.A., Selimbeyoglu, A., Hung, L.W., Deisseroth, K., and Malenka, R.C. (2018). 5-HT release in nucleus accumbens rescues social deficits in mouse autism model. *Nature* 560, 589–594.
- Wang, L., Dankert, H., Perona, P., and Anderson, D.J. (2008). A common genetic target for environmental and heritable influences on aggressiveness in *Drosophila*. *Proc. Natl. Acad. Sci. USA* 105, 5657–5663.
- Wang, F., Zhu, J., Zhu, H., Zhang, Q., Lin, Z., and Hu, H. (2011). Bidirectional control of social hierarchy by synaptic efficacy in medial prefrontal cortex. *Science* 334, 693–697.
- Weiss, R.S. (1973). *Loneliness: The Experience of Emotional and Social Isolation* (The MIT Press).
- Wendt, D., van Loon, L.J.C., and Lichtenbelt, W.D. van M. (2007). Thermoregulation during exercise in the heat: strategies for maintaining health and performance. *Sports Med.* 37, 669–682.
- Wiberg, G.S., and Grice, H.C. (1963). Long-term isolation stress in rats. *Science* 142, 507.
- Wichmann, R., Vander Weele, C.M., Yosafat, A.S., Schut, E.H.S., Verharen, J.P.H., Sridharma, S., Siciliano, C.A., Izadmehr, E.M., Farris, K.M., Wildes, C.P., et al. (2017). Acute stress induces long-lasting alterations in the dopaminergic system of female mice. *BioRxiv* 10.1101/168492.
- Williamson, C.M., Romeo, R.D., and Curley, J.P. (2017). Dynamic changes in social dominance and mPOA GnRH expression in male mice following social opportunity. *Horm. Behav.* 87, 80–88.
- Wongwitdech, N., and Marsden, C.A. (1996). Social isolation increases aggressive behaviour and alters the effects of diazepam in the rat social interaction test. *Behav. Brain Res.* 75, 27–32.
- Wu, X.-R., Zhang, Y., Liu, X.-D., Han, W.-B., Xu, N.-J., and Sun, S. (2020). EphB2 mediates social isolation-induced memory forgetting. *Transl. Psychiatry* 10, 389.
- Yamaguchi, T., Wei, D., Song, S.C., Lim, B., Tritsch, N.X., and Lin, D. (2020). Posterior amygdala regulates sexual and aggressive behaviors in male mice. *Nat. Neurosci.* 23, 1111–1124.



- Yao, S., Bergan, J., Lanjuin, A., and Dulac, C. (2017). Oxytocin signaling in the medial amygdala is required for sex discrimination of social cues. *eLife* 6, e31373.
- Young, W.S., Li, J., Wersinger, S.R., and Palkovits, M. (2006). The vasopressin 1b receptor is prominent in the hippocampal area CA2 where it is unaffected by restraint stress or adrenalectomy. *Neuroscience* 143, 1031–1039.
- Zelikowsky, M., Hui, M., Karigo, T., Choe, A., Yang, B., Blanco, M.R., Beadle, K., Gradinaru, V., Deverman, B.E., and Anderson, D.J. (2018). The Neuropeptide Tac2 Controls a Distributed Brain State Induced by Chronic Social Isolation Stress. *Cell* 173, 1265–1279.e19.
- Zhou, T., Zhu, H., Fan, Z., Wang, F., Chen, Y., Liang, H., Yang, Z., Zhang, L., Lin, L., Zhan, Y., et al. (2017). History of winning remodels thalamo-PFC circuit to reinforce social dominance. *Science* 357, 162–168.
- Zimmerman, C.A., Lin, Y.-C., Leib, D.E., Guo, L., Huey, E.L., Daly, G.E., Chen, Y., and Knight, Z.A. (2016). Thirst neurons anticipate the homeostatic consequences of eating and drinking. *Nature* 537, 680–684, advance online publication.
- Zimmerman, C.A., Huey, E.L., Ahn, J.S., Beutler, L.R., Tan, C.L., Kosar, S., Bai, L., Chen, Y., Corpuz, T.V., Madisen, L., et al. (2019). A gut-to-brain signal of fluid osmolarity controls thirst satiation. *Nature* 568, 98–102.

**Update**

**Cell**

Volume 184, Issue 10, 13 May 2021, Page 2794–2795

DOI: <https://doi.org/10.1016/j.cell.2021.04.044>

Correction

# The neural circuitry of social homeostasis: Consequences of acute versus chronic social isolation

Christopher R. Lee, Alon Chen, and Kay M. Tye\*

\*Correspondence: [tye@salk.edu](mailto:tye@salk.edu)

<https://doi.org/10.1016/j.cell.2021.04.044>

(Cell 184, 1500–1516; March 18, 2021)

In Figure 3 of this review as originally published, the arrow between VMHvl and PA<sup>Esr1</sup> should have been oriented in the opposite direction: PA<sup>Esr1</sup> neurons project to VMHvl and are active in intermale aggression. Additionally, the main text originally describes this neural circuit in the reverse direction (VMHvl → PA<sup>Esr1</sup>) and has now been corrected. The authors apologize for these errors and any confusion that may have resulted.

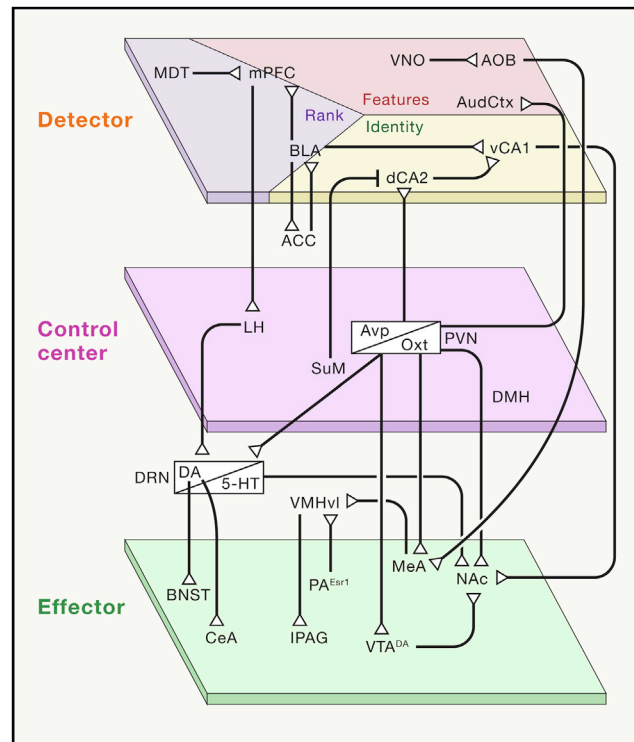


Figure 3. Neural circuits underlying homeostatic nodes (corrected)



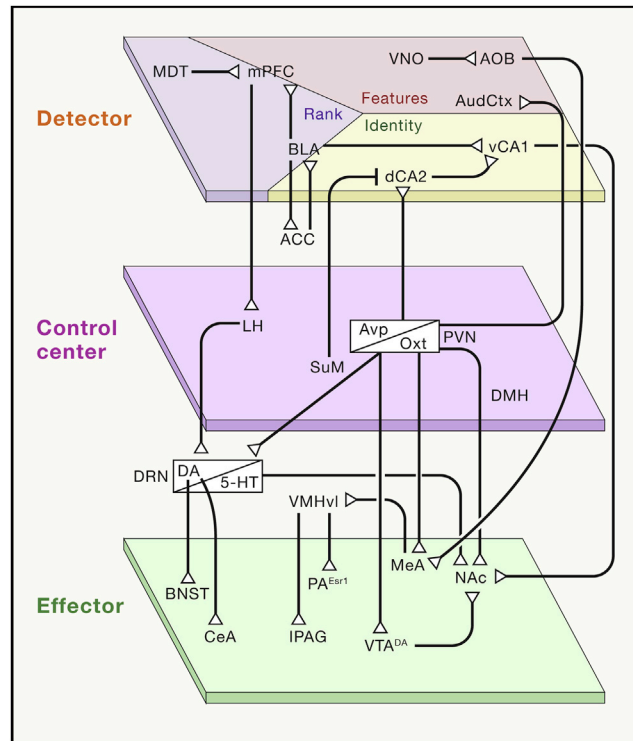


Figure 3. Neural circuits underlying homeostatic nodes (original)





## OPEN ACCESS

## EDITED BY

Karri P. Lamsa,  
University of Szeged, Hungary

## REVIEWED BY

Irmgard Tegeger,  
Goethe University Frankfurt, Germany  
Emilio J. Galván,  
Instituto Politécnico Nacional de México  
(CINVESTAV), Mexico  
Michael M. Kohl,  
University of Glasgow, United Kingdom  
Nelson Rebola,  
INSERM U1127 Institut du Cerveau et de la  
Moelle épinière (ICM), France

## \*CORRESPONDENCE

Charlotte Piette  
✉ charlotte.piette@college-de-france.fr  
Laurent Venance  
✉ laurent.venance@college-de-france.fr

RECEIVED 30 June 2023

ACCEPTED 20 November 2023

PUBLISHED 07 December 2023

## CITATION

Piette C, Gervasi N and Venance L (2023)  
Synaptic plasticity through a naturalistic lens.  
*Front. Synaptic Neurosci.* 15:1250753.  
doi: 10.3389/fnsyn.2023.1250753

## COPYRIGHT

© 2023 Piette, Gervasi and Venance. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Synaptic plasticity through a naturalistic lens

Charlotte Piette\*, Nicolas Gervasi and Laurent Venance\*

Center for Interdisciplinary Research in Biology (CIRB), College de France, CNRS, INSERM, Université PSL, Paris, France

From the myriad of studies on neuronal plasticity, investigating its underlying molecular mechanisms up to its behavioral relevance, a very complex landscape has emerged. Recent efforts have been achieved toward more naturalistic investigations as an attempt to better capture the synaptic plasticity underpinning of learning and memory, which has been fostered by the development of *in vivo* electrophysiological and imaging tools. In this review, we examine these naturalistic investigations, by devoting a first part to synaptic plasticity rules issued from naturalistic *in vivo*-like activity patterns. We next give an overview of the novel tools, which enable an increased spatio-temporal specificity for detecting and manipulating plasticity expressed at individual spines up to neuronal circuit level during behavior. Finally, we put particular emphasis on works considering brain-body communication loops and macroscale contributors to synaptic plasticity, such as body internal states and brain energy metabolism.

## KEYWORDS

synaptic plasticity, *in vivo*-like patterns, neuromodulation, learning, memory, spike-timing dependent plasticity, neuroenergetic, body internal states

## Background

There is an increasing body of evidence in favor of the neuronal plasticity (synaptic, intrinsic and/or structural) and memory hypothesis (Martin and Morris, 2002; Josselyn and Tonegawa, 2020). Synaptic plasticity rules were first investigated *in vitro*, in which neuronal activity patterns can be exactly controlled. Although some rules have been validated *in vivo*, further clarification is needed on how *in vivo* neuronal activity causes synaptic plasticity. Furthermore, the diversity of plasticity rules and profiles expressed within an individual neuron or a given circuit, the plasticityome (McFarlan et al., 2023), calls for a clearer understanding of their specific functions and also of their interplay during learning. In the light of recent studies, this review aims at highlighting how naturalistic investigations of synaptic plasticity can provide a critical insight into the plasticity and memory research field.

## From neuronal activity patterns to plasticity rules

Multiple plasticity induction protocols, more or less inspired by *in vivo* activity patterns, have been used both *in vitro* and *in vivo* to unveil the spatio-temporal constraints of synaptic plasticity expression in neuronal networks and dissect their molecular determinants. High-frequency stimulation (HFS) is still widely used because it induces reliable and (generally) potent plasticity (long-term potentiation, LTP, or long-term depression, LTD) LTP or LTD depending on brain areas, neuronal subtypes). Although HFS can somehow mimic some sensory epochs, HFS appears in most conditions rather as an artificial cell conditioning paradigm because of its high and regular stimulation frequency (100 Hz), and duration (typically

1 s repeated several times). Yet, it is crucial to study the effects of stimulation protocols using natural activity patterns, obtained from *in vivo* electrophysiological recordings (Paulsen and Sejnowski, 2000). Indeed, it will inform on which activity patterns are sufficient and effective at inducing plasticity during learning *in vivo*, hence uncovering naturalistic plasticity rules. In addition, it enables to identify molecular determinants (partially different from those recruited by HFS-induced plasticity), that could later be used for manipulating plasticity expression *in vivo*.

First attempts of realistic stimulations came with theta-burst stimulations, determined from *in vivo* recordings of place cells showing theta rhythm linked to memory storage. Later on, *in vivo* recordings of cortical neurons displaying low frequency firing (<5 Hz) and the discovery of backpropagating action potential (bAP), a signal that could bind presynaptic and postsynaptic activity for plasticity induction, led to spike-timing dependent plasticity (STDP) paradigms (Feldman, 2012; Debanne and Inglebert, 2023). Temporally ordered coincident neuronal activity was postulated by Donald Hebb as the critical driver of long-lasting modifications between neurons (Sejnowski, 1999). Its experimental validation came with the discovery that repeated presynaptic activity preceding post-synaptic activity within a few tens of milliseconds could induce LTP, while the converse temporal order led to LTD; aka Hebbian STDP. Since then, multiple polarity and forms of STDP have been described (Feldman, 2012). Classically, STDP is induced with 100–150 presynaptic and postsynaptic pairings at low frequency (1–2 Hz). However, various forms of STDP aiming at mimicking more *in vivo*-like activity (Debanne and Inglebert, 2023) were also evoked using smaller number of pairings (5–30) (Froemke et al., 2006; Cui et al., 2015, 2016; Cepeda-Prado et al., 2022) as expected in single-trial or one-shot learning (Piette et al., 2020), more complex spiking sequences, such as spike triplets or quadruplets (Froemke and Dan, 2002; Mendes et al., 2020), or *in vivo* spiking patterns replayed between two neighboring neurons *in vitro* (Isaac et al., 2009). STDP was also translated *in vivo* by associating natural sensory stimulation that activates afferents combined with evoked or spontaneous spiking of a single cortical neuron (Yao and Dan, 2001; Meliza and Dan, 2006; Jacob et al., 2007).

STDP rules usually rely on the repetition of precisely timed presynaptic and postsynaptic spikes *in vitro*. Yet the exact contribution of spike timing relative to firing rate in eliciting synaptic plasticity *in vivo* is still debated (Graupner et al., 2016). Related to this, it was shown that plasticity rules vary across cerebellar regions, with a precise time interval of 120 ms between parallel fiber and climbing fiber inputs allowing for plasticity expression in the flocculus (known to receive error signals at this delay during oculomotor learning), while a broader range of intervals are permissive for plasticity expression in the vermis, implicated in a wider variety of learning paradigms (Suvrathan et al., 2016). In addition, cortico-striatal STDPs show different sensitivity to spike timing jitter *in vitro*: endocannabinoid-dependent plasticities (endocannabinoid-LTD and endocannabinoid-LTP) are more robust to spike timing variability compared to NMDA-LTP (Cui et al., 2018). Such differential sensitivities between endocannabinoid-LTP and NMDA-LTP

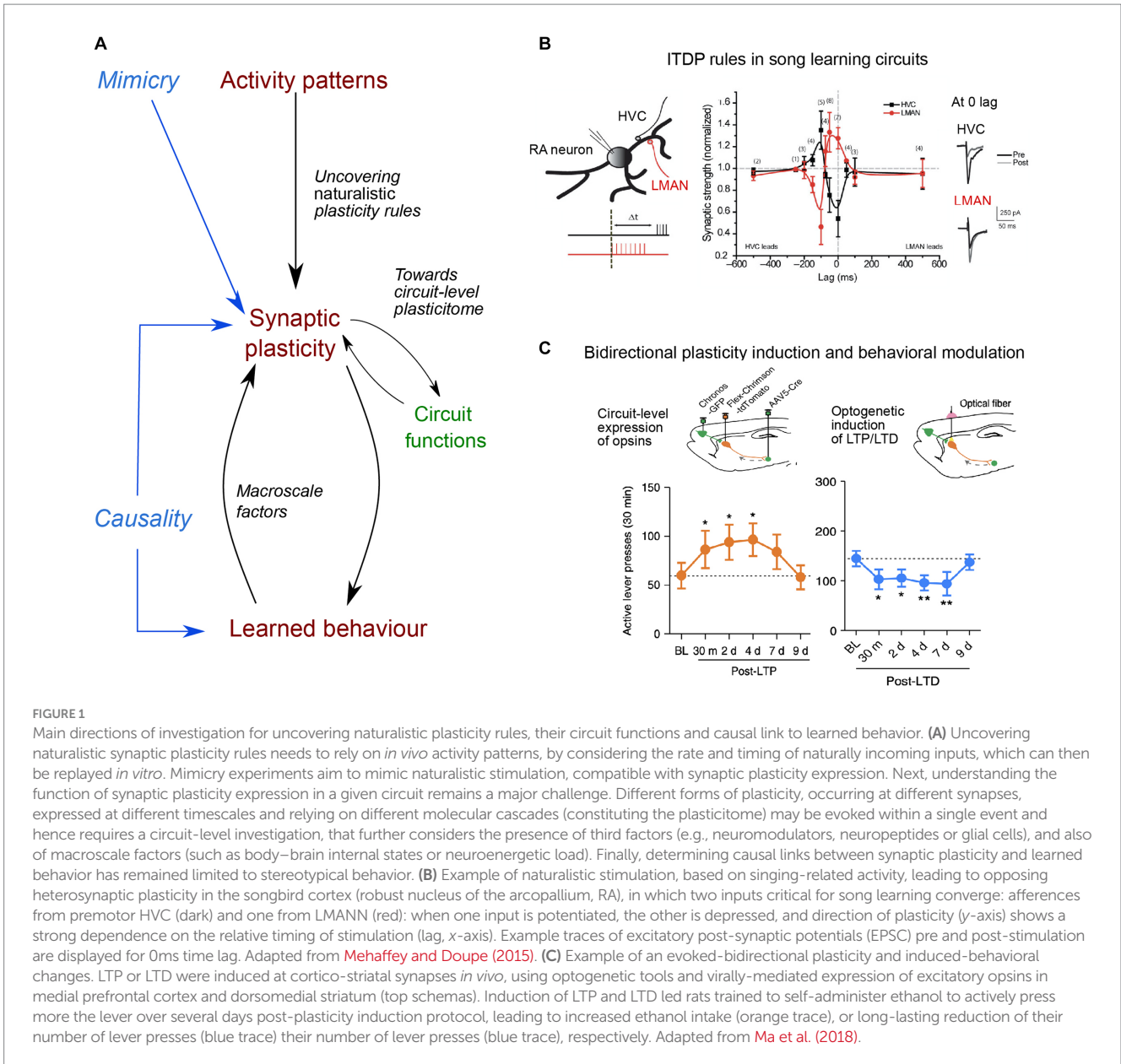
(induced by few pairings, 5–15, versus 100 pairings, respectively) could underlie different functions, at different stages of learning as sequential activity patterns become more and more stereotyped (Thorn et al., 2010; Figure 1).

Furthermore, STDP faces two major interrogations questioning its physiological relevance: (i) Can plasticity only be induced by a global feedback signal, such as a bAP, resulting from the activation of a critical number of synapses? This would contradict the fact that a limited number of synapses could be subjected to plasticity. Also, this implies that only digital (bAP) but not analog signals induce plasticity; (ii) How can the compressed timescale (typically, pairings intervals are <50 milliseconds and repeated at 1 Hz) in which STDP occurs be compatible with behavioral timescales (subseconds to minutes)? Indeed, if STDP is sensitive to correlations of tens or hundreds of milliseconds (Feldman, 2012), temporal associations between different stimuli during learning are typically in the subsecond/minute range (Drew and Abbott, 2006).

The first point has been addressed in different ways. Paired subthreshold events can induce long-term plasticity, such that bAP would not be necessary for plasticity expression (Fino et al., 2009; Brandalise and Gerber, 2014) and strong post-synaptic depolarization was shown to induce sufficient intracellular  $Ca^{2+}$  rise, due to dendritic non-linearities, to evoke long-term plasticity (Holthoff et al., 2004; Hardie and Spruston, 2009). Therefore, digital (bAPs) as well as analog (EPSPs) signals can induce *in vitro* plasticity. Because bAPs are attenuated differently among neuronal subtypes, and along the dendritic arbor, in part depending on the level of excitatory and inhibitory synaptic inputs, their contribution is likely to vary *in vivo* (Waters and Helmchen, 2004). Hence, delimiting the exact role and spatial extension of digital and analog signals for triggering synaptic plasticity *in vivo* requires further investigation. This consideration also led to the study of more naturalistic forms of STDP, dependent upon the temporal correlation between two distinct presynaptic afferences, the input-timing-dependent plasticity (ITDP), using timing rules that mapped synaptic delays caused by neural network architecture (Dudman et al., 2007; Cho et al., 2012; Basu et al., 2013; Mehaffey and Doupe, 2015; Leroy et al., 2017; Figure 1). ITDP can be viewed as a physiological upgrade of STDP since it does not rely on an artificial post-synaptic injection of current necessary to create the bAP (see the critical view of STDP in Lisman and Spruston, 2010), but on paired activation of presynaptic inputs triggering sub- or suprathreshold activity in the postsynaptic element.

Regarding the second point, even though the compressed (milliseconds) timescale of STDP is compatible with replay episodes in sleep for learning specific temporal sequences (Debanne and Inglebert, 2023; George et al., 2023), the search for factors contributing to enlarge its temporal window put a strong emphasis on nonlinear slow-timescale mechanisms (e.g. CaMKII and calcineurin) (O'Donnell, 2023) and on the three-factor learning rule (Frémaux and Gerstner, 2016; Edelmann et al., 2017; Foncelle et al., 2018; Brzosko et al., 2019). Indeed, STDP has been originally described as a two-factor rule relying on paired activity in the presynaptic and postsynaptic elements (two-factor) to fulfill Hebb's postulate, and was later augmented to a third-factor rule to include neuromodulators (e.g. dopamine, noradrenaline or acetylcholine), neurotransmitters (GABA or endocannabinoids), neuropeptides (BDNF) or glia (astrocytes), which efficiently modulate plasticity and can enlarge the temporal window of STDP expression. Indeed, repeated co-activation

Abbreviations: bAP, backpropagating action potential; BTSP, behavioral timescale plasticity; ITDP, input-timing-dependent plasticity; HFS, high-frequency stimulation; LTP, long-term potentiation; STDP, spike-timing dependent plasticity.



of synaptic activities, together with dopamine release, leave eligibility traces for about 1 s at cortico-striatal synapses ([Yagishita et al., 2014](#)), 5 s in the neocortex ([He et al., 2015](#)) or even up to 10 min in the hippocampus ([Brzosko et al., 2015](#); [Fuchsberger et al., 2022](#)).

Lastly, a naturalistic plasticity rule was recently uncovered at CA3-CA1 synapses, both *in vitro* and *in vivo*, which no longer requires repetitions and co-activation of presynaptic and postsynaptic elements ([Bittner et al., 2017](#); [Priestley et al., 2022](#); [Fan et al., 2023](#)). Synaptic inputs from CA3 place cells are potentiated by the occurrence of a single Ca<sup>2+</sup> dendritic plateau, produced at distal dendrites. Importantly, the temporal overlap between the two signals can span the second timescale and their temporal order does not impact plasticity expression. The dendritic plateau potential appears as an instructive signal, evoked by specific circumstances (reduced dendritic inhibition, permissive neuromodulatory signaling, strong inputs), occurring for instance during exploration of a novel environment ([Priestley et al., 2022](#)) or of a context in which specific task-related information is

carried by a given position ([Zhao et al., 2022](#)). Due to behavioral timescale plasticity (BTSP) asymmetric time course ([Bittner et al., 2017](#); [Magee and Grienberger, 2020](#)), predictive information might be encoded.

This overview presented several refinements aiming at approaching naturalistic synaptic plasticity rules. This effort should be continued, especially since our understanding of synaptic plasticity rules remains mostly defined at the scale of an entire synaptic pathway and focuses on a given population of excitatory or projecting neurons. In this direction, considering interactions between neighboring neurons has unveiled a variety of heterosynaptic plasticity mechanisms ([Chistiakova et al., 2015](#); [Mendes et al., 2020](#)). These can influence the net plasticity outcome of a given circuit when interactions between inhibitory and excitatory neurons are examined ([D'Amour and Froemke, 2015](#); [Hiratani and Fukai, 2017](#)). Heterogeneities in plasticity expression at the neuronal level can arise from a neuron's prior and ongoing activity ([Han et al., 2007](#)) or its dendritic architecture, in

particular the distribution of active inhibitory and excitatory synapses (Harvey and Svoboda, 2007; El-Boustani et al., 2018) or compartmentalized changes in dendritic excitability (Losonczy et al., 2008). These additional considerations reinforce the need to detect input-specific signals from both somatic and dendritic compartments, that could serve as proxies for local synaptic plasticity expression.

This overview also pointed out the importance of reinforcing the translation between *in vivo* and *in vitro* recordings, keeping in mind discrepancies relative to ionic composition (Inglebert et al., 2020), metabolic substrates (Dembitskaya et al., 2022), neuromodulator concentrations or spontaneous activity levels, as well as body-brain internal states, which constitute key factors affecting plasticity induction thresholds, as will be discussed below.

## Toward the uncovering and manipulation of a learning-induced plasticitome

In parallel to extracting synaptic plasticity rules evoked by natural *in vivo*-like activity patterns, a vast number of studies has uncovered learning-induced synaptic changes, hence directly examining synaptic plasticity in naturalistic settings. Detection of synaptic changes historically relied on *in vivo* electrophysiological recordings of synaptic efficacy using electrical stimulation yet lacking cell-type specificity and often restricted to a single circuit, or from *ex vivo* saturation/occlusion experiments or measures of AMPA/NMDAR ratio, which cannot provide a full account of synaptic temporal dynamics. The development of optical stimulation combined with spatial- and cell-specific expression of opsins partially lifted the first limitation: as an example, cortico-striatal plasticity monitored *in vivo* during an auditory discrimination task, based on the selective optogenetic stimulation of cortical neurons along the tonotopic axis, revealed spatially selective plasticity induction depending on reward contingencies (Xiong et al., 2015). Combined with c-Fos labeling of both presynaptic and postsynaptic cells active during fear conditioning, *ex vivo* recordings showed an occlusion of LTP between engram cells, along with changes in presynaptic release probability (Choi et al., 2018). Notably, the all-optical approach combined with imaging of subthreshold membrane potential dynamics and opto-stimulation of afferences allows plasticity detection at the circuit-level (Fan et al., 2023). It could further be extended, by probing multiple regions simultaneously using novel imaging tools, such as light beads microscopy, enabling large volumetric recording of neuronal activity (Demas et al., 2021). In addition, a key advantage of optical approaches is their combinatorial power, by juxtaposing the dynamics of optical sensors and labeling of active cells during behavioral tasks. Typically, to better capture the full temporal dynamics of synaptic plasticity expression, calcium and voltage sensors, which can track initial signatures of synaptic changes, could be combined with sensors related to downstream cascades such as CamKII (Lee et al., 2009) or PKA (Gervasi et al., 2010), which inform on long-term plasticity expression and maintenance *in vivo*. Furthermore, the development of presynaptic vesicular release sensors, which are currently effective *in vitro* (Ferro et al., 2017) should also help investigating the often-neglected presynaptic plasticity loci. Although these tools offer unprecedented access to detailed naturalistic plasticitomes, they also have their own limitations and caveats. Indeed, the expression of

opsins combined with viral vectors is not without cell specificity confounds and toxicity-related issues (Miyashita et al., 2013). In addition, opsins or fluorescent sensors can alter natural synaptic dynamics, depending for instance on their expression levels (Jackman et al., 2014).

Chronic tracking of structural dynamics in spine numbers and shapes using *in vivo* 2-photon imaging (Pfeiffer et al., 2018), i.e. structural plasticity, can be used as proxy for synaptic strength (Holtmaat and Svoboda, 2009). Dual-eGRASP, a split fluorescent protein that emits fluorescence only when presynaptic and postsynaptic eGRASP components are physically attached in the synaptic cleft (Choi et al., 2018, 2021), combined with the *Fos* promoter-driven tetracycline transactivator system (Mayford and Reijmers, 2015), allows to track longitudinally *in vivo* synapses between pre and post-synaptic neurons active or not during learning (Lee et al., 2023). Other avenues, down to *in vivo* tracking of receptor dynamics, have also been opened (Matsuo et al., 2008; Zhang et al., 2015) and can be envisaged simultaneously across thousands of synapses (Graves et al., 2021).

Beyond the detailed characterization of plasticity expression *in vivo* during learning, the demonstration of a causality between plasticity expression and learned behavior now represents a current grail in neurophysiology, such that bidirectional behavioral modifications can be caused by bidirectional manipulation of synaptic efficacy. Currently, only a couple of studies has achieved such bidirectional control: a conditioned fear response was, respectively, erased and restored upon depotentiation and re-potentiation of the auditory inputs to the lateral amygdala (Nabavi et al., 2014). Likewise, cortico-striatal opto-induced-LTP and -LTD promoted and decreased, respectively, alcohol-seeking behavior (Ma et al., 2018; Figure 1). Optogenetically-induced depotentiation of LTP, initially induced by auditory fear conditioning, suppressed fear responses to the conditioned stimulus (Kim et al., 2007). Furthermore, with two auditory stimuli underlying two different memories, opto-potentiation and -depotentiation of synapses shared within each specific cell assembly selectively restored or impaired the retrieval of one memory while sparing the other (Abdou et al., 2018). These causal manipulations should now aim at triggering reversible synaptic changes using naturalistic plasticity induction protocols, instead of classical low or high-frequency stimulation. In addition, to further nail down causality at the synapse-level, a specific ChR2 expression on recently activated synapses could allow more physiological excitation, by mimicking *in vivo* occurring calcium transients, compared to full somatic activation (Gobbo et al., 2017). In the future, one could even imagine modulating bidirectionally and reversibly the excitability of individual spines or dendritic branches during learning, for instance through targeted expression of both hyperpolarizing and excitatory opsins. Yet, an ongoing issue of manipulating synaptic efficacy relates to its specificity, and the absence of interference with other synaptic mechanisms or basal neurotransmission, as well as the possibility to cause other pathological changes or evoke compensatory mechanisms. Therefore, the development of spatio-temporally precise manipulations, with sensor expression impacting the least physiological dynamics, remains of critical importance.

Finally, instead of directly manipulating synaptic weights to cause behavioral changes, mimicry experiments currently (and somehow paradoxically) offer more naturalistic settings for testing



the memory and synaptic plasticity hypothesis. Indeed, mimicry consists of artificially stimulating neuronal circuits *in vivo* (without undergoing any kind of experience) and triggering behavioral changes. The stimulation mimics putative activity patterns during a real learning experience and can therefore bridge naturalistic synaptic plasticity rules described above and their behavioral relevance. At this day, only associations between conditioned and unconditioned stimuli were mimicked. Building up on previous works in which a partial sensory experience combined with opto-stimulation created artificial memories (Josselyn and Tonegawa, 2020), an artificial memory was generated by combining patterned stimulation of olfactory glomeruli with the stimulation of distinct inputs to the ventral tegmental area that mediated either aversion or reward (Vetere et al., 2019). A next challenge will be to move from neuronal assemblies down to the synaptic level in these same simple behavioral paradigms (using tools described above; Gobbo et al., 2017), and to generalize to more complex learning, using naturalistic sequences of neuronal activation.

## The (almost) overlooked of synaptic plasticity research: the body–brain communication loops and neuroenergetics

In synaptic plasticity, besides the two “Hebbian” factors (presynaptic and postsynaptic activities), a third “neoHebbian” factor allows the stabilization and shaping of plasticity maps. This third factor gathered well-defined elements such as neurotransmitters/neuromodulators, neuropeptides, fatty acids or glial cells (reviewed in: Frémaux and Gerstner, 2016; Foncelle et al., 2018; Brzosko et al., 2019). Here, we chose to focus on macroscale factors, defined by integrated body–brain communication loops, also in relation to external states (Kanwal et al., 2021; Flavell et al., 2022). Indeed, the brain receives massive sensorimotor feedback from the body, such as heartbeat (Hsueh et al., 2023), blood pressure, respiratory rate (Folschweiller and Sauer, 2023), gastric fullness, internal temperature or visceral pain. During active behavioral states, the integration of these feedbacks engages widespread circuits. Macroscale factors also include sleep–wake, circadian or seasonal-related rhythms, as well as metabolic (thirst and feeding) states. These factors are mediated by brain–body endocrine communication, metabolic substrates and associated signaling molecules, and recruit neuromodulators, glial or immune cells. As an illustration, the general body state changes during exercise can favor plasticity expression. Weak theta-pattern stimulation of the hippocampus, which does not produce LTP in control rats, induces LTP in rats housed with a running wheel (Farmer et al., 2004). Voluntary exercise, by increasing theta oscillation and lowering LTP induction threshold, may prime the network to promote synaptic plasticity *in vitro* and *in vivo* (van Praag et al., 1999).

Secretory molecules and vesicles released by organs such as skeletal muscle, adipose tissue, liver and gut are part of the body–brain feedback (Pedersen and Febbraio, 2012). These molecules, such as FND5/irisin, adiponectin, or IL-6, cross the blood–brain barrier, (i) induce changes in neurotrophins such as BDNF or EGF-1, associated with improvements in hippocampus plasticity, spatial memory, and object recognition (Vaynman et al., 2004; Gomes da Silva et al., 2010), (ii) modulate the cerebrovasculature, allowing improvements in

energy metabolism, delivery of oxygen, nutrients, neurotrophins and other factors promoting learning and memory, and (iii) act on plasticity through the increased number, cell body size and arborization length of astrocytes, (Saur et al., 2014), impairment of microglia (Vukovic et al., 2012) and increased neurogenesis (van Praag et al., 1999).

Among the macroscale factors that could control plasticity expression, neuroenergetics has received particular attention. Synaptic activity is the most energy-consuming process in the brain (Attwell and Laughlin, 2001). Synaptic energy supply is provided on-demand (Kasischke et al., 2004; Chuquet et al., 2010; Ruminot et al., 2017) by neuronal glycolysis and/or glial-derived lactate (via the astrocyte–neuron lactate shuttle) (Magistretti and Allaman, 2018; Bonvento and Bolanos, 2021). Synaptic plasticity and learned behavior depend on the metabolic reservoir. This is well illustrated under food restriction, in flies, with a trade-off between long-term memory establishment and survival (Mery and Kawecki, 2005; Plaçais and Preat, 2013). Long-term neuronal reconfigurations, leading to behavioral changes, mediated by feeding state have also been identified in *C. elegans* (Takeishi et al., 2020) and in the Etruscan shrew (Ray et al., 2020). In addition, top-down adaptations have also been identified between synaptic plasticity induction and energy uptake mechanisms following learning, with vascular adaptations (Lacoste et al., 2014), glial recruitment (Genoud et al., 2006), altered expression of insulin-sensitive glucose or lactate transporters (Tadi et al., 2015; Ashrafi et al., 2017) and mitochondrial activity (Todorova and Blokland, 2017). Interestingly, flies increase their energy intake following multiple trials training, leading to a dopamine-mediated upregulation of cellular metabolism driving LTP in the brain region involved in long-term memory (Plaçais et al., 2017). Dissecting precisely how the energy available and the nature of metabolic substrates – especially lactate and/or glucose – can control synaptic plasticity expression has been the subject of several works (Newman et al., 2011; Suzuki et al., 2011; Murphy-Royal et al., 2020); of note, lactate is also a signaling metabolite and as such acts on neuronal excitability and plasticity via NMDA receptors and/or hydroxycarboxylic acid receptor type-1 (HCAR1) (Magistretti and Allaman, 2018). In particular, during learning, the neuronal computational load at play may be more or less intense, raising the possibility that plasticity induction could be more or less metabolically demanding and therefore requires different metabolic pathways. Indeed, lactate supply was required for high stimulation load activity patterns (theta-burst-induced LTP) in CA3–CA1 circuit, whereas glucose was sufficient for less demanding neural computation (low-frequency STDP paradigm) (Dembitskaya et al., 2022). Interestingly, this switch in metabolic substrates was also visible *in vivo* when novel object exploration required a higher attentional and cognitive load and for the corresponding *in vivo* LTP expression (Dembitskaya et al., 2022; Figure 2). It remains to examine how glucose and lactate intervene as exclusive or combined fueling in various engrams depending on the body internal states (diet, emotions, effort) and how global energy is regionally redistributed to meet cellular metabolism (Bruckmaier et al., 2020). More generally, investigating how synaptic plasticity induction and maintenance can be controlled by specific diets (ketone-based or high-fat diet) and is altered in various metabolic diseases (obesity, diabetes, but also neurodegenerative diseases) might provide further mechanistic understanding of body–brain interactions. Yet, as these

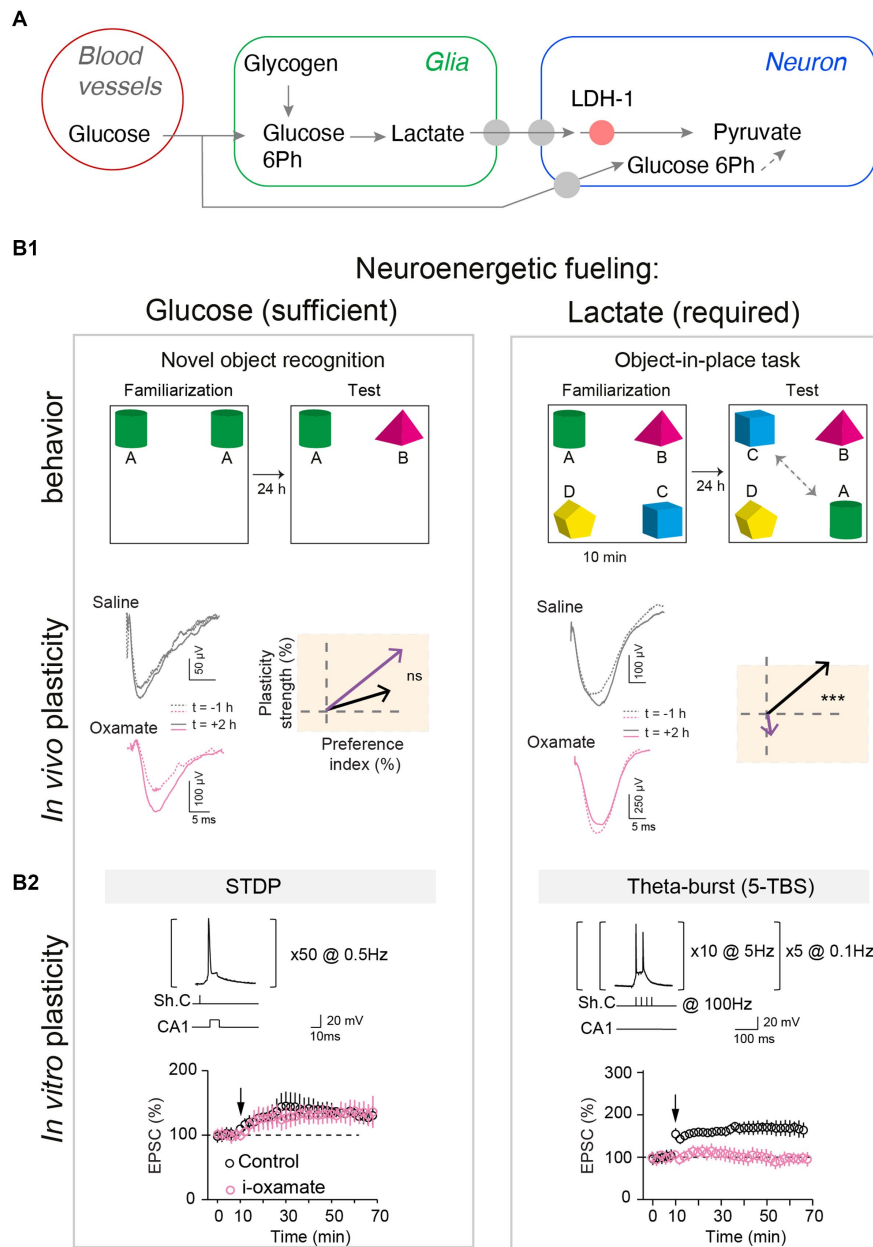


FIGURE 2

Glucose and lactate metabolisms are differently engaged in neuronal fueling for plasticity expression and memory. **(A)** Main steps of the glucose and the glia-derived lactate transports: astrocytic glycogen catalysis into glucose-6-phosphate and then lactate, lactate entry in neurons via monocarboxylate transporters, and lactate conversion into pyruvate by the neuronal lactate dehydrogenase (LDH-1). **(B1)** Lactate metabolism is necessary for learning cognitive tasks requiring high attentional load as exemplified in the object-in-place task (with four objects) and for expressing the corresponding *in vivo* hippocampal LTP, but glucose is sufficient for a less demanding task such as a simple novel object recognition (with two objects). Rats were injected bilaterally, via cannulas implanted above hippocampal CA1 layer, with either saline or oxamate (50 mM), an inhibitor of the neuronal LDH preventing the conversion of lactate into pyruvate, before familiarization step. Rats with saline performed equally well in both tasks whereas rats receiving oxamate did not detect novelty in the object-in-place task (illustrated by a low preference index value) and did not express LTP (averaged vectors: y-axis indicates LTP versus LTD expression and x-axis is the learning performance evaluated with the preference index). *In vivo* synaptic plasticity during behavioral task with evoked-field-EPSP recorded before familiarization (baseline) and 2 hours after familiarization to determine synaptic changes, in relation with behavior. **(B2)** Lactate metabolism is mandatory to fuel the demanding neural computations implicated in NMDA receptor-mediated LTP forms in hippocampus triggered by theta-burst stimulations, while glucose metabolism is sufficient for lighter forms of LTP, based on less and lower-frequency stimulations. The structure of the plasticity induction protocols and the averaged time-course of the synaptic weight after theta-burst stimulation and STDP protocols are illustrated. Oxamate was applied intracellularly (via the patch-clamp pipette) in the sole recorded neuron, and LDH inhibition shows distinct effects on theta-burst stimulation and STDP expression since it prevented theta-burst stimulation-induced LTP but not STDP-induced LTP. In conclusion, scaling of the computational and cognitive loads requires the metabolism of glia-derived lactate to match the neuroenergetic needs of sustained neuronal activity patterns and high cognitive load, and for less demanding plasticity and learning paradigms, glucose suffices as an energy substrate. Adapted from Dembitskaya et al. (2022).

body–brain communication loops and macroscale factors usually fluctuate on slower timescales relative to synaptic activation, evidence for their causal interplay might be particularly challenging to achieve.

## Conclusion

This mini-review presented key avenues, initiated in the synaptic plasticity and memory research field, that put forward a naturalistic viewpoint. This naturalistic lens was first directed at presenting naturalistic synaptic plasticity rules, based on *in vivo* neuronal activity patterns recorded during learning experience, which can then be dissected *in vitro* and/or *in vivo*. Next, it focused on current advances for uncovering naturalistic plasticities, i.e. induced by the animal's own experience, which can provide detailed spatio-temporal characterizations of synaptic plasticity. Finally, besides well-defined third factors (neuromodulators, neuropeptides or glia), this-mini review emphasized that macroscale factors (internal states and the neuronal energy fueling with glucose and lactate metabolisms) can interplay with synaptic plasticity, and hence participate in defining a complex naturalistic context that shapes synaptic plasticity expression during behavior. The next challenges will be to further nail down the relevant synaptic plasticity rules and associated signaling cascades engaged *in vivo*, by investigating causal interactions between neuronal activity patterns, plasticity maps and behavioral consequences. To further enlarge our naturalistic lens on synaptic plasticity, feasibility of mimicry and causality demonstrations should be tested on complex and natural behaviors, such as episodic-like memory or procedural learning. More attention should also be drawn to the existence and contribution of macroscale factors with the major difficulty of their inextricable bounds to natural behaviors and causal manipulations.

## References

- Abdou, K., Shehata, M., Choko, K., Nishizono, H., Matsuo, M., Muramatsu, S. I., et al. (2018). Synapse-specific representation of the identity of overlapping memory engrams. *Science* 360, 1227–1231. doi: 10.1126/science.aat3810
- Ashrafi, G., Wu, Z., Farrell, R. J., and Ryan, T. A. (2017). GLUT4 mobilization supports energetic demands of active synapses. *Neuron* 93, 606–615.e3. doi: 10.1016/j.neuron.2016.12.020
- Attwell, D., and Laughlin, S. B. (2001). An energy budget for signaling in the grey matter of the brain. *J. Cerebral Blood Flow Metabolism* 21, 1133–1145. doi: 10.1097/00004647-200110000-00001
- Basu, J., Srinivas, K. V., Cheung, S. K., Taniguchi, H., Huang, Z. J., and Siegelbaum, S. A. (2013). A cortico-hippocampal learning rule shapes inhibitory microcircuit activity to enhance hippocampal information flow. *Neuron* 79, 1208–1221. doi: 10.1016/j.neuron.2013.07.001
- Bittner, K. C., Milstein, A. D., Grienberger, C., Romani, S., and Magee, J. C. (2017). Behavioral time scale synaptic plasticity underlies CA1 place fields. *Science* 357, 1033–1036. doi: 10.1126/science.aan3846
- Bonvento, G., and Bolanos, J. P. (2021). Astrocyte-neuron metabolic cooperation shapes brain activity. *Cell Metab.* 33, 1546–1564. doi: 10.1016/j.cmet.2021.07.006
- Brandalise, F., and Gerber, U. (2014). Mossy fiber-evoked subthreshold responses induce timing-dependent plasticity at hippocampal CA3 recurrent synapses. *Proc. Natl. Acad. Sci.* 111, 4303–4308. doi: 10.1073/pnas.1317667111
- Bruckmaier, M., Tachtsidis, I., Phan, P., and Lavie, N. (2020). Attention and capacity limits in perception: a cellular metabolism account. *J. Neurosci.* 40, 6801–6811. doi: 10.1523/jneurosci.2368-19.2020
- Brzosko, Z., Mierau, S. B., and Paulsen, O. (2019). Neuromodulation of spike-timing-dependent plasticity: past, present, and future. *Neuron* 103, 563–581. doi: 10.1016/j.neuron.2019.05.041
- Brzosko, Z., Schultz, W., and Paulsen, O. (2015). Retroactive modulation of spike timing-dependent plasticity by dopamine. *Elife* 4:e09685. doi: 10.7554/elifelife.09685
- Cepeda-Prado, E. A., Khodaie, B., Quiceno, G. D., Beythien, S., Edelmann, E., and Lessmann, V. (2022). Calcium-permeable AMPA receptors mediate timing-dependent LTP elicited by low repeat coincident pre- and postsynaptic activity at schaffer collateral-CA1 synapses. *Cereb. Cortex* 32, 1682–1703. doi: 10.1093/cercor/bhab306
- Chistiakova, M., Bannon, N. M., Chen, J.-Y., Bazhenov, M., and Volgushev, M. (2015). Homeostatic role of heterosynaptic plasticity: models and experiments. *Front. Comput. Neurosci.* 9:89. doi: 10.3389/fncom.2015.00089
- Cho, J.-H., Bayazitov, I. T., Meloni, E. G., Myers, K. M., Carlezon, W. A., Zakharenko, S. S., et al. (2012). Coactivation of thalamic and cortical pathways induces input timing-dependent plasticity in amygdala. *Nat. Neurosci.* 15, 113–122. doi: 10.1038/nn.2993
- Choi, D. I., Kim, J., Lee, H., Kim, J., Sung, Y., Choi, J. E., et al. (2021). Synaptic correlates of associative fear memory in the lateral amygdala. *Neuron* 109, 2717–2726.e3. doi: 10.1016/j.neuron.2021.07.003
- Choi, J.-H., Sim, S.-E., Kim, J., Choi, D. I., Oh, J., Ye, S., et al. (2018). Interregional synaptic maps among engram cells underlie memory formation. *Science* 360, 430–435. doi: 10.1126/science.aas9204
- Chuquet, J., Quilichini, P., Nimchinsky, E. A., and Buzsáki, G. (2010). Predominant enhancement of glucose uptake in astrocytes versus neurons during activation of the somatosensory cortex. *J. Neurosci.* 30, 15298–15303. doi: 10.1523/JNEUROSCI.0762-10.2010
- Cui, Y., Paillé, V., Xu, H., Genet, S., Delord, B., Fino, E., et al. (2015). Endocannabinoids mediate bidirectional striatal spike-timing-dependent plasticity. *J. Physiol.* 593, 2833–2849. doi: 10.1113/jp270324

## Author contributions

CP wrote the first draft of the “From neuronal activity patterns to plasticity rules” and “Toward the uncovering of a learning-induced plasticity” sections, and design the Figure 1. LV and NG wrote the “The (almost) overlooked of synaptic plasticity research: the body–brain communication loops and neuroenergetics” section. LV redrafted the whole manuscript, and design the Figure 2. All authors wrote “background” and “conclusions” sections, and have edited and corrected the manuscript.

## Funding

This work was supported by grants from the Agence Nationale pour la Recherche (grants ANR EngFlea), the LabEx Paris-Sciences et Lettres (PSL) and Collège de France.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Cui, Y., Prokin, I., Mendes, A., Berry, H., and Venance, L. (2018). Robustness of STDP to spike timing jitter. *Sci. Rep.* 8:8139. doi: 10.1038/s41598-018-26436-y
- Cui, Y., Prokin, I., Xu, H., Delord, B., Genet, S., Venance, L., et al. (2016). Endocannabinoid dynamics gate spike-timing dependent depression and potentiation. *Elife* 5:e13185. doi: 10.7554/elifelife.13185
- D'Amour, J. A., and Froemke, R. C. (2015). Inhibitory and excitatory spike-timing-dependent plasticity in the auditory cortex. *Neuron* 86, 514–528. doi: 10.1016/j.neuron.2015.03.014
- Debanne, D., and Inglebert, Y. (2023). Spike timing-dependent plasticity and memory. *Curr. Opin. Neurobiol.* 80:102707. doi: 10.1016/j.comb.2023.102707
- Demas, J., Manley, J., Tejera, F., Barber, K., Kim, H., Traub, F. M., et al. (2021). High-speed, cortex-wide volumetric recording of neuroactivity at cellular resolution using light beads microscopy. *Nat. Methods* 18, 1103–1111. doi: 10.1038/s41592-021-01239-8
- Dembitskaya, Y., Piette, C., Perez, S., Berry, H., Magistretti, P. J., and Venance, L. (2022). Lactate supply overtakes glucose when neural computational and cognitive loads scale up. *Proc. Natl. Acad. Sci.* 119:e2212004119. doi: 10.1073/pnas.2212004119
- Drew, P. J., and Abbott, L. F. (2006). Extending the effects of spike-timing-dependent plasticity to behavioral timescales. *Proc. Natl. Acad. Sci. U. S. A.* 103, 8876–8881. doi: 10.1073/pnas.0600676103
- Dudman, J. T., Tsay, D., and Siegelbaum, S. A. (2007). A role for synaptic inputs at distal dendrites: instructive signals for hippocampal long-term plasticity. *Neuron* 56, 866–879. doi: 10.1016/j.neuron.2007.10.020
- Edelmann, E., Cepeda-Prado, E., and Leßmann, V. (2017). Coexistence of multiple types of synaptic plasticity in individual hippocampal CA1 pyramidal neurons. *Front. Synaptic Neurosci.* 9:7. doi: 10.3389/fnsyn.2017.00007
- El-Boustani, S., Ip, J. P., Breton-Provencher, V., Knott, G. W., Okuno, H., Bito, H., et al. (2018). Locally coordinated synaptic plasticity of visual cortex neurons in vivo. *Science* 360, 1349–1354. doi: 10.1126/science.aao0862
- Fan, L. Z., Kim, D. K., Jennings, J. H., Tian, H., Wang, P. Y., Ramakrishnan, C., et al. (2023). All-optical physiology resolves a synaptic basis for behavioral timescale plasticity. *Cells* 18, 543–559. doi: 10.1016/j.cell.2022.12.035
- Farmer, J., Zhao, X., van Praag, H., Wodtke, K., Gage, F. H., and Christie, B. R. (2004). Effects of voluntary exercise on synaptic plasticity and gene expression in the dentate gyrus of adult male Sprague–Dawley rats in vivo. *Neuroscience* 124, 71–79. doi: 10.1016/j.neuroscience.2003.09.029
- Feldman, D. E. (2012). The spike-timing dependence of plasticity. *Neuron* 75, 556–571. doi: 10.1016/j.neuron.2012.08.001
- Ferro, M., Lamanna, J., Ripamonti, M., Racchetti, G., Arena, A., Spadini, S., et al. (2017). Functional mapping of brain synapses by the enriching activity-marker SynaptoZip. *Nat. Commun.* 8:1229. doi: 10.1038/s41467-017-01335-4
- Fino, E., Deniau, J.-M., and Venance, L. (2009). Brief subthreshold events can act as Hebbian signals for long-term plasticity. *PLoS One* 4:e6557. doi: 10.1371/journal.pone.0006557
- Flavell, S. W., Gogolla, N., Lovett-Barron, M., and Zelikowsky, M. (2022). The emergence and influence of internal states. *Neuron* 110, 2545–2570. doi: 10.1016/j.neuron.2022.04.030
- Folschweiller, S., and Sauer, J.-F. (2023). Behavioral state-dependent modulation of prefrontal cortex activity by respiration. *J. Neurosci.* 43, 4795–4807. doi: 10.1523/jneurosci.2075-22.2023
- Foncelle, A., Mendes, A., Jędrzejewska-Szmek, J., Valtcheva, S., Berry, H., Blackwell, K. T., et al. (2018). Modulation of spike-timing dependent plasticity: towards the inclusion of a third factor in computational models. *Front. Comput. Neurosci.* 12:49. doi: 10.3389/fncom.2018.00049
- Frémaux, N., and Gerstner, W. (2016). Neuromodulated spike-timing-dependent plasticity, and theory of three-factor learning rules. *Front. Neural Circ.* 9:85. doi: 10.3389/fnirc.2015.00085
- Froemke, R. C., and Dan, Y. (2002). Spike-timing-dependent synaptic modification induced by natural spike trains. *Nature* 416, 433–438. doi: 10.1038/416433a
- Froemke, R. C., Tsay, I. A., Raad, M., Long, J. D., and Dan, Y. (2006). Contribution of individual spikes in burst-induced long-term synaptic modification. *J. Neurophysiol.* 95, 1620–1629. doi: 10.1152/jn.00910.2005
- Fuchsberger, T., Clopath, C., Jarzebowsky, P., Brzosko, Z., Wang, H., and Paulsen, O. (2022). Postsynaptic burst reactivation of hippocampal neurons enables associative plasticity of temporally discontinuous inputs. *Elife* 11:e81071. doi: 10.7554/elifelife.81071
- Genoud, C., Quairiaux, C., Steiner, P., Hirling, H., Welker, E., and Knott, G. W. (2006). Plasticity of astrocytic coverage and glutamate transporter expression in adult mouse cortex. *PLoS Biol.* 4:e343. doi: 10.1371/journal.pbio.0040343
- George, T. M., de Cothi, W., Stachenfeld, K. L., and Barry, C. (2023). Rapid learning of predictive maps with STDP and theta phase precession. *Elife* 12:e80663. doi: 10.7554/elifelife.80663
- Gervasi, N., Tchéno, P., and Preat, T. (2010). PKA dynamics in a drosophila learning center: coincidence detection by rutabaga adenyl cyclase and spatial regulation by dunce phosphodiesterase. *Neuron* 65, 516–529. doi: 10.1016/j.neuron.2010.01.014
- Gobbo, F., Marchetti, L., Jacob, A., Pinto, B., Binini, N., Pecoraro Bisogni, F., et al. (2017). Activity-dependent expression of channelrhodopsin at neuronal synapses. *Nat. Commun.* 8:1629. doi: 10.1038/s41467-017-01699-7
- Gomes da Silva, S., Unsain, N., Mascó, D. H., Toscano-Silva, M., de Amorim, H. A., Silva Araújo, B. H., et al. (2010). Early exercise promotes positive hippocampal plasticity and improves spatial memory in the adult life of rats. *Hippocampus* 22, 347–358. doi: 10.1002/hipo.20903
- Graupner, M., Wallisch, P., and Ostojic, S. (2016). Natural firing patterns imply low sensitivity of synaptic plasticity to spike timing compared with firing rate. *J. Neurosci.* 36, 11238–11258. doi: 10.1523/jneurosci.0104-16.2016
- Graves, A. R., Roth, R. H., Tan, H. L., Zhu, Q., Bygrave, A. M., Lopez-Ortega, E., et al. (2021). Visualizing synaptic plasticity in vivo by large-scale imaging of endogenous AMPA receptors. *Elife* 10:e66809. doi: 10.7554/elifelife.66809
- Han, J.-H., Kushner, S. A., Yiu, A. P., Cole, C. J., Matynia, A., Brown, R. A., et al. (2007). Neuronal competition and selection during memory formation. *Science* 316, 457–460. doi: 10.1126/science.1139438
- Hardie, J., and Spruston, N. (2009). Synaptic depolarization is more effective than back-propagating action potentials during induction of associative long-term potentiation in hippocampal pyramidal neurons. *J. Neurosci.* 29, 3233–3241. doi: 10.1523/JNEUROSCI.6000-08.2009
- Harvey, C. D., and Svoboda, K. (2007). Locally dynamic synaptic learning rules in pyramidal neuron dendrites. *Nature* 450, 1195–1200. doi: 10.1038/nature06416
- He, K., Huertas, M., Hong, S. Z., Tie, X., Hell, J. W., Shouval, H., et al. (2015). Distinct eligibility traces for LTP and LTP in cortical synapses. *Neuron* 88, 528–538. doi: 10.1016/j.neuron.2015.09.037
- Hiratani, N., and Fukui, T. (2017). Detailed dendritic excitatory/inhibitory balance through heterosynaptic spike-timing-dependent plasticity. *J. Neurosci.* 37, 12106–12122. doi: 10.1523/jneurosci.0027-17.2017
- Holthoff, K., Kovalchuk, Y., Yuste, R., and Konnerth, A. (2004). Single-shock LTP by local dendritic spikes in pyramidal neurons of mouse visual cortex. *J. Physiol.* 560, 27–36. doi: 10.1113/jphysiol.2004.072678
- Holtmaat, A., and Svoboda, K. (2009). Experience-dependent structural synaptic plasticity in the mammalian brain. *Nat. Rev. Neurosci.* 10, 647–658. doi: 10.1038/nrn2699
- Hsueh, B., Chen, R., Jo, Y., Tang, D., Raffiee, M., Kim, Y. S., et al. (2023). Cardiogenic control of affective behavioural state. *Nature* 615, 292–299. doi: 10.1038/s41586-023-05748-8
- Inglebert, Y., Aljaffef, J., Brunel, N., and Debanne, D. (2020). Synaptic plasticity rules with physiological calcium levels. *Proc. Natl. Acad. Sci.* 117, 33639–33648. doi: 10.1073/pnas.2013663117
- Isaac, J. T., Buchanan, K. A., Muller, R. U., and Mellor, J. R. (2009). Hippocampal place cell firing patterns can induce long-term synaptic plasticity in vitro. *J. Neurosci.* 29, 6840–6850. doi: 10.1523/jneurosci.0731-09.2009
- Jackman, S. L., Beneduce, B. M., Drew, I. R., and Regehr, W. G. (2014). Achieving high-frequency optical control of synaptic transmission. *J. Neurosci.* 34, 7704–7714. doi: 10.1523/JNEUROSCI.4694-13.2014
- Jacob, V., Brasier, D. J., Erchova, I., Feldman, D., and Shulz, D. E. (2007). Spike timing-dependent synaptic depression in their vivobarrel cortex of the rat. *J. Neurosci.* 27, 1271–1284. doi: 10.1523/jneurosci.4264-06.2007
- Josselyn, S. A., and Tonegawa, S. (2020). Memory engrams: recalling the past and imagining the future. *Science* 367:eaaw4325. doi: 10.1126/science.aaw4325
- Kanwal, J. K., Coddington, E., Frazer, R., Limbania, D., Turner, G., Davila, K. J., et al. (2021). Internal state: dynamic, interconnected communication loops distributed across body, brain, and time. *Integr. Comp. Biol.* 61, 867–886. doi: 10.1093/icb/ibab101
- Kasichke, K. A., Vishwasrao, H. D., Fisher, P. J., Zipfel, W. R., and Webb, W. W. (2004). Neural activity triggers neuronal oxidative metabolism followed by astrocytic glycolysis. *Science* 305, 99–103. doi: 10.1126/science.1096485
- Kim, J., Lee, S., Park, K., Hong, I., Song, B., Son, G., et al. (2007). Amygdala depotentiation and fear extinction. *Proc. Natl. Acad. Sci.* 104, 20955–20960. doi: 10.1073/pnas.0710548105
- Lacoste, B., Comin, C. H., Ben-Zvi, A., Kaeser, P. S., Xu, X., Costa Lda, F., et al. (2014). Sensory-related neural activity regulates the structure of vascular networks in the cerebral cortex. *Neuron* 83, 1117–1130. doi: 10.1016/j.neuron.2014.07.034
- Lee, S.-J. R., Escobedo-Lozoya, Y., Szatmari, E. M., and Yasuda, R. (2009). Activation of camkii in single dendritic spines during long-term potentiation. *Nature* 458, 299–304. doi: 10.1038/nature07842
- Lee, C., Lee, B. H., Jung, H., Lee, C., Sung, Y., Kim, H., et al. (2023). Hippocampal engram networks for fear memory recruit new synapses and modify pre-existing synapses in vivo. *Curr. Biol.* 33, 507–516.e3. doi: 10.1016/j.cub.2022.12.038
- Leroy, F., Brann, D. H., Meira, T., and Siegelbaum, S. A. (2017). Input-timing-dependent plasticity in the hippocampal CA2 region and its potential role in social memory. *Neuron* 95, 1089–1102.e5. doi: 10.1016/j.neuron.2019.03.021



- Lisman, J., and Spruston, N. (2010). Questions about STDP as a general model of synaptic plasticity. *Front. Synaptic Neurosci.* 2:140. doi: 10.3389/fnsyn.2010.00140
- Losonczy, A., Makara, J. T., and Magee, J. C. (2008). Compartmentalized dendritic plasticity and input feature storage in neurons. *Nature* 452, 436–441. doi: 10.1038/nature06725
- Ma, T., Cheng, Y., Roltsch Hellard, E., Wang, X., Lu, J., Gao, X., et al. (2018). Bidirectional and long-lasting control of alcohol-seeking behavior by corticostriatal LTP and LTD. *Nat. Neurosci.* 21, 373–383. doi: 10.1038/s41593-018-0081-9
- Magee, J. C., and Grienberger, C. (2020). Synaptic plasticity forms and functions. *Annu. Rev. Neurosci.* 43, 95–117. doi: 10.1146/annurev-neuro-090919-022842
- Magistretti, P. J., and Allaman, I. (2018). Lactate in the brain: from metabolic end-product to signalling molecule. *Nat. Rev. Neurosci.* 19, 235–249. doi: 10.1038/nrn.2018.19
- Martin, S. J., and Morris, R. G. M. (2002). New life in an old idea: the synaptic plasticity and memory hypothesis revisited. *Hippocampus* 12, 609–636. doi: 10.1002/hipo.10107
- Matsuo, N., Reijmers, L., and Mayford, M. (2008). Spine-type-specific recruitment of newly synthesized AMPA receptors with learning. *Science* 319, 1104–1107. doi: 10.1126/science.1149967
- Mayford, M., and Reijmers, L. (2015). Exploring memory representations with activity-based genetics. *Cold Spring Harb. Perspect. Biol.* 8:a021832. doi: 10.1101/cshperspect.a021832
- McFarlan, A. R., Chou, C. Y., Watanabe, A., Cherepacha, N., Haddad, M., Owens, H., et al. (2023). The plasticity of cortical interneurons. *Nat. Rev. Neurosci.* 24, 80–97. doi: 10.1038/s41583-022-00663-9
- Mehaffey, W. H., and Doupe, A. J. (2015). Naturalistic stimulation drives opposing heterosynaptic plasticity at two inputs to songbird cortex. *Nat. Neurosci.* 18, 1272–1280. doi: 10.1038/nn.4078
- Meliza, C. D., and Dan, Y. (2006). Receptive-field modification in rat visual cortex induced by paired visual stimulation and single-cell spiking. *Neuron* 49, 183–189. doi: 10.1016/j.neuron.2005.12.009
- Mendes, A., Vignoud, G., Perez, S., Perrin, E., Touboul, J., and Venance, L. (2020). Concurrent thalamostriatal and corticostriatal spike-timing-dependent plasticity and heterosynaptic interactions shape striatal plasticity map. *Cereb. Cortex* 30, 4381–4401. doi: 10.1093/cercor/bhaa024
- Mery, F., and Kawecki, T. J. (2005). A cost of long-term memory in drosophila. *Science* 308:1148. doi: 10.1126/science.1111331
- Miyashita, T., Shao, Y. R., Chung, J., Pourzia, O., and Feldman, D. E. (2013). Long-term channelrhodopsin-2 (ChR2) expression can induce abnormal axonal morphology and targeting in cerebral cortex. *Front. Neural Circ.* 7:8. doi: 10.3389/fncir.2013.00008
- Murphy-Royal, C., Johnston, A. D., Boyce, A. K., Diaz-Castro, B., Institoris, A., Peringod, G., et al. (2020). Stress gates an astrocytic energy reservoir to impair synaptic plasticity. *Nat. Commun.* 11:2014. doi: 10.1038/s41467-020-15778-9
- Nabavi, S., Fox, R., Proulx, C. D., Lin, J. Y., Tsien, R. Y., and Malinow, R. (2014). Engineering a memory with LTD and LTP. *Nature* 511, 348–352. doi: 10.1038/nature13294
- Newman, L. A., Korol, D. L., and Gold, P. E. (2011). Lactate produced by glycogenolysis in astrocytes regulates memory processing. *PLoS One* 6:e28427. doi: 10.1371/journal.pone.0028427
- O'Donnell, C. (2023). Nonlinear slow-timescale mechanisms in synaptic plasticity. *Curr. Opin. Neurobiol.* 82:102778. doi: 10.1016/j.conb.2023.102778
- Paulsen, O., and Sejnowski, T. J. (2000). Natural patterns of activity and long-term synaptic plasticity. *Curr. Opin. Neurobiol.* 10, 172–180. doi: 10.1016/s0959-4388(00)00076-3
- Pedersen, B. K., and Febbraio, M. A. (2012). Muscles, exercise and obesity: skeletal muscle as a secretory organ. *Nat. Rev. Endocrinol.* 8, 457–465. doi: 10.1038/nrendo.2012.49
- Pfeiffer, T., Poll, S., Bancelin, S., Angibaud, J., Inavalli, V. K., Keppler, K., et al. (2018). Chronic 2p-sted imaging reveals high turnover of dendritic spines in the hippocampus in vivo. *Elife* 7. doi: 10.7554/elife.34700
- Piette, C., Touboul, J., and Venance, L. (2020). Engrams of fast learning. *Front. Cell. Neurosci.* 14:575915. doi: 10.3389/fncel.2020.575915
- Plaças, P.-Y., de Tredern, É., Scheunemann, L., Trannoy, S., Goguel, V., Han, K.-A., et al. (2017). Upregulated energy metabolism in the drosophila mushroom body is the trigger for long-term memory. *Nat. Commun.* 8:15510. doi: 10.1038/ncomms15510
- Plaças, P.-Y., and Preat, T. (2013). To favor survival under food shortage, the brain disables costly memory. *Science* 339, 440–442. doi: 10.1126/science.1226018
- Priestley, J. B., Bowler, J. C., Rolotti, S. V., Fusi, S., and Losonczy, A. (2022). Signatures of rapid plasticity in hippocampal CA1 representations during novel experiences. *Neuron* 110, 1978–1992.e6. doi: 10.1016/j.neuron.2022.03.026
- Ray, S., Li, M., Koch, S. P., Mueller, S., Boehm-Sturm, P., Wang, H., et al. (2020). Seasonal plasticity in the adult somatosensory cortex. *Proc. Natl. Acad. Sci. U. S. A.* 117, 32136–32144. doi: 10.1073/pnas.1922888117
- Ruminot, I., Schmalzle, J., Leyton, B., Barros, L. F., and Deitmer, J. W. (2017). Tight coupling of astrocyte energy metabolism to synaptic activity revealed by genetically encoded FRET nanosensors in hippocampal tissue. *J. Cereb. Blood Flow Metab.* 39, 513–523. doi: 10.1177/0271678X17737012
- Saur, L., Baptista, P. P., de Senna, P. N., Paim, M. F., Nascimento, P., Ilha, J., et al. (2014). Physical exercise increases GFAP expression and induces morphological changes in hippocampal astrocytes. *Brain Struct. Funct.* 219, 293–302. doi: 10.1007/s00429-012-0500-8
- Sejnowski, T. J. (1999). The book of hebb. *Neuron* 24, 773–776. doi: 10.1016/s0896-6273(00)81025-9
- Suvrathan, A., Payne, H. L., and Raymond, J. L. (2016). Timing rules for synaptic plasticity matched to behavioral function. *Neuron* 92, 959–967. doi: 10.1016/j.neuron.2016.10.022
- Suzuki, A., Stern, S. A., Bozdagi, O., Huntley, G. W., Walker, R. H., Magistretti, P. J., et al. (2011). Astrocyte-neuron lactate transport is required for long-term memory formation. *Cells* 144, 810–823. doi: 10.1016/j.cell.2011.02.018
- Tadi, M., Allaman, I., Lengacher, S., Grenningloh, G., and Magistretti, P. J. (2015). Learning-induced gene expression in the hippocampus reveals a role of neuron-astrocyte metabolic coupling in long term memory. *PLoS One* 10:e0141568. doi: 10.1371/journal.pone.0141568.eCollection
- Takeishi, A., Yeon, J., Harris, N., Yang, W., and Sengupta, P. (2020). Feeding state functionally reconfigures a sensory circuit to drive thermosensory behavioral plasticity. *Elife* 9:e61167. doi: 10.7554/eLife.61167
- Thorn, C. A., Atallah, H., Howe, M., and Graybiel, A. M. (2010). Differential dynamics of activity changes in dorsolateral and dorsomedial striatal loops during learning. *Neuron* 66, 781–795. doi: 10.1016/j.neuron.2010.04.036
- Todorova, V., and Blokland, A. (2017). Mitochondria and synaptic plasticity in the mature and aging nervous system. *Curr. Neuropharmacol.* 15, 166–173. doi: 10.2174/1570159x14666160414111821
- van Praag, H., Christie, B. R., Sejnowski, T. J., and Gage, F. H. (1999). Running enhances neurogenesis, learning, and long-term potentiation in mice. *Proc. Natl. Acad. Sci.* 96, 13427–13431. doi: 10.1073/pnas.96.23.13427
- Vaynman, S., Ying, Z., and Gomez-Pinilla, F. (2004). Hippocampal BDNF mediates the efficacy of exercise on synaptic plasticity and cognition. *Eur. J. Neurosci.* 20, 2580–2590. doi: 10.1111/j.1460-9568.2004.03720.x
- Vetere, G., Tran, L. M., Moberg, S., Steadman, P. E., Restivo, L., Morrison, F. G., et al. (2019). Memory formation in the absence of experience. *Nat. Neurosci.* 22, 933–940. doi: 10.1038/s41593-019-0389-0
- Vukovic, J., Colditz, M. J., Blackmore, D. G., Ruitenberg, M. J., and Bartlett, P. F. (2012). Microglia modulate hippocampal neural precursor activity in response to exercise and aging. *J. Neurosci.* 32, 6435–6443. doi: 10.1523/JNEUROSCI.5925-11.2012
- Waters, J., and Helmchen, F. (2004). Boosting of action potential backpropagation by neocortical network activity in vivo. *J. Neurosci.* 24, 11127–11136. doi: 10.1523/jneurosci.2933-04.2004
- Xiong, Q., Znamenskiy, P., and Zador, A. M. (2015). Selective Corticostriatal plasticity during acquisition of an auditory discrimination task. *Nature* 521, 348–351. doi: 10.1038/nature14225
- Yagishita, S., Hayashi-Takagi, A., Ellis-Davies, G. C. R., Urakubo, H., Ishii, S., and Kasai, H. (2014). A critical time window for dopamine actions on the structural plasticity of dendritic spines. *Science* 345, 1616–1620. doi: 10.1126/science.1255514
- Yao, H., and Dan, Y. (2001). Stimulus timing-dependent plasticity in cortical processing of orientation. *Neuron* 32, 315–323. doi: 10.1016/s0896-6273(01)00460-3
- Zhang, Y., Cudmore, R. H., Lin, D.-T., Linden, D. J., and Huganir, R. L. (2015). Visualization of NMDA receptor-dependent AMPA receptor synaptic plasticity in vivo. *Nat. Neurosci.* 18, 402–407. doi: 10.1038/nm.3936
- Zhao, X., Hsu, C.-L., and Spruston, N. (2022). Rapid synaptic plasticity contributes to a learned conjunctive code of position and choice-related information in the hippocampus. *Neuron* 110, 96–108.e4. doi: 10.1016/j.neuron.2021.10.003



## Adolescent Nutrition 1

## Nutrition in adolescent growth and development

Shane A Norris\*, Edward A Frongillo\*, Maureen M Black, Yanhui Dong, Caroline Fall, Michelle Lampl, Angela D Liese, Mariam Naguib, Ann Prentice, Tamsen Rochat, Charles B Stephensen, Chiwoneso B Tinago, Kate A Ward, Stephanie V Wrottesley, George C Patton†

Lancet 2022; 399: 172–84

Published Online  
November 29, 2021  
[https://doi.org/10.1016/S0140-6736\(21\)01590-7](https://doi.org/10.1016/S0140-6736(21)01590-7)

This is the first in a **Series** of three papers on adolescent nutrition

\*Co-lead authors

†Co-lead of the Series (all authors between the co-lead authors and the Series co-lead are listed in alphabetical order)

SAMRC Developmental Pathways for Health Research Unit, Department of Paediatrics, University of the Witwatersrand, Johannesburg, South Africa (S A Norris PhD, T Rochat PhD, S V Wrottesley PhD); Global Health Research Institute, School of Health and Human Development (S A Norris) and MRC Lifecourse Epidemiology Unit (C Fall DM, K A Ward PhD), University of Southampton, Southampton, UK; Department of Health Promotion, Education, and Behavior (E A Frongillo PhD) and Department of Epidemiology and Biostatistics (A D Liese PhD), Arnold School of Public Health, University of South Carolina, Columbia, SC, USA; Department of Pediatrics, University of Maryland School of Medicine, Baltimore, MD, USA (M M Black PhD); RTI International, Research Triangle Park, NC, USA (M M Black); Institute of Child and Adolescent Health, School of Public Health, Peking University, Beijing, China (Y Dong PhD); Emory Center for the Study of Human Health, Emory University, Atlanta, GA, USA (M Lampl MD); Department of Medicine, McGill University, Montreal, QC, Canada (M Naguib MD); MRC Nutrition and Bone Health Group, Cambridge, UK (A Prentice PhD); MRC Unit The Gambia, London School of Hygiene & Tropical Medicine, London, UK (A Prentice,

During adolescence, growth and development are transformative and have profound consequences on an individual's health in later life, as well as the health of any potential children. The current generation of adolescents is growing up at a time of unprecedented change in food environments, whereby nutritional problems of micronutrient deficiency and food insecurity persist, and overweight and obesity are burgeoning. In a context of pervasive policy neglect, research on nutrition during adolescence specifically has been underinvested, compared with such research in other age groups, which has inhibited the development of adolescent-responsive nutritional policies. One consequence has been the absence of an integrated perspective on adolescent growth and development, and the role that nutrition plays. Through late childhood and early adolescence, nutrition has a formative role in the timing and pattern of puberty, with consequences for adult height, muscle, and fat mass accrual, as well as risk of non-communicable diseases in later life. Nutritional effects in adolescent development extend beyond musculoskeletal growth, to cardiorespiratory fitness, neurodevelopment, and immunity. High rates of early adolescent pregnancy in many countries continue to jeopardise the growth and nutrition of female adolescents, with consequences that extend to the next generation. Adolescence is a nutrition-sensitive phase for growth, in which the benefits of good nutrition extend to many other physiological systems.

### Introduction

Adolescence is a transformative life phase, with growth and maturation of all organs and physiological systems. On average, 10–19 year olds gain 20% of their final adult height and 50% of adult weight during this phase, with a considerable remodelling of the skeleton and an increase in bone mass of up to 40%.<sup>1</sup> Inevitably, the link between nutrition and adolescent development is strong. For example, particularly in girls, iron requirements increase sharply during adolescence to meet additional needs relating to menstruation. Iron deficiency in adolescents results in compromised growth, decreased cognitive function, and depressed immune function.<sup>2</sup> Despite this understanding, iron deficiency anaemia remains prevalent worldwide, showing little reduction over three decades, and is the third most important cause of lost disability-adjusted life-years in adolescents.<sup>3</sup>

Not only are there more adolescents nowadays than at any other timepoint in human history but they are also growing up at a time of momentous shift—ie, rapid urbanisation, climate change, food systems shifting

towards foods with an increased caloric and decreased nutritional value, the COVID-19 pandemic, and growing socioeconomic inequality. The consequences of these changing contexts have profound impacts on adolescent nutrition and development. Figure 1 presents data from 54 million children and adolescents (aged 5–19 years) and shows the effects that varying nutrition and living conditions can have on height and adiposity (ie, body-mass index [BMI]) over age and time, and across countries. There is a difference of at least 20 cm in the mean height of individuals aged 19 years between the tallest and shortest populations. The data highlight that, for many countries, linear growth in children and adolescents still falls below the WHO reference. Despite this evidence of persisting undernutrition, overweight and obesity are now widespread. Since height and BMI have been considered together over the past two decades, the unhealthiest changes of gaining too little height, too much weight, or both, have been prevalent in both high-income countries and low-income and middle-income countries (LMICs).<sup>4</sup> Consequences include an increased risk of non-communicable diseases (NCDs) and a suboptimal start to life in the next generation.<sup>5</sup>

Understanding adolescent biology and its relationship to nutrition is essential for identifying the best timing and form of action, and for avoiding potentially negative consequences. Therefore, this first Series paper synthesises our understanding of adolescent biological development and its relationship with nutrition.

### Pubertal maturation

The adolescent growth phase starts with puberty, which drives linear growth; accrual of bone, muscle, and fat mass; and maturation of biological systems. The onset and

#### Search strategy and selection criteria

For this narrative review, we searched Pubmed, MEDLINE, Google Scholar, and Embase, without date or language restrictions, from Jan 31, 2020, to March 30, 2021, for literature pertaining to the general domains of puberty, physical growth, body composition, neurodevelopment, cardiorespiratory fitness, immune development, and adolescent pregnancy and intergenerational consequences. We also sought longitudinal studies to illustrate further effects of nutrition on adolescent growth and development.

duration of puberty differ markedly between adolescents living in environments with varying childhood nutrition.<sup>6</sup> Pubertal timing, as indicated by the late pubertal event of menstruation (menarche) in girls, has decreased by 1·0 year in high-income countries over time, from a mean of 13·5 years for births before 1930 to 12·6 years for births between 1970 and 1984.<sup>7</sup> Among healthy girls in LMICs during 2009–17, mean age at menarche was estimated to be 12·3 years.<sup>8</sup> In some LMIC populations, where nutrition has improved to a lesser extent than typical LMIC populations, the mean age of menarche is significantly later; for example, 15·1 years in rural parts of The Gambia.

Adiposity is associated with pubertal form. For girls, the mean age of thelarche (ie, breast budding)—an early indicator of gonadal maturation—is 10·2 years for individuals with underweight, 10·4 years for individuals with normal weight, and 8·4 years for individuals with overweight.<sup>8</sup> In boys, mean age of puberty onset—indicated by the scrotum becoming pendulous—is 11·3 years for individuals with underweight, 11·0 years for individuals with normal weight, and 10·3 years for individuals with overweight.<sup>8</sup> Nutritional status not only affects onset of puberty but also its duration.<sup>9</sup> In Australian children aged 8–9 years, high androgen concentrations, reflecting adrenal maturation as the earliest pubertal change, were associated with an increased BMI and waist circumference.<sup>10</sup> In turn, pubertal form has implications for obesity in later life, with early onset and short duration predicting increased adiposity in adulthood (aged  $\geq 40$  years).<sup>11,12</sup>

Furthermore, previous parental and childhood nutrition influences pubertal form. For example, maternal obesity before conception predicts early pubertal onset in offspring.<sup>13</sup> Children who were breastfed for 6 months or longer have a later onset of pubertal development than do those who were not breastfed or were breastfed for less than 6 months, perhaps in part reflecting different growth patterns in infancy.<sup>14</sup> A high intake of animal protein in children at age 5–6 years and 12 years predicted an earlier onset of the pubertal growth spurt, whereas a high intake of vegetable protein predicted a later onset.<sup>15–17</sup> A high dietary intake of carbohydrates and fats in girls aged 8 years predicted earlier gonadal maturation and menarche, and faster pubertal tempo than did a high intake of protein.<sup>18</sup> Consumption of sugar-sweetened beverages advances onset of menarche in girls.<sup>19</sup> Given the extent to which pubertal form is a marker of growth, development, and NCD risk in later life, there is a need for research to develop a comprehensive lifecourse understanding of its nutritional and other, potentially modifiable, determinants.

## Linear growth

Adolescent linear growth has the highest velocity after infancy and occurs at the growth plate in a two-step cellular process. First, bone elongation cells—chondrocytes—sequentially proliferate, secrete matrix, and

### Key messages

- Adolescence is a time of transformative growth when both undernutrition and obesity affect the maturation of multiple physiological systems
- Adolescent malnutrition is multiplicative in that, if any one physiological system is affected, the development of other systems will also be compromised
- Nutrition in childhood and early adolescence affects the timing and form of puberty with consequences on linear growth, body composition, and maturation of other physiological systems
- Although some catch-up growth in height can occur in late childhood and early adolescence, it rarely happens if the adverse nutritional environment of early life persists into adolescence
- Across late childhood and early adolescence, the pubertal transition offers a nutrition-sensitive window to promote healthy growth and reduce risk of obesity in later life
- Given that nutrition is a cornerstone of investments in human capital, scaling up research into the effects of nutrition on adolescent growth and development is a pressing need

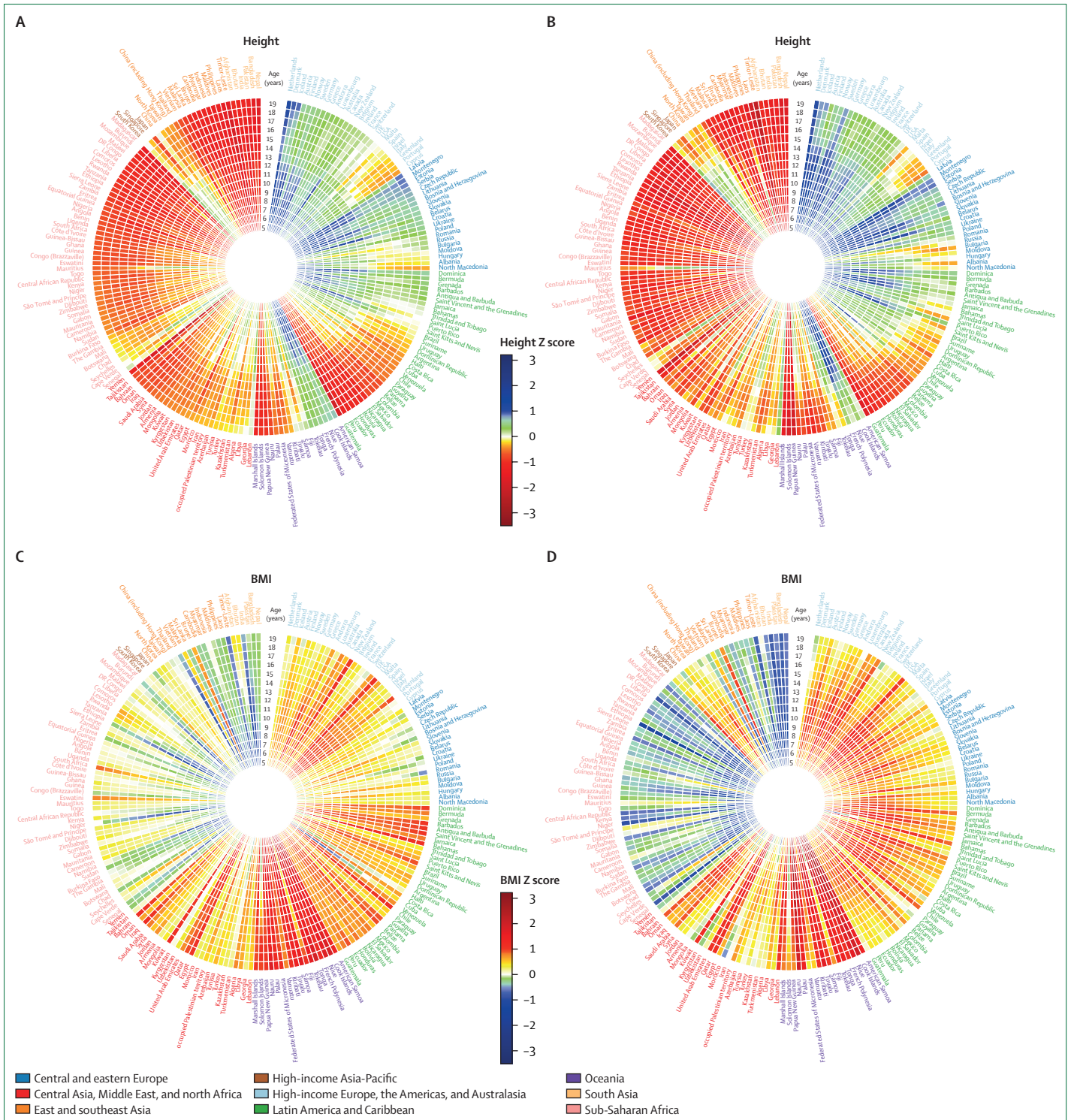
undergo hypertrophy, hydraulically propelling bone elongation and producing a protein model of the lengthened bone. Second, bone-secreting cells—osteoblasts—secrete a mineral matrix on the newly created protein model to consolidate the new growth into bone.<sup>20–22</sup> Without the first step, linear growth cannot occur; without the second step, new growth is lost, and the protein model is resorbed. Mechanisms underlying progress across the phases of the chondrocytic lifecycle, from stem cells to hypertrophic transition, involve prompts and inhibitions from complex networks of regulatory proteins<sup>23,24</sup> and endocrine signals.<sup>25</sup> Many nutrients are important for chondrocytic function and for ensuring mineral consolidation.<sup>26–29</sup> Any nutritional intervention to ameliorate retardation in linear growth should consider both of these steps, with the added challenge that the underlying cause originates from past conditions in which the child lived and might be neither evident nor reparable due to missed opportunity, epigenetic effects, or both. Albeit incomplete, some restoration of lost linear growth can occur; however, this can only happen if the intervention substantially improves socioeconomic and living conditions, such as through adoption. Nutrition-specific interventions alone are not likely to restore lost growth.<sup>30</sup>

Height has increased in all populations over decades.<sup>31,32</sup> In high-income countries, this trend is modest in children aged 6 years and largest in adolescents aged 10–14 years; in LMICs, trends vary.<sup>33</sup> Preschool children (aged <60 months) living in conditions conducive to good health and development grow similarly. For preadolescent children in favourable conditions, height across global populations differs by 3–5 cm,<sup>34</sup> and Asian populations are slightly shorter.<sup>31</sup> Both nutrition and living conditions contribute to attained height.<sup>35</sup> South Asian children living in the Netherlands grew taller between 1992 and 2010, but remained shorter than their Dutch peers at each age, with greater divergence during

K A Ward); USDA Western Human Nutrition Research Center and Nutrition Department, University of California, Davis, CA, USA (C B Stephensen PhD); Department of Health, West Chester University, West Chester, PA, USA (C B Tinago PhD); Murdoch Children's Research Institute, University of Melbourne, Melbourne, VIC, Australia (G C Patton MD)

Correspondence to Prof Shane A Norris, SAMRC Developmental Pathways for Health Research Unit, Department of Paediatrics, University of the Witwatersrand, Johannesburg 2193, South Africa [shane.norris@wits.ac.za](mailto:shane.norris@wits.ac.za)





**Figure 1: Z scores for mean height and BMI of 54 million children and adolescents globally**  
 Z scores for mean height of girls (A) and boys (B). Z scores for mean BMI for girls (C) and boys (D). Individuals were born in 2000 and data were collected every year from age 5 years to 19 years. Each cell represents the Z score, derived from the WHO growth reference for a given age. Countries are ordered by region. For height, the heat map represents Z scores ranging from up to -3 (dark red) to above 3 (dark blue). For many countries, children and adolescents are shorter (stunted <2 Z score) than the WHO standard, as seen through the proliferation of red across the dial. For BMI, the heat map represents Z scores ranging from up to -3 (dark blue) to above 3 (dark red). For an increasing number of countries, children and adolescents are becoming overweight or obese (>1 Z score). BMI=body-mass index.



adolescence.<sup>36</sup> Economic hardship during preadolescent and adolescent periods is associated with short adult height.<sup>37</sup> Preference to have boys in China is associated with greater sex differences in height during childhood and adolescence than in the Philippines, where preference for boys exists to a lesser extent.<sup>38</sup> In Japan, day length predicts a regional gradient in height in late adolescence.<sup>39</sup> This mechanism might relate to regional gradients in photoperiod (ie, day length), which affects secretion of melatonin, inhibiting sexual and skeletal maturation, and inducing an increase in height.

In preschool children from Belarus and the USA, high BMI was associated with an increased velocity of upper body length and height in the following 4–5 years and with decreased height velocity during the next 5-year period.<sup>40</sup> Higher BMI in middle childhood (aged 6–8 years) was associated with earlier puberty and increased standing height and trunk length in adolescence. Data for the roles of specific nutrients or foods in adolescent height are scarce. In a cohort study of children aged 2–17 years in Iowa, USA, a high dietary intake of milk throughout childhood and adolescence (adjusted for nutrient adequacy, energy intake, and baseline socioeconomic status) was associated with greater height in adulthood than a low intake of milk.<sup>41</sup> Whether this association is specifically due to milk or to other attributes of the family or child is not known. Exposure to the Dutch famine of 1944–45 in young children during gestation or aged 1–2 years was associated with 3–4 cm deficits in adult height; however, inconsistent, smaller associations were seen for exposure at older ages (2–15 years).<sup>42</sup> Exposure to famines in Nigeria and Cambodia during adolescence reduced adult height more than exposure during younger ages (aged <12 years).<sup>43,44</sup> In Alabama (USA), early undernourishment delayed skeletal growth and menarche, and prolonged the period of growth in girls, with no difference in final adult height.<sup>45</sup> In Guatemala, receipt of a high protein-energy supplement improved nutrition, resulting in increased growth during the preschool period.<sup>46</sup> At adolescence, these children had greater height, muscle, and bone mass than did adolescents who had not received the supplement and, for boys only, skeletal maturation had advanced by 0.5 months.<sup>46</sup> A follow-up study in The Gambia explored the effect of calcium supplementation on the timing of puberty in children, and found a negative effect on attained height (panel 1).

Data from three decades of research in China suggest the interplay between socioeconomic context and the prevalence of stunting, thinness, and overweight or obesity over time. These findings highlight that linear growth restriction is reduced when environmental constraints are lifted (appendix p 1). These same environmental transitions have a substantial effect on the prevalence of overweight and obesity among adolescents. Given the consequences of undernourishment on health, such as an increased risk of NCDs

#### Panel 1: Long-term effects of calcium supplementation on pubertal timing and skeletal growth

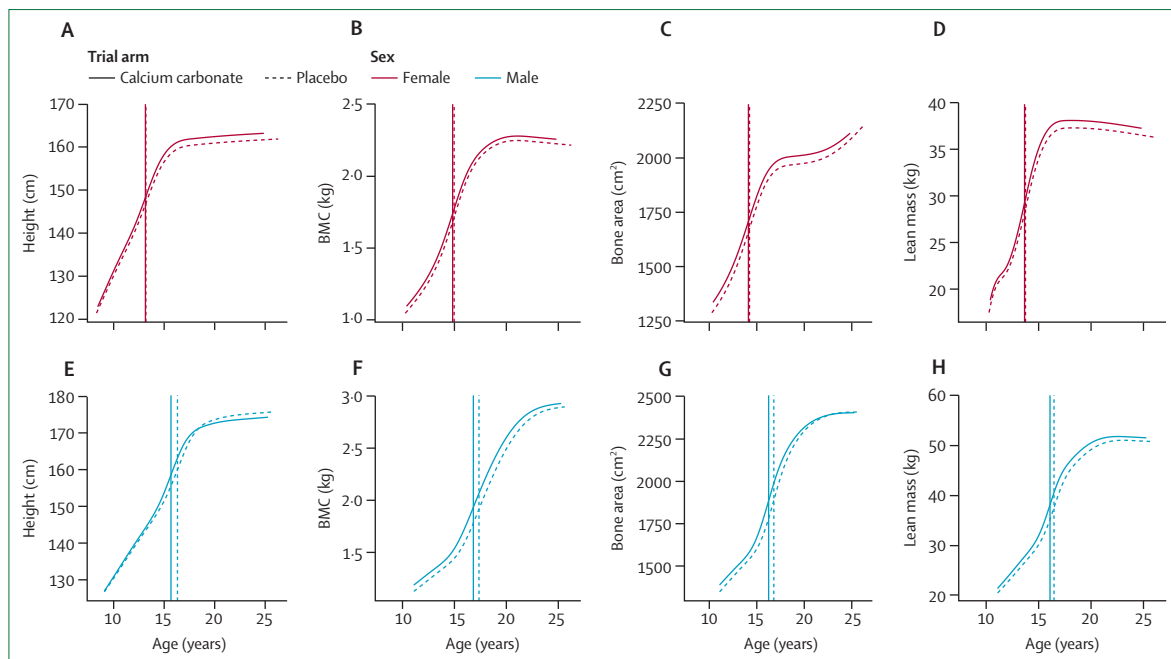
Most studies on calcium supplementation have been done in populations with adequate habitual calcium intakes. Therefore, in populations with extremely low calcium intake, interventions might be beneficial to skeletal development. Although most studies reported an initial increase in bone mineral density or size-adjusted bone mineral content (BMC), after a period of follow-up, the differences between intervention and control groups were attenuated.<sup>47–49</sup> To date, the study with the longest period of follow-up following supplementation is the 11-year follow-up study in The Gambia, in which calcium intakes were, on average, 300 mg/day. Pre-pubertal children aged 8–11 years were given 1000 mg of calcium or placebo for 5 days per week over 1 year.<sup>49</sup> The participants were then followed up until the end of growth, approximately 12 years later. At the end of the trial and 1 year and 2 years after supplementation, the calcium group had higher size-adjusted BMC at the midshaft radius than did the placebo group; the mean difference in size-adjusted BMC at the end of the trial was 4.6% (SE 0.9), reduced to 2.5% (1.3) by 2 years after supplementation. After modelling longitudinal growth for the entire follow-up period, group differences in pubertal timing, the velocity of growth, and final size were found, split by sex. In girls, no significant differences were found between the intervention groups in the amount of bone accrued or in the timing of puberty. In boys, pubertal timing (age at peak height velocity) was brought forward by approximately 7 months in participants in the calcium group and, although they transitioned through puberty at the same velocity as the placebo group, they stopped growing earlier (figure 2). Consequently, the boys in the calcium group were taller and had greater BMC in mid-adolescence compared with their counterparts in the placebo group; however, on average, they were 3.5 cm shorter at the end of the follow-up period. There were no significant group differences in bone outcomes at the end of growth, which could suggest that the supplementation had a negative effect on longitudinal growth with no direct benefit on bone mineralisation.

(eg, diabetes and hypertension), as well as the rising incidence of overweight and obesity, achieving a balance between optimising linear growth and avoiding the negative consequences of excessive weight gain is needed to reduce the burden of NCDs.

#### Body composition

During adolescence, changes in the proportions and distribution of bone, muscle, and fat form the foundation of metabolic and musculoskeletal health.<sup>50</sup> The timing of onset, duration, and velocity of these indicators of body composition are important for nutrition-sensitive interventions to optimise body composition trajectories. Body composition is commonly calculated with dual-energy x-ray absorptiometry measures of total body fat mass, fat free mass, and bone mineral content (BMC), which is a marker of bone strength and fracture risk. Lean mass is used as a surrogate of muscle mass and is derived by fat free mass minus BMC.<sup>51</sup> According to data from high-income countries, girls reach peak height velocity (PHV)—ie, the period of time with the fastest upward growth (8.3 cm/year for girls and 9.5 cm/year for boys)—at an average age of 11.8 years, which is earlier than boys. By contrast, boys reach PHV at an average age of 13.5 years.<sup>1,52</sup> Additionally, girls have lower total body

See Online for appendix



**Figure 2: Effect of calcium supplementation on distance curves for linear, bone, and muscle growth in adolescents from The Gambia**

Distance curves per year plotted for peak height (A), whole-body BMC (B), whole-body bone area (C), and lean mass (D) in female participants, and peak height (E), whole-body BMC (F), whole-body bone area (G), and lean mass (H) in male participants. The vertical line indicates age at peak accrual. Order of growth is height, lean mass, bone area, and BMC in both sexes. Male adolescents appear to continue accruing bone mineral after age 25 years. For more detail on this study, see panel 1. BMC=bone mineral content.

lean mass but greater fat mass than do boys.<sup>1,52</sup> Alongside greater lean mass, boys exhibit less total fat mass but similar (or greater in some cases) central fat mass than do girls.<sup>52</sup> These generalised values do not apply to all populations; for example, the age of PHV in The Gambia is approximately 16 years for boys and 13 years for girls (panel 1; figures 2, 3).

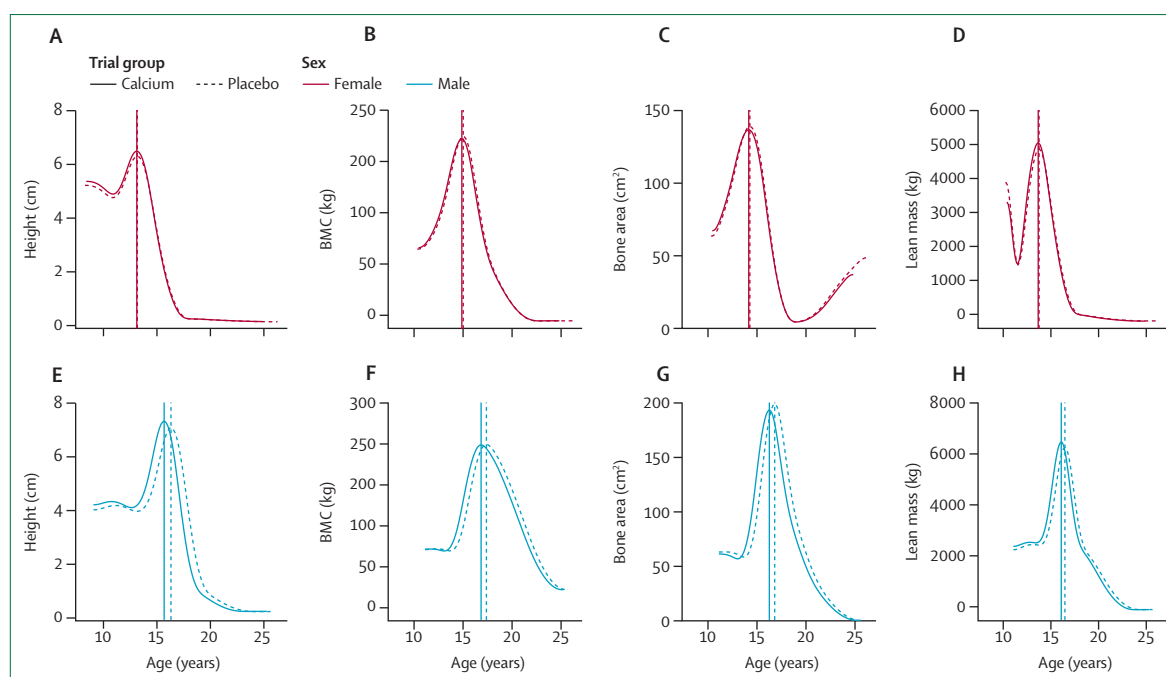
As height increases in girls and boys (for approximately 3 years after reaching PHV), there are corresponding increases in bone area and BMC.<sup>1</sup> Patterns of bone acquisition are relatively consistent between girls and boys; however, final BMC is higher<sup>1,53</sup> and reaches its plateau approximately 2 years later (at an average age of 18 years in girls and 20 years in boys) in boys than in girls.<sup>1</sup> Furthermore, ethnic differences are evident, with data suggesting that African American children have a higher BMC than do White children, despite similarities in height.<sup>53</sup> The onset and duration of puberty and nutrition can affect peak bone mass. A late onset of puberty has been associated with 10% decrease in bone mineral density and an increased risk of hip fracture in later life.<sup>54,55</sup>

Lean mass increases in girls and boys during adolescence; however, the rate of lean mass acquisition is higher in boys.<sup>54</sup> On average, girls attain stable, adult levels of lean mass at approximately 15–16 years of age.<sup>45,54</sup> In boys, steady acquisition of lean mass occurs from approximately 8–18 years of age, with more rapid increases at 12–15 years.<sup>50,56</sup> Independent of

chronological age, puberty is associated with an average 1.14 kg/year increase in absolute fat mass in girls.<sup>56,57</sup> In boys, absolute fat mass is relatively stable over the pubertal period, which results in a decrease in body fat percentage during adolescence as a result of rapid increases in lean mass.<sup>56</sup> There are no significant sex differences in peripheral fat mass in the upper body compartments (ie, arm and torso), suggesting that differences in lower body (ie, legs) fat mass are the primary contributor to the sexual dimorphism in adiposity.<sup>52</sup> In general, boys have been shown to have higher amounts of visceral fat mass in later adolescence than do girls.<sup>52</sup> Panel 2 and figure 4 detail the trajectories of body composition in adolescents from South Africa, and show the altered trajectories of fat mass in individuals who have obesity as young adults. These results suggest that efforts to prevent obesity need to start earlier in adolescence (age 9–11 years). Furthermore, given the variations in timing and duration of puberty between girls and boys, interventions should be tailored by sex.

### Cardiorespiratory fitness

High cardiorespiratory fitness (ie, reduced oxygen uptake during exercise, as measured by a maximal oxygen consumption test) attained during adolescence might decrease risk of cardiovascular disease in adulthood. A 2018 review concluded that, regardless of sex, cardiorespiratory fitness in childhood and



**Figure 3: Effect of calcium supplementation on velocity curves for linear, bone, and muscle growth in adolescents from The Gambia**

Measurement velocity curves per year plotted for peak height (A), whole-body BMC (B), whole-body bone area (C), and lean mass (D) in female participants, and peak height (E), whole-body BMC (F), whole-body bone area (G), and lean mass (H) in male participants. Velocity curves show the offset in peak velocity for each measure. The vertical line indicates age at peak accretion. Order of growth is height, lean mass, bone area, and BMC in both sexes. Age at peak height velocity (ie, onset of puberty) was 13.3 years (girls) and 14.4 years (boys) in the calcium group and 13.2 years (girls) and 14.8 years (boys) in the placebo group. For more detail on this study, see panel 1. BMC=bone mineral content.

adolescence was associated with decreased fat mass over time.<sup>58</sup> Additionally, analyses of the Swedish military conscription register indicated that low cardiorespiratory fitness at conscription strongly predicted being on a disability pension in later life due to ischaemic heart disease, cerebrovascular diseases, or heart failure.<sup>59,60</sup> Cardiorespiratory fitness in adolescence predicts a favourable risk factor profile for cardiovascular disease during adulthood, including reduced blood pressure, a favourable lipid profile, and reduced plasma fasting glucose concentrations.<sup>61</sup> Although cardiorespiratory fitness has a strong genetic component, high amounts of moderate-to-vigorous activity during adolescence have been associated with increased cardiorespiratory fitness.<sup>62,63</sup> The beneficial effects of cardiorespiratory fitness on body composition and adiposity, as well as the early establishment of healthy physical activity habits, could be jointly responsible for these health benefits in the long term (appendix pp 2–4).

### Neurodevelopment

The brain reaches approximately 90% of its adult size by age 6 years, but the grey and white matter subcomponents continue to undergo dynamic changes throughout adolescence.<sup>5</sup> Considerable brain growth and development occur during adolescence in the construction and strengthening of regional neurocircuitry, with rewiring accomplished through dendritic pruning

and myelination. In particular, the prefrontal cortex continually reconstructs, consolidates, and matures.<sup>64</sup> The adolescent brain is characterised by neuroplasticity, which is the ability of neural networks to reorganise in response to different social, learning, and nutritional environments.<sup>65</sup> On one hand, plasticity enables learning and adaptation; on the other hand, it brings a susceptibility to adverse environmental exposures, such as poor nutrition and stressful experiences.<sup>66,67</sup> This susceptibility raises the possibility of lasting changes in neurocircuitry, perhaps one explanation for why many psychiatric disorders first manifest in adolescence.<sup>64</sup>

Adolescent nutrition can have direct and indirect effects on the maturing brain. The severe undernutrition of anorexia nervosa can interrupt pubertal development, with impairment of cognitive flexibility and working memory.<sup>68</sup> Extended undernutrition results in a reduction in grey and white matter of the brain,<sup>68,69</sup> especially the frontoparietal network, with effects on higher executive functions.<sup>68</sup> These changes are also associated with poor emotional regulation, poor processing of social cues, and altered responses to reward.<sup>68,70</sup> Changes in brain structure in people with non-chronic anorexia nervosa seem largely reversible in response to improved nutrition and weight gain, with one study showing that the volume of grey and white matter normalised within 2–8 years of remission;<sup>69</sup> however, there might be less reversibility in chronic disorders.

### Panel 2: Body composition of adolescents from Soweto, South Africa

As part of the Birth to Twenty Plus Birth Cohort, longitudinal sub-cohort data on the body composition of children born in 1990 in Soweto, Johannesburg, South Africa, were derived from dual-energy x-ray absorptiometry. Data from 3067 scans, performed in 174 girls and 196 boys annually from age 9 years to 18 years, highlighted variation in timing and development of body composition between the sexes (figure 3). The peak velocity for bone mineral content (BMC) and fat-free soft-tissue mass (surrogate for lean mass) in boys occurred significantly later than in girls (BMC 14.6 years vs 12.2 years; fat-free soft-tissue mass 14.3 years vs 11.4 years). By contrast, peak velocity for fat mass occurred earlier in boys (10.9 years vs 13.9 years), although the magnitude of the mass and velocity for fat is significantly less in boys than in girls. However, after standardising for puberty, similar patterns for bone mass accrual were evident in boys and girls, and occurred approximately 1 year following peak height velocity (PHV), with boys having greater bone mass accrual. This finding was similar for lean mass, but not for fat mass. The peak fat mass velocity in boys occurred approximately 2.0 years before PHV, whereas for girls it was 2.5 years after, with significant differences in fat mass accrual between the sexes. This result aligns with the deposition of post-menarche fat mass in female adolescents in preparation for pregnancy. We know from longitudinal data that over 40% of female participants and 15% of male participants in the Birth to Twenty Plus Birth Cohort had overweight or obesity by adulthood. Using body-mass index in young adulthood (aged 20 years) to classify overweight or obesity, we examined the adolescent profile of fat mass accrual in young adults with or without overweight or obesity (figure 3). Unlike in adolescents without overweight, male adolescents with overweight or obesity have similar profiles to female adolescents with or without overweight or obesity in terms of peak fat mass velocity occurring after PHV. These data suggest that prevention should start in early adolescence to minimise excess accumulation of fat mass.

Excessive consumption of energy-dense foods can alter self-regulatory processes by affecting brain function.<sup>71</sup> High-fat and high-sugar diets might affect neurodevelopment through alterations in two neurotransmitter systems: dopamine-mediated reward signalling and inhibitory neurotransmission controlled by  $\gamma$ -aminobutyric acid.<sup>71</sup> Consequently, modifications of these two systems during adolescence could lead to dysregulated eating and impulsive behaviours.

Neurodevelopment seems to be linked to the maturation of other biological systems. For example, there appears to be a bidirectional communication between the gut microbiome and the brain. Dysbiosis (ie, change in the gut microbiome composition with metabolic and inflammatory effects) seems to affect neural function in

vitro, in vivo, and in human studies, raising the possibility of neurodevelopmental consequences.<sup>72</sup> Additionally, musculoskeletal growth has consequences for neurocognitive development, with absence of the bone-derived hormone, osteocalcin, linked to anxiety and depression, as well as inhibited exploration, spatial learning, and memory.<sup>73,74</sup>

### Immune system development

In infancy, passively acquired maternal immunity and breastfeeding provide protection against pathogens. Both innate (eg, neutrophils, monocytes, macrophages, and dendritic cells) and adaptive (eg, B and T lymphocytes) components of the immune system deliver tempered responses to pathogens and commensal microorganisms. In childhood, this pattern changes to provide more robust innate responses to pathogens and to allow for the development of protective immunological memory to pathogens through memory B and T cells, as well as pathogen-specific antibody responses. By late childhood, adult-like innate and adaptive responses are typically observed: the number of memory B and T cells reach adult numbers, and the output of naive T cells by the thymus diminishes substantially as immune memory to childhood infectious diseases has developed.<sup>75</sup> Therefore, adolescents have adult-like innate and adaptive immune responses, with adult-like sex differences in these responses.<sup>76</sup> Although some sex differences result from X-linked immune system genes and are seen throughout life, the differences that develop after puberty are caused primarily by the different actions of androgens and oestrogen on immune cells.<sup>77</sup> Sex can also influence the development of the immune system due to gender-specific differences in behaviour that affect exposure to environmental factors, including diet.<sup>76,78–80</sup>

Thus, nutritional status might affect adolescent health in a sex-specific manner, in which these effects are mediated by immune function. For example, as children, girls have a more robust adaptive immune response to infection than do boys and, consequently, lower mortality rates from infectious disease.<sup>81–83</sup> However, these mortality rates are similar for adolescent girls and boys, and are higher in adult women than in adult men, highlighting

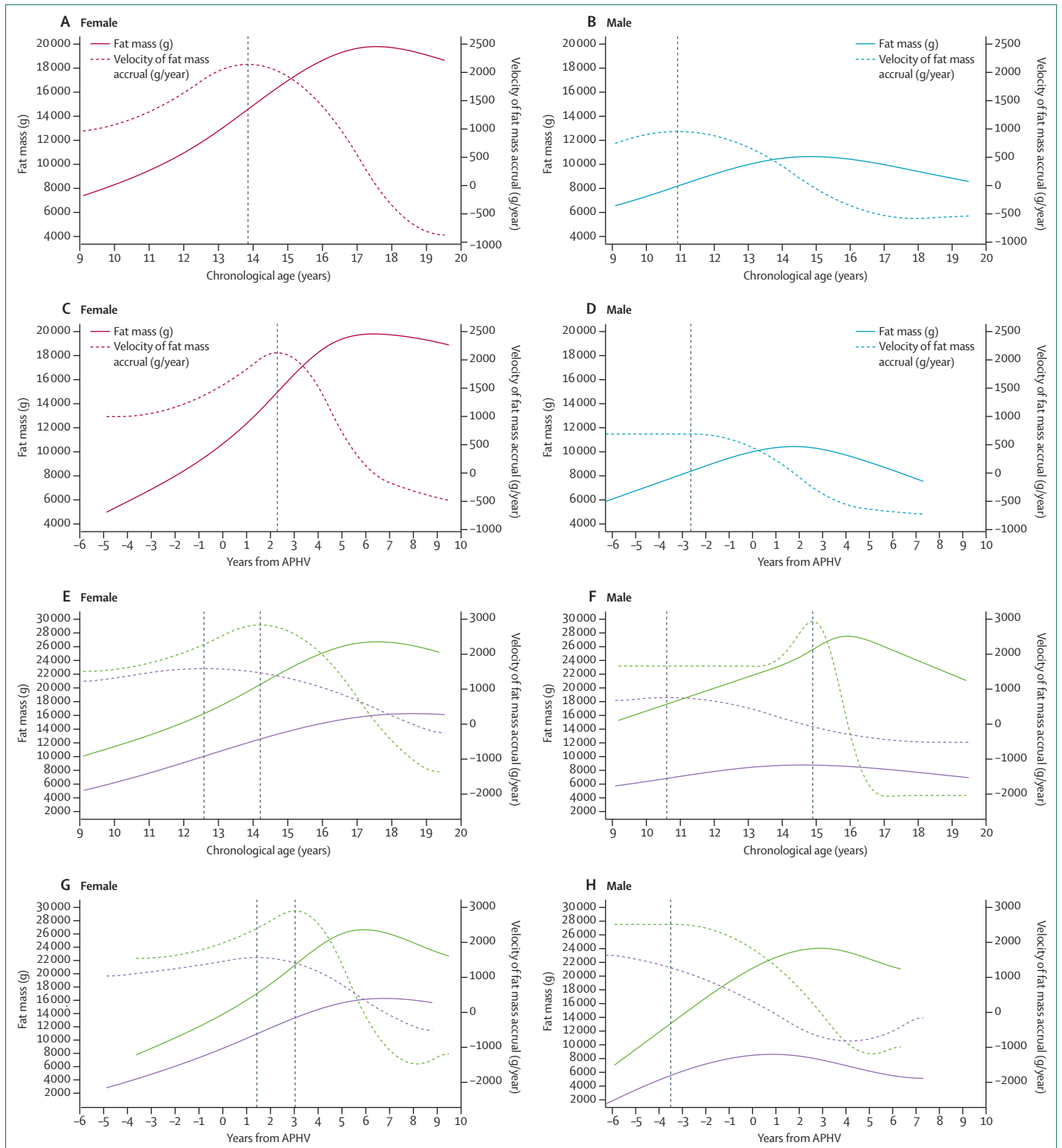
### Figure 4: Longitudinal modelling of fat mass and velocity of fat mass accrual by chronological age and APHV

Whole-body fat mass (solid line) and velocity of fat mass accrual (dashed line) in female and male adolescents by chronological age (A, B) and by years from APHV (C, D) from the Birth to Twenty Plus Birth Cohort in South Africa. Longitudinal modelling of whole-body fat mass and velocity of fat mass accrual in female and male adolescents by chronological age (E, F) and years from APHV (G, H), stratified by individuals with (green) or without (purple) overweight or obesity at age 20 years. Unlike in adolescents with healthy weight, overweight and obesity in male adolescents have similar profiles to female adolescents, with peak velocity of fat mass accrual occurring after peak height velocity. In individuals with overweight or obesity, fat mass accrues early in adolescence and continues to increase until late adolescence. For more detail on this study, see panel 2. APHV=age at peak height velocity.



the impact of nutrition and social influences on biology (appendix p 4). In populations with a high HIV prevalence in adolescents, infection exacerbates undernutrition,

which can further impair immunity. Dietary deficiencies in both macronutrients (eg, too little dietary protein) and micronutrients (eg, deficiencies in vitamins B12, C, and D)



can impair most aspects of immune function, including compromising epithelial barriers (particularly relevant in HIV and other sexually transmitted infections) and impairing the development and function of innate and adaptive immune cells, with the predictable result of increasing the severity of common infectious diseases. For example, in adolescents with a dietary deficiency, macrophages and neutrophils have a diminished ability to take up and kill pathogenic bacteria, lymphocyte cell counts in the spleen and lymph nodes are reduced, and development of memory T and B cells is impaired.<sup>84</sup> One example is seen with protein-energy malnutrition, which particularly impairs the T-cell arm of adaptive immunity by diminishing thymic function to reduce the supply of naive T cells to peripheral lymphoid tissue. Therefore, this reduction might impair development of immunological memory, leading to an increased risk of death from infectious disease in childhood.<sup>84</sup> Nevertheless, studies in adolescence are scarce. Nutritional interventions that support resistance to infectious disease could benefit girls and boys.

Chronic inflammation caused by activation of the immune system during adolescence can decrease linear growth, partly due to the activity of proinflammatory cytokines (including IL-1 $\beta$ , TNF $\alpha$ , and IL-6) on the growth plate of long bones.<sup>85</sup> Obesity in adolescence stimulates chronic inflammation that increases the risk of various NCDs during adulthood, including fatty liver disease, type 2 diabetes (also in adolescence; appendix pp 2–4), and cardiovascular disease.<sup>86</sup> The cause of inflammation in obesity is complex, probably involving activation of innate immune cells in adipose tissue depots because of metabolic or cellular stress. The mechanism might involve diet-induced disruption of the intestinal barrier, perhaps initially causing changes to the intestinal microbiome that lead to increased exposure to microbial products (eg, bacterial lipopolysaccharides), which trigger systemic or local inflammation in abdominal adipose tissue.<sup>87</sup> During adolescence, the inflammation observed in obesity is associated with increased risk of chronic inflammatory diseases, including asthma.<sup>88</sup> Thus, preventing or treating obesity in adolescence could have clinically significant benefits by preventing immune-mediated exacerbations of infectious or chronic inflammatory diseases.

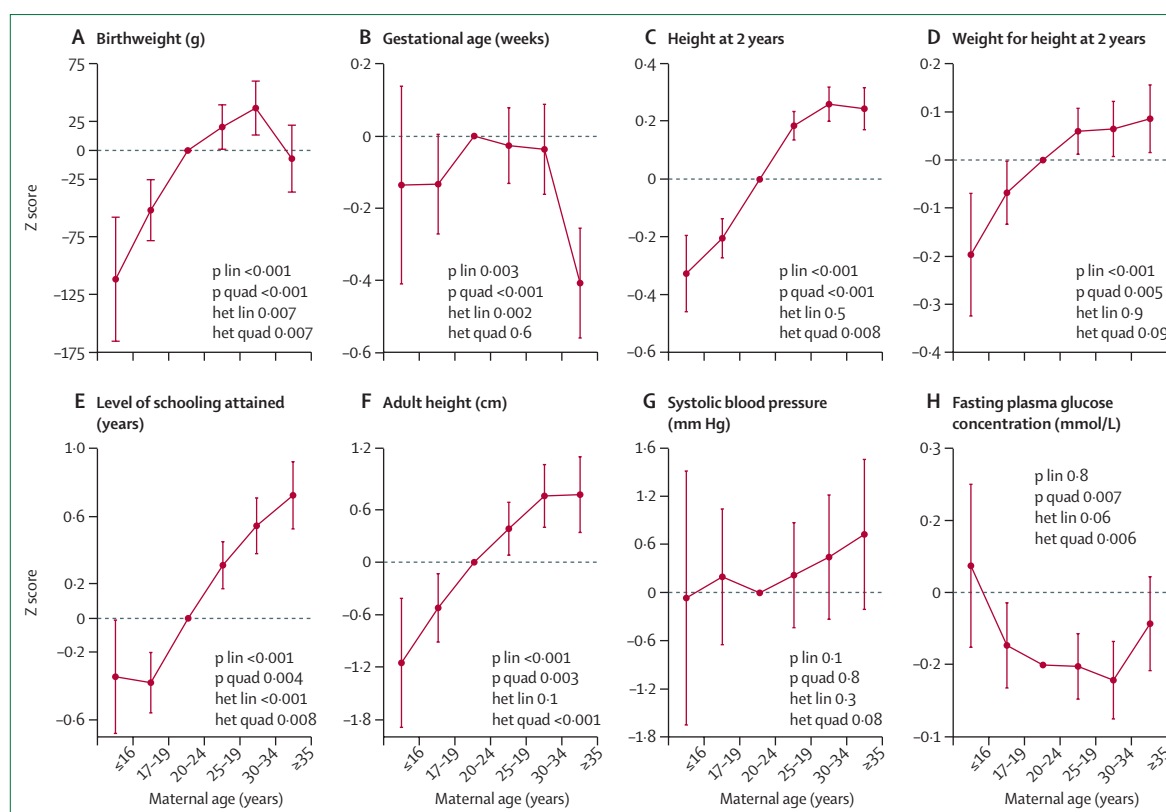
### Adolescent pregnancy, nutrition, and intergenerational effects

Sexual maturation and relationships during adolescence set the scene for future parenthood. Reproductive success and optimal upbringing of children are best achieved after parents have largely completed the physical, mental, social, and emotional development of adolescence. Nevertheless, WHO estimates that around 16 million adolescent girls become mothers every year in LMICs.<sup>89</sup> Although the rate of adolescent pregnancy has decreased globally, an increasing number of adolescents overall

means that the absolute number of adolescent pregnancies is increasing, particularly in settings with the greatest nutritional disadvantage.

The occurrence of adolescent pregnancies varies greatly across regions and within countries, but the number tends to be high in groups facing nutritional disadvantage, including rural and Indigenous populations.<sup>90</sup> These pregnancies occur more frequently in socioeconomically disadvantaged populations and among girls with unstable relationships and financial resources.<sup>89</sup> Adolescent pregnancy compounds disadvantages for girls by leaving education, limiting life chances (eg, employment), and perpetuating the cycle of poverty.<sup>91</sup> Neonates of adolescent mothers in LMICs are at increased risk of low birthweight and short birth length, at least partly because of maternal stunting and competition for nutrients between the mother and fetus during pregnancy.<sup>92,93</sup> Neonates of adolescent mothers are also at increased risk of preterm delivery,<sup>94,95</sup> with heightened risks for poor childhood growth and nutritional status, low educational attainment, and increased fasting glucose concentrations in adulthood.<sup>94,95</sup> These risks are most pronounced among children of the youngest adolescent mothers (figure 5),<sup>95</sup> and are likely to result from the biological immaturity of their mothers and their socioeconomic context.<sup>94</sup> Even though there are almost no data available from LMICs, scarce evidence suggests that adolescent fathers have similar offspring outcomes to adolescent mothers in terms of low birthweight, increased risk of preterm birth and infant mortality, and poor childhood health overall.<sup>97</sup>

When considered in the context of pregnancy and parenthood, the growing burden of adolescent malnutrition is of concern.<sup>98</sup> Undernutrition, food insecurity, and poor quality, monotonous diets remain common, especially in sub-Saharan Africa and south Asia. Gender inequality in nutrition often emerges in adolescence.<sup>99</sup> Both undernutrition and overweight or obesity in mothers before conception or during pregnancy predict altered growth and health in their offspring. Maternal height is positively associated with birthweight, adult stature, and educational attainment and income in the offspring.<sup>100</sup> Low maternal folate, vitamin B12, and vitamin D status in pregnancy have been associated with reduced cognitive function and changes in glucose and insulin concentrations in offspring, which indicate an increased future risk of diabetes.<sup>101–103</sup> Mothers with overweight or obesity are at an increased risk of developing gestational diabetes.<sup>104</sup> In turn, gestational glucose intolerance risks congenital malformations in the fetus, increasing the child's risk of increased adiposity and insulin resistance, elevated blood pressure, and early onset type 2 diabetes.<sup>105,106</sup> Although none of these associations are specific to adolescent pregnancy, stunting, micronutrient deficiencies, and overweight or obesity among adolescents all persist into later pregnancies, and shape fetal programming, development in early life, and cardiometabolic health of the offspring in the long term.



**Figure 5: Associations between maternal age and outcomes in offspring**

Z scores provided for birthweight, gestational age, height at 2 years, weight for height at 2 years, years of schooling attained, adult height, adult systolic blood pressure, and adult fasting plasma glucose concentration. Data taken from the COHORTS collaboration of five birth cohorts from low-income and middle-income countries.<sup>96</sup> For each maternal age group, the amount (95% CI) by which the outcome differs from offspring of mothers aged 20–24 years was obtained using linear regression of a pooled dataset from 19 403 women from five cohorts in Brazil, Guatemala, India, the Philippines, and South Africa, adjusted for offspring sex, maternal height, parity, marital status, schooling, wealth, race (Brazil and South Africa), urbanicity (the Philippines), breastfeeding duration (postnatal outcomes only), and offspring age (adult outcomes only). p values were derived using maternal age as a continuous variable. p lin is the p value from a test for linear trends in the outcome with maternal age; p quad is the p value from a test for quadratic trends; het lin is the F test p value for heterogeneity in the linear trends between the five cohorts; and het quad is the p value for heterogeneity in the quadratic trends.

There is growing research interest into whether paternal nutritional status has similar intergenerational effects through epigenetic changes in sperm, although most available evidence currently comes from animal studies.<sup>107,108</sup> In rodents, changes in paternal diet or exposure to stress between weaning and sexual maturity have been shown to alter the metabolism of offspring (ie, glucose tolerance and lipid metabolism), stress responsiveness, and mood. Although other epigenetic mechanisms could be involved, micro RNAs carried in sperm are strong candidates for messengers that link paternal nutritional state before conception to offspring phenotype.<sup>107</sup>

## Conclusion

Biological development during adolescence involves a finely tuned orchestration of maturation of different physiological systems, with varying onsets and durations. Furthermore, this orchestration differs between girls and boys. Although undernutrition and overnutrition have

diverse and different effects on biological development during adolescence, research has been scarce and there is still much to learn, particularly around adolescent growth and development in LMICs. Future studies into adolescent growth and nutrition should move beyond a focus on a single physiological system, towards integrated system-wide approaches over the lifecourse. Such research should include a better understanding of the relationships between pubertal development and nutrition, physical activity, and metabolic state, which could give rise to strategies that optimise growth and prevent diseases (eg, type 2 diabetes, osteoporosis and other musculoskeletal disorders, and cardiovascular disease) in later life. At a time when a rapid nutrition transition is shifting diets for most young people globally, improving adolescent nutrition provides an opportunity to shape the health and wellbeing of this generation and the next.

## Contributors

SAN, EAF, and GCP conceptualised and coordinated the paper, and incorporated all revisions until submission. SAN, YD, CF, AP, and KAW

contributed figures to the paper. All authors contributed to writing designated sections of the paper and editing the paper and have reviewed and approved the final version of the manuscript.

#### Declaration of interests

AP declares grants from Medical Research Council (UK) during the conduct of The Gambia study. KAW declares personal fees from Abbott Laboratories, Pfizer Consumer Healthcare, and *Journal of Bone and Mineral Research*, outside of the submitted work. All other authors declare no competing interests.

#### Acknowledgments

This work received funding support from Fondation Botnar and the Wellcome Trust. Neither organisation played any role in writing the manuscript or the decision to submit for publication. We thank Majid Ezzati for sharing the data for figure 1. We thank Lukhanyo Nyati for assisting with the modelling of body composition data from the Birth to Twenty Plus Cohort. We thank the principal investigators of the COHORTS collaboration in Brazil, India, Philippines, Guatemala, and South Africa for permission to show the data in figure 4. SAN is supported by the DSI-NRF Centre of Excellence in Human Development at the University of the Witwatersrand and the South African Medical Research Council. GCP is supported by a National Health and Medical Research Council Senior Principal Research Fellowship. AP and KAW received funding for The Gambian studies described in panel 1 from the UK Medical Research Council (programme codes U105960371 and U123261351) and the UK Department for International Development, under the Medical Research Council–Department for International Development Concordat agreement. TR is supported by a Wellcome Trust Intermediate Fellowship In Public Health and Tropical Medicine (211374/Z/18/Z) and receives salary support from Joint Global Health Trials within the UK Department for International Development, Wellcome Trust, and the UK Medical Research Council grant (MR/P006965/1). MMB is supported by a grant from the National Institutes of Health (R01 DK106424).

Editorial note: the *Lancet* Group takes a neutral position with respect to territorial claims in published figures and institutional affiliations.

#### References

- Baxter-Jones ADG, Faulkner RA, Forwood MR, Mirwald RL, Bailey DA. Bone mineral accrual from 8 to 30 years of age: an estimation of peak bone mass. *J Bone Miner Res* 2011; **26**: 1729–39.
- Lassi ZS, Mansoor T, Salam RA, Bhutta SZ, Das JK, Bhutta ZA. Review of nutrition guidelines relevant for adolescents in low- and middle-income countries. *Ann N Y Acad Sci* 2017; **1393**: 51–60.
- Kyu HH, Pinho C, Wagner JA, et al. Global and national burden of diseases and injuries among children and adolescents between 1990 and 2013: findings from the Global Burden of Disease 2013 Study. *JAMA Pediatr* 2016; **170**: 267–87.
- Abarca-Gómez L, Abdeen ZA, Hamid ZA, et al. Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128·9 million children, adolescents, and adults. *Lancet* 2017; **390**: 2627–42.
- Patton GC, Sawyer SM, Santelli JS, et al. Our future: a *Lancet* commission on adolescent health and wellbeing. *Lancet* 2016; **387**: 2423–78.
- Chang H-P, Yang S-F, Wang S-L, Su P-H. Associations among IGF-1, IGF2, IGF-1R, IGF-2R, IGFBP-3, insulin genetic polymorphisms and central precocious puberty in girls. *BMC Endocr Disord* 2018; **18**: 66.
- InterLACE Study Team. Variations in reproductive events across life: a pooled analysis of data from 505 147 women across 10 countries. *Hum Reprod* 2019; **34**: 881–93.
- Moodie JL, Campisi SC, Salena K, Wheatley M, Vandermorris A, Bhutta ZA. Timing of pubertal milestones in low- and middle-income countries: a systematic review and meta-analysis. *Adv Nutr* 2020; **11**: 951–59.
- Corley RP, Beltz AM, Wadsworth SJ, Berenbaum SA. Genetic influences on pubertal development and links to behavior problems. *Behav Genet* 2015; **45**: 294–312.
- Goddings A-L, Viner RM, Mundy L, et al. Growth and adrenarche: findings from the CATS observational study. *Arch Dis Child* 2021; **106**: 967–74.
- Trikudanathan S, Pedley A, Massaro JM, et al. Association of female reproductive factors with body composition: the Framingham Heart Study. *J Clin Endocrinol Metab* 2013; **98**: 236–44.
- Cheng TS, Day FR, Lakshman R, Ong KK. Association of puberty timing with type 2 diabetes: a systematic review and meta-analysis. *PLoS Med* 2020; **17**: e1003017.
- Brix N, Ernst A, Lauridsen LLB, et al. Maternal pre-pregnancy obesity and timing of puberty in sons and daughters: a population-based cohort study. *Int J Epidemiol* 2019; **48**: 1684–94.
- Aghaee S, Deardorff J, Greenspan LC, Quesenberry CP Jr, Kushi LH, Kubo A. Breastfeeding and timing of pubertal onset in girls: a multiethnic population-based prospective cohort study. *BMC Pediatr* 2019; **19**: 277.
- Günther ALB, Karaolis-Danckert N, Kroke A, Remer T, Buyken AE. Dietary protein intake throughout childhood is associated with the timing of puberty. *J Nutr* 2010; **140**: 565–71.
- Remer T, Shi L, Buyken AE, Maser-Gluth C, Hartmann MF, Wudy SA. Prepubertal adrenarchal androgens and animal protein intake independently and differentially influence pubertal timing. *J Clin Endocrinol Metab* 2010; **95**: 3002–09.
- Rahimi A, Rahimi M, Norouzy A, et al. Association of dietary pattern and body size with early menarche among elementary school girls in west of Iran. *Int J Pediatr* 2019; **7**: 10583–93.
- Cheng HL, Raubenheimer D, Steinbeck K, Baur L, Garnett S. New insights into the association of mid-childhood macronutrient intake to pubertal development in adolescence using nutritional geometry. *Br J Nutr* 2019; **122**: 274–83.
- Carwile JL, Willett WC, Spiegelman D, et al. Sugar-sweetened beverage consumption and age at menarche in a prospective study of US girls. *Hum Reprod* 2015; **30**: 675–83.
- Noonan KJ, Hunziker EB, Nessler J, Buckwalter JA. Changes in cell, matrix compartment, and fibrillar collagen volumes between growth-plate zones. *J Orthop Res* 1998; **16**: 500–08.
- Wilsman NJ, Farnum CE, Leiferman EM, Fry M, Barreto C. Differential growth by growth plates as a function of multiple parameters of chondrocytic kinetics. *J Orthop Res* 1996; **14**: 927–36.
- Day TF, Guo X, Garrett-Beal L, Yang Y. Wnt/beta-catenin signaling in mesenchymal progenitors controls osteoblast and chondrocyte differentiation during vertebrate skeletogenesis. *Dev Cell* 2005; **8**: 739–50.
- Lampl M, Schoen M. How long bones grow children: mechanistic paths to variation in human height growth. *Am J Hum Biol* 2017; **29**: e22983.
- Kobayashi T, Kronenberg H. Minireview: transcriptional regulation in development of bone. *Endocrinology* 2005; **146**: 1012–17.
- Nilsson O, Marino R, De Luca F, Phillip M, Baron J. Endocrine regulation of the growth plate. *Horm Res* 2005; **64**: 157–65.
- Callahan BP, Wang C. Hedgehog cholesterolysis: specialized gatekeeper to oncogenic signaling. *Cancers (Basel)* 2015; **7**: 2037–53.
- Koren N, Simsa-Maziel S, Shahar R, Schwartz B, Monsonego-Ornan E. Exposure to omega-3 fatty acids at early age accelerate bone growth and improve bone quality. *J Nutr Biochem* 2014; **25**: 623–33.
- Mason EF, Rathmell JC. Cell metabolism: an essential link between cell growth and apoptosis. *Biochim Biophys Acta* 2011; **1813**: 645–54.
- Zhong L, Huang X, Karperien M, Post JN. The regulatory role of signaling crosstalk in hypertrophy of MSCs and human articular chondrocytes. *Int J Mol Sci* 2015; **16**: 19225–47.
- Leroy JL, Frongillo EA, Dewan P, Black MM, Waterland RA. Can children catch up from the consequences of undernourishment? Evidence from child linear growth, developmental epigenetics, and brain and neurocognitive development. *Adv Nutr* 2020; **11**: 1032–41.
- Ulijaszek SJ. Secular trends in growth: the narrowing of ethnic differences in stature. *Nutr Bull* 2001; **26**: 43–51.
- Perkins JM, Subramanian SV, Davey Smith G, Özaltn E. Adult height, nutrition, and population health. *Nutr Rev* 2016; **74**: 149–65.
- McCullough JM, McCullough CS. Age-specific variation in the secular trend for stature: a comparison of samples from industrialized and nonindustrialized regions. *Am J Phys Anthropol* 1984; **65**: 169–80.



- 34 Haas JD, Campirano F. Interpopulation variation in height among children 7 to 18 years of age. *Food Nutr Bull* 2006; 27 (suppl): S212–23.
- 35 Benefice E, Luna Monrroy SJ, Lopez Rodriguez RW, Ndiaye G. Fat and muscle mass in different groups of pre-pubertal and pubertal rural children. Cross-cultural comparisons between Sahelian (rural Senegal) and Amazonian (Beni River, Bolivia) children. *Ann Hum Biol* 2011; 38: 500–07.
- 36 de Wilde JA, van Dommelen P, van Buuren S, Middelkoop BJC. Height of south Asian children in the Netherlands aged 0-20 years: secular trends and comparisons with current Asian Indian, Dutch and WHO references. *Ann Hum Biol* 2015; 42: 38–44.
- 37 Isasi CR, Jung M, Parrinello CM, et al. Association of childhood economic hardship with adult height and adult adiposity among Hispanics/Latinos. The HCHS/SOL Socio-Cultural Ancillary Study. *PLoS One* 2016; 11: e0149923.
- 38 Song S, Burgard SA. Does son preference influence children's growth in height? A comparative study of Chinese and Filipino children. *Popul Stud (Camb)* 2008; 62: 305–20.
- 39 Yokoya M, Shimizu H, Higuchi Y. Geographical distribution of adolescent body height with respect to effective day length in Japan: an ecological analysis. *PLoS One* 2012; 7: e50994.
- 40 Aris IM, Rifas-Shiman SL, Zhang X, et al. Association of BMI with linear growth and pubertal development. *Obesity (Silver Spring)* 2019; 27: 1661–70.
- 41 Marshall TA, Curtis AM, Cavanaugh JE, Warren JJ, Levy SM. Higher longitudinal milk intakes are associated with increased height in a birth cohort followed for 17 years. *J Nutr* 2018; 148: 1144–49.
- 42 Portrait FRM, van Wingerden TF, Deeg DJH. Early life undernutrition and adult height: the Dutch famine of 1944–45. *Econ Hum Biol* 2017; 27: 339–48.
- 43 Akresh R, Bhalotra S, Leone M, Osili UO. War and stature: growing up during the Nigerian civil war. *Am Econ Rev* 2012; 102: 273–77.
- 44 Holmqvist G, Pereira A. Famines and stunting: are adolescents the hardest hit? March 27, 2017. <https://blogs.unicef.org/evidence-for-action/famines-and-stunting-are-adolescents-the-hardest-hit/> (accessed Sept 12, 2020).
- 45 Dreizen S, Spirakis CN, Stone RE. A comparison of skeletal growth and maturation in undernourished and well-nourished girls before and after menarche. *J Pediatr* 1967; 70: 256–63.
- 46 Rivera JA, Martorell R, Ruel MT, Habicht JP, Haas JD. Nutritional supplementation during the preschool years influences body size and composition of Guatemalan adolescents. *J Nutr* 1995; 125 (suppl): 1068S–77S.
- 47 Dibba B, Prentice A, Ceesay M, Stirling DM, Cole TJ, Poskitt EM. Effect of calcium supplementation on bone mineral accretion in Gambian children accustomed to a low-calcium diet. *Am J Clin Nutr* 2000; 71: 544–49.
- 48 Dibba B, Prentice A, Ceesay M, et al. Bone mineral contents and plasma osteocalcin concentrations of Gambian children 12 and 24 mo after the withdrawal of a calcium supplement. *Am J Clin Nutr* 2002; 76: 681–86.
- 49 Ward KA, Cole TJ, Laskey MA, et al. The effect of prepubertal calcium carbonate supplementation on skeletal development in Gambian boys—a 12-year follow-up study. *J Clin Endocrinol Metab* 2014; 99: 3169–76.
- 50 Siervogel RM, Demerath EW, Schubert C, et al. Puberty and body composition. *Horm Res* 2003; 60 (suppl 1): 36–45.
- 51 Ward K. Musculoskeletal phenotype through the life course: the role of nutrition. *Proc Nutr Soc* 2012; 71: 27–37.
- 52 Katzmarzyk PT, Shen W, Baxter-Jones A, et al. Adiposity in children and adolescents: correlates and clinical consequences of fat stored in specific body depots. *Pediatr Obes* 2012; 7: e42–61.
- 53 McCormack SE, Cousminer DL, Chesi A, et al. Association between linear growth and bone accrual in a diverse cohort of children and adolescents. *JAMA Pediatr* 2017; 171: e171769.
- 54 Kuh D, Muthuri SG, Moore A, et al. Pubertal timing and bone phenotype in early old age: findings from a British birth cohort study. *Int J Epidemiol* 2016; 45: 1113–24.
- 55 Javaid MK, Eriksson JG, Kajantie E, et al. Growth in childhood predicts hip fracture risk in later life. *Osteoporos Int* 2011; 22: 69–73.
- 56 Veldhuis JD, Roemmich JN, Richmond EJ, et al. Endocrine control of body composition in infancy, childhood, and puberty. *Endocr Rev* 2005; 26: 114–46.
- 57 Rodríguez G, Moreno LA, Blay MG, et al. Body composition in adolescents: measurements and metabolic aspects. *Int J Obes* 2004; 28 (suppl 3): S54–58.
- 58 Mintjens S, Menting MD, Daams JG, van Poppel MNN, Roseboom TJ, Gemke RBJ. Cardiorespiratory fitness in childhood and adolescence affects future cardiovascular risk factors: a systematic review of longitudinal studies. *Sports Med* 2018; 48: 2577–605.
- 59 Henriksson H, Henriksson P, Tynelius P, et al. Cardiorespiratory fitness, muscular strength, and obesity in adolescence and later chronic disability due to cardiovascular disease: a cohort study of 1 million men. *Eur Heart J* 2020; 41: 1503–10.
- 60 Henriksson P, Henriksson H, Tynelius P, et al. Fitness and body mass index during adolescence and disability later in life. *Ann Intern Med* 2019; 170: 230–39.
- 61 Ortega FB, Ruiz JR, Castillo MJ, Sjöström M. Physical fitness in childhood and adolescence: a powerful marker of health. *Int J Obes* 2008; 32: 1–11.
- 62 Marques A, Santos R, Ekelund U, Sardinha LB. Association between physical activity, sedentary time, and healthy fitness in youth. *Med Sci Sports Exerc* 2015; 47: 575–80.
- 63 De Baere S, Philippaerts R, De Martelaer K, Lefevre J. Associations between objectively assessed components of physical activity and health-related fitness in 10- to 14-year-old children. *J Phys Act Health* 2016; 13: 993–1001.
- 64 Arain M, Haque M, Johal L, et al. Maturation of the adolescent brain. *Neuropsychiatr Dis Treat* 2013; 9: 449–61.
- 65 Giedd JN, Blumenthal J, Jeffries NO, et al. Brain development during childhood and adolescence: a longitudinal MRI study. *Nat Neurosci* 1999; 2: 861–63.
- 66 Blakemore S-J. Adolescence and mental health. *Lancet* 2019; 393: 2030–31.
- 67 Blakemore S-J, Robbins TW. Decision-making in the adolescent brain. *Nat Neurosci* 2012; 15: 1184–91.
- 68 Scharner S, Stengel A. Alterations of brain structure and functions in anorexia nervosa. *Clin Nutr Exp* 2019; 28: 22–32.
- 69 Seitz J, Bühren K, von Polier GG, Heussen N, Herpertz-Dahlmann B, Konrad K. Morphological changes in the brain of acutely ill and weight-recovered patients with anorexia nervosa: a meta-analysis and qualitative review. *Z Kinder Jugendpsychiatr Psychother* 2014; 42: 7–18.
- 70 Olivo G, Gaudio S, Schiöth HB. Brain and cognitive development in adolescents with anorexia nervosa: a systematic review of fMRI studies. *Nutrients* 2019; 11: 1907.
- 71 Lowe CJ, Morton JB, Reichelt AC. Adolescent obesity and dietary decision making—a brain-health perspective. *Lancet Child Adolesc Health* 2020; 4: 388–96.
- 72 Mohajeri MH, La Fata G, Steinert RE, Weber P. Relationship between the gut microbiome and brain function. *Nutr Rev* 2018; 76: 481–96.
- 73 Obri A, Khirmian L, Karsenty G, Oury F. Osteocalcin in the brain: from embryonic development to age-related decline in cognition. *Nat Rev Endocrinol* 2018; 14: 174–82.
- 74 Khirmian L, Obri A, Karsenty G. Modulation of cognition and anxiety-like behavior by bone remodeling. *Mol Metab* 2017; 6: 1610–15.
- 75 Simon AK, Hollander GA, McMichael A. Evolution of the immune system in humans from infancy to old age. *Proc Biol Sci* 2015; 282: 20143085.
- 76 Klein SL, Flanagan KL. Sex differences in immune responses. *Nat Rev Immunol* 2016; 16: 626–38.
- 77 Sperling M. *Pediatric endocrinology*, 4th edn. Philadelphia, PA: Elsevier, 2014.
- 78 Libert C, Dejager L, Pinheiro I. The X chromosome in immune functions: when a chromosome makes the difference. *Nat Rev Immunol* 2010; 10: 594–604.
- 79 Gubbels Bupp MR, Potluri T, Fink AL, Klein SL. The confluence of sex hormones and aging on immunity. *Front Immunol* 2018; 9: 1269.
- 80 Markle JG, Fish EN. Sex matters in immunity. *Trends Immunol* 2014; 35: 97–104.
- 81 Zimmermann P, Curtis N. Factors that influence the immune response to vaccination. *Clin Microbiol Rev* 2019; 32: e00084-18.
- 82 Rhines AS. The role of sex differences in the prevalence and transmission of tuberculosis. *Tuberculosis (Edinb)* 2013; 93: 104–07.

- 83 Flanagan KL, Fink AL, Plebanski M, Klein SL. Sex and gender differences in the outcomes of vaccination over the life course. *Annu Rev Cell Dev Biol* 2017; **33**: 577–99.
- 84 Ibrahim MK, Zambruni M, Melby CL, Melby PC. Impact of childhood malnutrition on host defense and infection. *Clin Microbiol Rev* 2017; **30**: 919–71.
- 85 Sederquist B, Fernandez-Vojvodich P, Zaman F, Sävdahl L. Recent research on the growth plate: impact of inflammatory cytokines on longitudinal bone growth. *J Mol Endocrinol* 2014; **53**: T35–44.
- 86 Singer K, Lumeng CN. The initiation of metabolic inflammation in childhood obesity. *J Clin Invest* 2017; **127**: 65–73.
- 87 Cox AJ, West NP, Cripps AW. Obesity, inflammation, and the gut microbiota. *Lancet Diabetes Endocrinol* 2015; **3**: 207–15.
- 88 Kelishadi R, Roufsharshaf M, Soheili S, Payghambarzadeh F, Masjedi M. Association of childhood obesity and the immune system: a systematic review of reviews. *Child Obes* 2017; **13**: 332–46.
- 89 WHO. Adolescent pregnancy. Jan 31, 2020 <https://www.who.int/en/news-room/fact-sheets/detail/adolescent-pregnancy> (accessed Sept 11, 2020).
- 90 UN. World fertility patterns 2015. <https://www.un.org/en/development/desa/population/publications/pdf/fertility/world-fertility-patterns-2015.pdf> (accessed Sept 11, 2020).
- 91 Wodon QT, Male C, Nayihouba KA, et al. Economic impacts of child marriage: global synthesis report. Washington, DC: The International Bank for Reconstruction and Development, The World Bank, and The International Center for Research on Women, 2017.
- 92 Scholl TO, Hediger ML. A review of the epidemiology of nutrition and adolescent pregnancy: maternal growth during pregnancy and its effect on the fetus. *J Am Coll Nutr* 1993; **12**: 101–07.
- 93 Hsu JW, Thame MM, Gibson R, et al. Unlike pregnant adult women, pregnant adolescent girls cannot maintain glycine flux during late pregnancy because of decreased synthesis from serine. *Br J Nutr* 2016; **115**: 759–63.
- 94 Fall CHD, Osmond C, Haazen DS, et al. Disadvantages of having an adolescent mother. *Lancet Glob Health* 2016; **4**: e787–88.
- 95 Fall CHD, Sachdev HS, Osmond C, et al. Association between maternal age at childbirth and child and adult outcomes in the offspring: a prospective study in five low-income and middle-income countries (COHORTS collaboration). *Lancet Glob Health* 2015; **3**: e366–77.
- 96 Richter LM, Victora CG, Hallal PC, et al. Cohort profile: the consortium of health-orientated research in transitioning societies. *Int J Epidemiol* 2012; **41**: 621–26.
- 97 Bamishigbin ON Jr, Dunkel Schetter C, Stanton AL. The antecedents and consequences of adolescent fatherhood: a systematic review. *Soc Sci Med* 2019; **232**: 106–19.
- 98 Akseer N, Al-Gashm S, Mehta S, Mokdad A, Bhutta ZA. Global and regional trends in the nutritional status of young people: a critical and neglected age group. *Ann N Y Acad Sci* 2017; **1393**: 3–20.
- 99 Aurino E. Do boys eat better than girls in India? Longitudinal evidence from young lives. Oxford, UK: Young Lives, 2016.
- 100 Victora CG, Adair L, Fall C, et al. Maternal and child undernutrition: consequences for adult health and human capital. *Lancet* 2008; **371**: 340–57.
- 101 Yajnik CS, Deshpande SS, Jackson AA, et al. Vitamin B12 and folate concentrations during pregnancy and insulin resistance in the offspring: the Pune Maternal Nutrition Study. *Diabetologia* 2008; **51**: 29–38.
- 102 Veena SR, Krishnaveni GV, Srinivasan K, et al. Higher maternal plasma folate but not vitamin B-12 concentrations during pregnancy are associated with better cognitive function scores in 9- to 10-year-old children in south India. *J Nutr* 2010; **140**: 1014–22.
- 103 Krishnaveni GV, Veena SR, Winder NR, et al. Maternal vitamin D status during pregnancy and body composition and cardiovascular risk markers in Indian children: the Mysore Parthenon Study. *Am J Clin Nutr* 2011; **93**: 628–35.
- 104 Lagerros YT, Cnattingius S, Granath F, Hanson U, Wikström A-K. From infancy to pregnancy: birth weight, body mass index, and the risk of gestational diabetes. *Eur J Epidemiol* 2012; **27**: 799–805.
- 105 Dabelea D, Mayer-Davis EJ, Lamichhane AP, et al. Association of intrauterine exposure to maternal diabetes and obesity with type 2 diabetes in youth: the SEARCH Case-Control Study. *Diabetes Care* 2008; **31**: 1422–26.
- 106 Krishnaveni GV, Veena SR, Hill JC, Kehoe S, Karat SC, Fall CHD. Intrauterine exposure to maternal diabetes is associated with higher adiposity and insulin resistance and clustering of cardiovascular risk markers in Indian children. *Diabetes Care* 2010; **33**: 402–04.
- 107 Sharma U. Paternal contributions to offspring health: role of sperm small RNAs in intergenerational transmission of epigenetic information. *Front Cell Dev Biol* 2019; **7**: 215.
- 108 Klastrup LK, Bak ST, Nielsen AL. The influence of paternal diet on snRNA-mediated epigenetic inheritance. *Mol Genet Genomics* 2019; **294**: 1–11.

Copyright © 2021 Elsevier Ltd. All rights reserved.



# Major depressive disorder as a neuro-immune disorder: Origin, mechanisms, and therapeutic opportunities

Gislaine Z. Réus<sup>a,\*,1</sup>, Luana.M. Manosso<sup>a,1</sup>, João Quevedo<sup>a,b</sup>, André F. Carvalho<sup>c,d</sup>

<sup>a</sup> Translational Psychiatry Laboratory, Graduate Program in Health Sciences, University of Southern Santa Catarina (UNESC), Criciúma, SC, Brazil

<sup>b</sup> Center of Excellence on Mood Disorders, Fairlance Department of Psychiatry and Behavioral Sciences, McGovern Medical School, The University of Texas Health Science Center at Houston (UTHealth), Houston, TX, USA

<sup>c</sup> Department of Psychiatry, University of Toronto, Toronto, ON, Canada

<sup>d</sup> Centre for Addiction and Mental Health (CAMH), Toronto, ON, Canada

## ARTICLE INFO

### Keywords:

Neuroinflammation  
Gut-brain-axis  
Microglia  
Astrocyte  
Kynurenine pathway  
Major depressive disorder

## ABSTRACT

Notwithstanding advances in understanding the pathophysiology of major depressive disorder (MDD), no single mechanism can explain all facets of this disorder. An expanding body of evidence indicates a putative role for the inflammatory response. Several meta-analyses showed an increase in systemic peripheral inflammatory markers in individuals with MDD. Numerous conditions and circumstances in the modern world may promote chronic systemic inflammation through mechanisms, including alterations in the gut microbiota. Peripheral cytokines may reach the brain and contribute to neuroinflammation through cellular, humoral, and neural pathways. On the other hand, antidepressant drugs may decrease peripheral levels of inflammatory markers. Anti-inflammatory drugs and nutritional strategies that reduce inflammation also could improve depressive symptoms. The present study provides a critical review of recent advances in the role of inflammation in the pathophysiology of MDD. Furthermore, this review discusses the role of glial cells and the main drivers of changes associated with neuroinflammation. Finally, we highlight possible novel neurotherapeutic targets for MDD that could exert antidepressant effects by modulating inflammation.

## 1. Introduction

Major depressive disorder (MDD) is a common mental illness that causes significant disability and burden worldwide. This mood disorder has a global reach, affecting people of all ages, races, sex, and economic strata (Malhi and Mann, 2018; WHO, 2020). Several clinical symptoms may be present, including depressed mood, anhedonia, fatigue or loss of energy nearly every day, feelings of worthlessness, and diminished ability to think or concentrate (American Psychiatric Association, 2013).

Despite meaningful advances in the current understanding of the pathophysiology of MDD, no single mechanism appears to explain all facets of this devastating disorder. The pharmacotherapy of this disorder rests on the well-known monoaminergic theory of depression, which postulates that MDD occurs due to deficiencies of serotonin, dopamine, and noradrenaline in the synaptic cleft (Malhi and Mann, 2018). However, only approximately a third of patients with MDD appears to achieve remission after an adequate trial with a monoaminergic

antidepressant (Rush et al., 2006). Thus indicating that other possible neurobiological mechanisms may underpin this chronic and heterogeneous condition, including but not limited to genetic and epigenetic factors, increased glutamate levels, decreased brain-derived neurotrophic factor (BDNF) and BDNF hypermethylation, deregulation in the hypothalamic–pituitary–adrenal (HPA) axis, oxidative stress, changes in the microbiota-gut-brain axis, as well as neuroinflammation (Foster and McVey Neufeld, 2013; Malhi and Mann, 2018; Miller and Raison, 2016; Zhu et al., 2023). Changes in the gut microbiota could impact the gut barrier and produce higher periphery inflammation, leading to changes in inflammatory pathways that can affect brain communication and to involved in stress and MDD (Carlessi et al., 2021). In treatment-resistant depression, the vagus nervous has been suggested as a primary communication for neuroinflammation mediated by gut microbiota (Hashimoto, 2023).

Although these factors may mechanistically overlap, an ever-increasing body of evidence indicates that the immune system,

\* Correspondence to: Translational Psychiatry Laboratory, University of Southern Santa Catarina, Criciúma, SC 88806-000, Brazil.

E-mail address: [gislainereus@unesc.net](mailto:gislainereus@unesc.net) (G.Z. Réus).

<sup>1</sup> Authors have the same contribution and share first authorship.

particularly the inflammatory response, may be involved in the pathophysiology of MDD. As reviewed by Miller and Raison (2016), considering an evolutionary perspective, the inflammatory response provided an advantage to our ancestors, protecting them from pathogens and predators. However, in our time, the cross-talk between inflammatory pathways and the brain appears to drive the development of several neuropsychiatric disorders, including MDD (Miller and Raison, 2016). In addition,

Leboyer et al. (2016) suggest that in the clinical practice for psychiatric conditions, a reconciled body–brain interface should be integrated, leading to a new “immuno-psychiatry-based nosology.” Indeed, immune and environmental alterations play a vital role in the etiology of MDD, providing a promising opportunity to define hidden subgroups that share genetic, brain, and immune changes (Leboyer et al., 2016).

In this context, we critically review recent efforts investigating the role of inflammation in MDD and the mechanism by which neuroinflammation is integrated. In addition, we will highlight possible neurotherapeutic targets for MDD that could exert antidepressant effects by modulating inflammation.

## 2. Methods

We used the PubMed database to identify relevant articles published until 2023 to construct this narrative review. We focused on: (a) systematic reviews, meta-analyses, and narrative reviews; (b) clinical studies; and (c) preclinical studies. The search strategy comprised the use of the following terms, alone or in combination: “major depressive disorder” OR “depression” OR “inflammation” OR “neuroinflammation” OR “neuroplasticity” OR “microglia” OR “astrocyte” OR “microbiota-gut-brain axis” OR “blood-brain barrier” OR “nutraceuticals” OR “gut-brain axis” OR “kynurenine pathway” OR “probiotics” OR “antidepressant” OR “anti-inflammatory.”

The treatment options addressed in this review were based on the mechanisms of action associated with MDD and anti-inflammatory effects. In addition, clinical studies with therapeutic responses were preferred.

## 3. Depression and inflammation

The field has witnessed an increased research interest in the role of the neuro-immune system in the path-etiology of MDD (Anders et al., 2013; Milaneschi et al., 2020; Raison et al., 2006). It is also worth noting that individuals with MDD have changes in immune cells. A meta-analysis revealed that subjects with MDD had a higher neutrophil-to-lymphocyte ratio when compared to healthy controls (Mazza et al., 2018). Moreover, several meta-analyses showed an increase in systemic inflammatory markers in depressed individuals compared to healthy controls (Dowlati et al., 2010; Howren et al., 2009; Köhler et al., 2017; Smith et al., 2018). Importantly, the evidence was most consistent for interleukin (IL)-6, tumor necrosis factor-alpha (TNF- $\alpha$ ), and C-reactive protein (CRP) (Dowlati et al., 2010; Howren et al., 2009; Köhler et al., 2014; Osimo et al., 2019; Smith et al., 2018). Furthermore, based on the meta-analysis of 30 studies, the prevalence of low-grade inflammation (CRP >3 mg/L) in depression was 27%, and the majority of elevated CRP (>1 mg/L) in depression was 58% (Osimo et al., 2019). Another meta-analysis also revealed that CRP and IL-6 were associated with depression in children and adolescents (Colasanto et al., 2020).

In addition to increased peripheral inflammation, individuals with MDD are at increased risk of neuroinflammation. A systematic review and meta-analysis of studies examining cytokine cerebrospinal fluid (CSF) concentrations revealed an increase in IL-6 and TNF- $\alpha$  in individuals with MDD compared to controls (Enache et al., 2019). Furthermore, CSF levels of IL-6 were increased in suicide attempters regardless of their psychiatric diagnosis. Interestingly, the same meta-analysis also revealed that the translocator protein (TSPO), a

positron emission tomography marker of central inflammation, was elevated in the anterior cingulate cortex and temporal cortex of patients with MDD compared with controls (Enache et al., 2019). Likewise, another meta-analysis demonstrated that IL-1 $\beta$ , IL-6, and TNF- $\alpha$  levels were increased in postmortem brain samples of suicide victims compared to healthy control individuals who did not die from suicide (Black and Miller, 2015).

More recently, a systematic review and meta-analysis, which included 3075 subjects, revealed that neopterin concentrations (produced by monocytes, macrophages, and dendritic cells activated by pro-inflammatory origins) were higher among individuals with MDD compared to healthy controls (Cavaleri et al., 2023). In addition, these results were maintained when it was considered only drug-free subjects (Cavaleri et al., 2023).

Several meta-analyses also showed that antidepressant treatments, especially selective serotonin reuptake inhibitors (SSRIs), decreased peripheral levels of inflammatory markers, including IL-6, IL-1 $\beta$ , TNF- $\alpha$ , and others (Köhler et al., 2018; Wang et al., 2019; Więdołcha et al., 2018). Following this, data from another meta-analysis showed that IL-6 levels decreased with antidepressant treatment regardless of the outcome. In contrast, persistently elevated TNF- $\alpha$  was associated with prospectively determined treatment resistance, thus suggesting peripheral inflammation may at least partly contribute to treatment resistance (Strawbridge et al., 2015). Furthermore, another meta-analysis evidenced that antidepressant treatment significantly decreased levels of TNF- $\alpha$  only in responders, and MDD patients who responded to antidepressant treatment had lower baseline IL-8 levels than the non-responders (Liu et al., 2020).

It is also well established that inflammatory cytokines acting in the brain may promote endocrine, autonomic, and behavioral changes referred to as “sickness behaviors,” which may include manifestations like lethargy, weakness, fatigue, malaise, loss of interest in usual activities, inability to concentrate, anorexia, sleep changes, poor appetite, significant weight loss, and fever (Dantzer et al., 2008; Kelley et al., 2003). It is important to note that several of these symptoms are also commonly observed in individuals with MDD (American Psychiatric Association, 2013). Allied to this context, as reviewed by Anders et al. (2013), most signs and symptoms of depression aid the immune system’s ability to fight infections. Furthermore, Milaneschi et al. (2020) highlight that immuno-metabolic dysregulations can be associated with some atypical depressive symptoms, especially those reflecting altered energy intake/expenditure balance. Likewise, some clinical studies evidenced increased depressive symptoms following acute or chronic infections (Arseniou et al., 2014; Murray et al., 2007). Noteworthy, a recent meta-analysis evidenced that the overall prevalence of depression among COVID-19 patients is 45% (Deng et al., 2021). Recently, we reported that COVID-19 was associated with greater severity of depressive symptoms, stress, and C reactive protein (CRP) levels (de Azevedo Cardoso et al., 2023). Moreover, CRP levels were more significantly increased in individuals with major depressive disorder (MDD) and COVID-19 than in individuals with COVID-19 without MDD (de Azevedo Cardoso et al., 2023).

Another clinical association that demonstrates the relationship between inflammation and depression involves interferon- $\alpha$  (IFN- $\alpha$ ). IFN- $\alpha$  is a pleiotropic cytokine released by the innate immune system, and due to its wide range of action, IFN- $\alpha$  is widely administered as a therapeutic in several medical conditions, including chronic hepatitis C virus infection (Hoofnagle, 2002). Some studies have evidenced that antiviral treatment in chronic hepatitis C patients induces depressive symptoms and MDD (Bonaccorso et al., 2002; Hauser et al., 2002; Horikawa et al., 2003). Notably, a meta-analysis observed that administering SSRIs before starting antiviral treatment reduces the incidence of IFN-induced depression (Udina et al., 2014).

Noteworthy, individuals with illnesses or conditions associated with low-grade inflammation also appear at an increased risk of depressive symptoms and even MDD. Hence, several meta-analyses demonstrate



that obesity is associated with a higher prevalence of depression (ranging from 18% to 57% to more risk, depending on the study) (Abou Abbas et al., 2015; de Wit et al., 2010; Luppino et al., 2010; Mannan et al., 2016; Pereira-Miranda et al., 2017). Some meta-analyses also evidenced that diabetes increases the prevalence of depression (Anderson et al., 2001; Pashaki et al., 2019). However, not all individuals with MDD have an increase in inflammatory cytokines. An interesting clinical study identified two clustered immunophenotypes (using principal components analysis for the 14 absolute cell counts), uninfamed depression ( $n = 125$ ) and inflamed depression ( $n = 81$ ) (Lynall et al., 2020). After investigation, Lynall et al. (2020) showed that the subgroup of cases with increased immune cell counts also had significantly increased inflammatory protein concentrations (CRP and IL-6) and had significantly higher severity of depressive symptoms (analyzing by Hamilton Rating Scale for Depression (HAM-D) and Beck Depression Inventory-II (BDI)) compared with the uninfamed depression cases. Moreover, they also used Gaussian finite multivariate mixture modeling and consensus clustering, without prior constraint on the number of distributions in the mixture, to identify four subgroups of cases, each characterized by a distinct profile of absolute immune cell counts: uninfamed subgroup (S0 - with low counts for all cells and low CRP and IL-6 levels;  $n = 58$ ), subgroup 1 (S1 - with low CRP and IL-6 levels but with higher numbers of B cells, natural killer cells, eosinophils, nonclassical monocytes, and intermediate monocytes;  $n = 10$ ), subgroup 2 (S2 - with significantly increased inflammatory proteins and with significantly higher numbers of classical monocytes, intermediate monocytes, nonclassical monocytes, and neutrophils;  $n = 100$ ), and subgroup 3 (S3 - with increased considerably inflammatory proteins and with a more substantial myeloid bias compared with subgroup S2;  $n = 38$ ) (Lynall et al., 2020). Subgroups 2 and 3 had significantly increased depressive symptom severity scores compared with the uninfamed subgroup (S0). Subgroup 3 (inflamed, myeloid biased) has had significantly increased observer-rated depressive symptoms (HAM-D), self-reported depressive symptoms (BDI), and anhedonia (Snaith-Hamilton Pleasure Scale (SHAPS)) compared with the uninfamed subgroup. Cases in subgroup 2 (inflamed, lymphoid biased) likewise had significantly increased self-reported depressive symptoms (BDI), anhedonia (SHAPS), and fatigue ratings (Chalder Fatigue Scale) compared with the uninfamed subgroup (Lynall et al., 2020).

#### 4. Sources of peripheral inflammation

A normal inflammatory response is essential when a threat is present. The acute inflammatory response is typically initiated during infections via an interaction between pattern recognition receptors expressed on innate immune cells and pathogen-associated molecular patterns (PAMPs) or damage-associated molecular patterns (DAMPs). In contrast, DAMPs without acute infection typically trigger chronic systemic inflammation and are low-grade and persistent (Furman et al., 2019). It is this low-grade and chronic inflammation that is correlated with several diseases, such as cardiovascular disease (Gisterå and Hansson, 2017), type 2 diabetes mellitus (Wellen and Hotamisligil, 2005), cancer (Taniguchi and Karin, 2018), and neuropsychiatric disorders (Miller and Raison, 2016).

Numerous conditions in the modern world can contribute to chronic systemic inflammation and immune activation, such as stress, obesity, smoking, physical inactivity, diet, sleep disorders, decreased vitamin D levels, altered gut permeability, and gut microbiota changes (Berk et al., 2013). Noteworthy, a continually emerging body of evidence supports the role of the gut microbiota in systemic inflammation (Kamada et al., 2013; Maslowski et al., 2009) and psychiatric disorders, including MDD (Foster and McVey Neufeld, 2013; Morais et al., 2020; Sanada et al., 2020).

The gastrointestinal tract is the largest immune organ and the largest area of the human body that interacts with the environment (Powell et al., 2017). The intestinal barrier function is essential against

pathogens and restricts the free exchange of water, ions, and macromolecules between the intestinal lumen and the underlying tissues (Odenwald and Turner, 2017). The intestinal barrier is comprised of physical (epithelial cells, their intercellular junctions, and the mucus layer), chemical (pH, bile salts, enzymes, antibacterial proteins), immunological (immune cells, immunoglobulins, mucosa-associated lymphoid tissue) barriers, and microbial components (the microbiota) (Chassaing et al., 2014; Turner, 2009).

The intercellular junctions (tight junction, adheren junction, and desmosome) comprise the apical junctional complex important to average intestinal permeability (Powell et al., 2017). The increased intestinal permeability, also called leaky gut, allows that small molecule bacterial metabolic components, toxins, and food particles, can translocate or diffuse systemically and leak into the bloodstream (Camilleri, 2019; Obrenovich, 2018). For example, the translocation of bacterial endotoxin lipopolysaccharide (LPS - the main molecular component of the cell wall of Gram-negative bacteria) can activate toll-like receptor 4 (TLR-4), initiating the innate immune response and activating factor nuclear kappa B (NF- $\kappa$ B), which increase inflammatory cytokines (Kawai and Akira, 2010). Thus, a leaky gut can stimulate a systemic inflammatory response.

In line with this, a clinical study showed that biomarkers of the leaky gut were related to inflammation in older adults with cardiometabolic disease (Kavanagh et al., 2019). Moreover, LPS-binding protein and LPS-binding protein/soluble CD14 (endotoxin biomarkers) were associated with greater CRP production in individuals with more hostile marital interactions. Furthermore, combining more hostile marital interactions with a mood disorder history was associated with a higher LPS-binding protein/soluble CD14 ratio (Kiecolt-Glaser et al., 2018).

The gut microbiota also plays a vital role in gut physiology and the modulation of inflammation, having crucial roles in health and disease. As reviewed by Cox, West, and Cripps (2015), the gut microbiota composition can influence intestinal permeability, the number of circulating LPS, and the production of short-chain fatty acids. In addition, these factors can contribute to immune activation and inflammation (Cox et al., 2015). Allied to this context, a preclinical study showed that fecal microbiota transplantation from chronic unpredictable mild, stress mice donors induced depression-like behavior in recipient mice via the gut microbiota-inflammation-brain axis. In addition, mice given the microbiota had significant elevations of IFN- $\gamma$  and TNF- $\alpha$  in the hippocampus and upregulated indoleamine 2,3 dioxygenase (IDO) in the same brain region (N. Li et al., 2019; C. Li et al., 2019).

It is worth mentioning that chronic gut inflammation is linked to a considerable burden of psychological morbidity. Individuals with inflammatory bowel disease have more depressive symptoms than controls (Goodhand et al., 2012; Häuser et al., 2011; Kurina et al., 2001). On the other hand, baseline depression is associated with a higher risk for aggressive inflammatory bowel disease at follow-up (Kochhar et al., 2018), a lower remission rate (Persoons et al., 2005), and an increase in the total number of relapses after 12 and 18 months of follow-up (Mittermaier et al., 2004). Maes, Kubera, and Leunis (2008) revealed that individuals with MDD have more leaky gut and increased LPS translocation. Moreover, clinical studies (Jiang et al., 2015; Kelly et al., 2016) and meta-analyses (Sanada et al., 2020) have also shown that individuals with MDD have altered gut microbiota composition compared to controls. It is worth noting that the microbiota composition is host-specific and can change according to age, sex, medication, diet, exercise, and other factors (Cryan et al., 2019; Stilling et al., 2014).

In a cohort study, Ciocan et al., 2021 demonstrated that unmedicated patients experiencing a major depressive episode had a distinct blood microbiome. In addition, some bacteria (mainly a lower proportion of bacteria from the Saccharibacteria and Fusobacteria phyla) and metabolites were associated with response to antidepressant treatment (Ciocan et al., 2021). However, a lack of information on blood microbiome origin could be a limitation, and future studies in this field may be promising.

## 5. Neuroinflammation: how peripheral inflammation reaches in the brain

Peripheral cytokines can cross the blood-brain barrier (BBB) and reach the central nervous system (CNS), a process dependent on saturable influx transporters or the retrograde axonal transport system. Moreover, peripheral cytokines can increase BBB permeability or be transported through the circumventricular organs (regions where the BBB is less restrictive). PAMPs also can reach the brain at the circumventricular organ level and induce the production and release of the pro-inflammatory cytokine, a process involving TLRs. In addition, peripheral cytokines and PAMPs also can activate vagal afferent nerves, which project to the nucleus tractus solitarius in the CNS (Dantzer et al., 2008; Kaufmann et al., 2017). Additionally, immune cells, such as leukocytes, can cross the BBB and enter the brain (Becher et al., 2017).

It is worth mentioning that the BBB integrity can also be affected by gut microbiota. For example, a preclinical study by Braniste et al. (2014) revealed that germ-free mice, beginning with intrauterine life, displayed increased BBB permeability compared to pathogen-free mice with normal gut microbiota. Moreover, the increased BBB permeability was maintained in germ-free mice after birth and during adulthood and was associated with reduced expression of the tight junction proteins (occludin and claudin-5). Interestingly, the same study also showed that the exposure of germ-free adult mice to a pathogen-free gut microbiota has a decrease in BBB permeability and an up-regulation in the expression of tight junction proteins (Braniste et al., 2014).

Although data are mainly from preclinical studies, evidence in humans shows that peripheral cytokines can reach the CNS. For example, a clinical study revealed that peripherally administered cytokine (IFN- $\alpha$ ) for about 12 weeks can activate a CNS inflammatory response. Besides, IFN- $\alpha$  was detected in the cerebrospinal fluid (CSF) of all IFN- $\alpha$ -treated patients and only in a control subject (Raison et al., 2009). BBB integrity also can be affected in individuals with MDD. A clinical study showed that serum S100B (a marker associated with BBB damage) levels were higher in individuals with MDD than in controls (Arora et al., 2019).

Interestingly, some authors have investigated the occurrence of neuroinflammation from positron emission tomography (PET) images. As recently reviewed by Meyer et al. (2020), analyzing PET imaging of neuroinflammation can be promising for neuropsychiatric disorders, including MDD, despite some limitations. For example, during the MDD episode, it was observed an increase in 18 kDa translocator protein (TSPO), a biomarker of neuroinflammation, binding particularly in unmedicated individuals with a long history of untreated MDD (Meyer et al., 2020).

## 6. Microglia and other cells involved in neuroinflammation

Contrary to what was thought of in the past, today, we know that microglia do not have a neuroectodermal origin but that they belong to the immune system. Microglia are derived from early erythroid myeloid progenitors in the yolk sac, and they populate the brain early during embryogenesis (Ginhoux et al., 2010; Prinz et al., 2019). Although these cells are macrophages, microglial cells found in the adult mouse brain are remnants of these original early erythroid myeloid progenitor-derived cells, which differ from peripheral macrophages (where embryo-derived macrophages are progressively replaced by circulating blood monocytes) (Ginhoux et al., 2010; Prinz et al., 2019). Perivascular and meningeal macrophages also appear to have the exact embryonic origin of CNS microglia [82,83]. On the other hand, the choroid plexus is an exception among the CNS macrophage niches because, in this place, the macrophages originate from monocytes of hematopoietic stem cells and have a shorter life span (Prinz et al., 2019).

Although much remains to be uncovered, microglia have broad functions for the physiology and physiopathology of CNS and have phenotypical heterogeneity (Masuda et al., 2019; Morris et al., 2020;

Tan et al., 2020). Furthermore, several microglial subtypes have already been identified, involving different functions (Masuda et al., 2019; Stratoulis et al., 2019). Notably, it is well-established that microglia have an essential role in neuroinflammation (Leng and Edison, 2021; Miller and Raison, 2016). In the presence of insult (endogenous or exogenous), microglia can activate receptors for DAMPs and PAMPs on the microglial surface. This activation can, among several mechanisms, increase the expression of genes involved with neuroinflammation, including the expression of inflammatory cytokine genes and their receptors (Bajetto et al., 2002; Leng and Edison, 2021). These actions are mediated, at least in part, by activation of pro-inflammatory signaling transducers NOD-like receptor pyrin domain-containing 3 (NLRP3) inflammasome, NF $\kappa$ B, and signaling pathways such as mitogen-activated protein kinases (MAPKs) and phosphatidylinositol 3-kinases (PI3K)-protein kinase B (AKT) (Kaufmann et al., 2017; N. Li et al., 2019; C. Li et al., 2019; Savage et al., 2012; Yu et al., 2019). Interestingly, a recent preclinical study evidenced that activation of microglia in the dorsal striatum induced local cytokine expression (including IL-6 and prostaglandin synthesis) and a negative affective state characterized by anhedonia and aversion, whereas inactivation of microglia blocked aversion caused by systemic inflammation (Klawonn et al., 2021).

In the presence of inflammatory cytokines in the brain and DAMP receptors activation, microglia can also produce reactive oxygen and nitrogen species, generating oxidative stress (Miller et al., 2009; Simpson and Oliver, 2020). In microglia, reactive oxygen species are caused primarily by NADPH oxidase 2 (NOX2) (Choi et al., 2012; Simpson and Oliver, 2020). An interestingly preclinical study showed that stimulation of microglia with LPS also activates caspase-8 and caspase-3/7 in microglia and that this activation regulates microglia activation through a protein kinase C (PKC)- $\delta$ -dependent pathway (Burguillos et al., 2011). Moreover, another study revealed that microglia-driven neuroinflammation involves, at least in part, an upregulation of several micro (mi)RNAs, especially miRNA-342 (Brás et al., 2020). Allied to this context, a systematic review with postmortem studies showed that two studies found increased markers of microglia in MDD brains, while four studies found no MDD-related changes (Enache et al., 2019). Moreover, a clinical study also demonstrated increased microglial densities in the dorsolateral prefrontal cortex, anterior cingulate cortex, and mediadorsal thalamus of suicide patients with MDD or schizophrenia (Steiner et al., 2008).

It is important to note that other cells are also involved in neuroinflammation in addition to microglia. Astrocytes are critical glial cells that regulate the CNS's innate and adaptive immune responses. They can exercise beneficial or detrimental responses for the brain, depending on the timing and context of stimuli/injury (Colombo and Farina, 2016). An exciting study evidenced that reactive microglia can activate reactive astrocytes and that reactive astrocytes lose the ability to promote neuronal survival, synaptogenesis, phagocytosis and induce the death of neurons and oligodendrocytes (Liddel et al., 2017).

Notably, astrocytes expressed several TLRs, especially TLR3. When TLR3 is activated, there is an increase in inflammatory cytokines (Jack et al., 2005). Moreover, this increase in inflammatory cytokine by TLR3 can involve, at least in part, some signaling pathways, including MAPK, glycogen synthase kinase-3beta (GSK-3beta), and PI3K pathways (Park et al., 2006). A preclinical study also demonstrated that in chronic mild stress-induced depressive mice, the level of kynurenine increased (signaling an immune dysfunction - see details in the following topic) in the hippocampus, accompanied by the activation of astrocytic NLRP2 inflammasome (Zhang et al., 2020). Astrocytes also can generate reactive oxygen species, contributing to oxidative stress (Chen et al., 2020). Mitochondria can produce reactive oxygen species generation (ROS) in the astrocytes. A preclinical study revealed that mitochondrial production of reactive oxygen species is mediated by NLRP3 inflammasome activation (Alfonso-Loeches et al., 2014). Like microglia, NOX activation contributes to reactive oxygen species generation in astrocytes (Yang

et al., 2020). It is worth noting that some histological studies of post-mortem brain tissue from individuals with MDD revealed a decrease in the number or packing density of the glial cells (Cotter et al., 2002; Gittins and Harrison, 2011; Rajkowska et al., 1999). Moreover, post-mortem brain tissue from subjects with MDD also reported lower levels of glial fibrillary acidic protein (GFAP), an astrocytic marker, as compared to controls (Gittins and Harrison, 2011; Miguel-Hidalgo et al., 2010; Si et al., 2004), suggesting that astrocytes can be involved in the pathophysiology of MDD.

Finally, the third glial cell that deserves to be highlighted is the oligodendrocytes. Oligodendrocytes are the myelinating cells of the CNS and, thus, are important to rapid impulse propagation and integrity of axons (Moore et al., 2020; Nave and Werner, 2014). Although this glial cell is less studied in terms of neuroinflammation processes, data from the literature demonstrate that oligodendrocytes also contribute to neuroinflammation and that this depends, at least in part, on the peroxisomes (Kassmann et al., 2007). Noteworthy, oligodendroglial abnormalities are also observed in individuals with MDD. Clinical studies showed, for instance, that the expression of several genes related to oligodendrocyte function was significantly decreased in subjects with MDD (Aston et al., 2005). In addition, Uranova et al. (2004) also revealed a lowered density of oligodendroglial cells in subjects with MDD. Moreover, a study using quantitative magnetic resonance imaging demonstrated that MDD participants had lower levels of myelin than the control participants (Sacchet and Gotlib, 2017). Following this, pre-clinical research highlighted that the oligodendrocytes in the hippocampus are involved in the pathogenesis of MDD and that fluoxetine (an SSRI) might prevent the loss of this glial cell (Wang et al., 2020).

## 7. Neuromodulatory mechanisms of cytokines

In CNS, cytokine signals participate in several pathways involved in the development of MDD (Raison et al., 2006). One of the potential mechanisms by which inflammation can induce MDD is associated with the kynurenine pathway (Dantzer et al., 2011). Usually, more than 90% of dietary tryptophan is oxidized through the kynurenine pathway catalyzed by tryptophan dioxygenase (TDO) in the liver or extra-hepatically by IDO (Dantzer et al., 2008). IDO is highly inducible by inflammatory cytokines, especially IFN- $\gamma$  and TNF- $\alpha$  (O'Connor et al., 2009a). Thus, one of the consequences of neuroinflammation is the activation of IDO and a decrease in tryptophan bioavailability to synthesize serotonin (Dantzer et al., 2011).

Moreover, in the brain, kynurenine can be catabolized into kynurenic acid (in astrocytes) or 3-hydroxykynurenine and quinolinic acid (in microglia) (Dantzer et al., 2008; Guillemain et al., 2005). These metabolites involved in the kynurenine pathway are neuroactive and may contribute to behavioral changes. Kynurenic acid antagonizes glutamate release and blocks excitatory neurotransmission. In contrast, quinolinic acid is an excitotoxic neuromodulator that promotes glutamate release by directly stimulating N-methyl-D-aspartate (NMDA) receptors and inducing oxidative stress (Guillemain et al., 2005; Rios and Santamaria, 1991; Schwarcz and Pellicciari, 2002). Allied with this context pre-clinical study showed that peripheral administration of LPS induced a depressive-like behavior and activated IDO (O'Connor et al., 2009b). On the other hand, the same study demonstrated that inhibition of IDO blocked the development of depressive-like behaviors in mice in response to LPS. LPS also increased the plasma and brain kynurenine/tryptophan ratio (indexing IDO activity) (O'Connor et al., 2009b). A clinical study also showed an increased plasma kynurenine/tryptophan ratio in MDD compared to control subjects (Myint et al., 2007). The plasma kynurenine/tryptophan ratio is also higher in suicidal MDD than in non-suicidal MDD subjects and controls (Messaoud et al., 2019). However, there are controversial data. A recent clinical study found a decreased kynurenine/tryptophan ratio in first-episode drug-naïve MDD (Chiu et al., 2021). Other changes in plasma are also observed. For example, a meta-analysis of clinical studies found decreased kynurenine

levels in individuals with MDD (Arnone et al., 2018). Using a meta-analysis of curated metabolic characterization data from a large sample of patients (Pu et al., 2021) identified differential metabolites in the peripheral blood of MDD patients and revealed disturbances of amino acid and lipid metabolism, especially the tryptophan-kynurenine pathway and fatty acid metabolism. Another meta-analysis showed that IFN- $\alpha$  treatment in patients with chronic illnesses was also associated with reduced tryptophan, increased kynurenine, and kynurenine/tryptophan ratio levels, and depression scores (Hunt et al., 2020). In subjects with MDD plasma kynurenine/tryptophan ratio is associated with CSF kynurenine/tryptophan ratio (Haroon et al., 2020). Moreover, inflammatory measures in the plasma, especially TNF- $\alpha$ , were correlated with plasma kynurenine pathway metabolites (Haroon et al., 2020). Another study found that subjects treated with IFN- $\alpha$  exhibited significantly higher concentrations of CSF kynurenine, quinolinic acid, and kynurenic acid. Moreover, CSF kynurenine and quinolinic acid were associated with depressive symptoms (Raison et al., 2010).

Another mechanism that could explain the relationship between inflammation and MDD is related to the serotonin transporter (SERT / 5-HTT). It was reported that IL-1 $\beta$  and TNF- $\alpha$  could up-regulate 5-HTT (Zhu et al., 2006). In humans, a positive correlation was also evidenced between 5-HTT and cytokines mRNA expression of leukocytes (Tsao et al., 2006). Finally, it is worth noting that 5-HTT is located on the presynaptic membrane, is responsible for removing serotonin from the synaptic cleft (Mohammad-Zadeh et al., 2008), and represents an initial action site for SSRI (Stahl, 1998).

The HPA axis is also involved in the effects of inflammatory cytokines. A clinical study assessed the relationship between the HPA axis response to IFN- $\alpha$  and the development of MDD during IFN- $\alpha$  treatment. Interestingly, an increase in adrenocorticotropic hormone (ACTH) and cortisol levels was observed after the initial injection of IFN- $\alpha$ . Nevertheless, patients who subsequently met symptom criteria for MDD exhibited significantly higher ACTH and cortisol responses than patients who did not become depressed, especially 3 h after IFN- $\alpha$  injection (Capuron et al., 2003). The preclinical study also showed that LPS-induced depressive-like behavior was accompanied by an altered HPA axis (Adzic et al., 2015).

Proinflammatory cytokines also influence HPA regulation by disturbing glucocorticoid receptor dysregulation (Kim et al., 2016). For example, in vitro study evidenced that IL-1 $\beta$  inhibits glucocorticoid receptor translocation from the cytoplasm to the nucleus, inducing glucocorticoid resistance (Raddatz et al., 2001). Noteworthy, the dysregulation of the HPA axis in MDD is believed to be related, at least in part, to the failure of cortisol to suppress corticotropin-releasing hormone (CRH) production through negative feedback regulation (Pace et al., 2007; Pace and Miller, 2009; Pariante and Miller, 2001). On the other hand, glucocorticoids also influence inflammatory cytokines. Besides the well-known anti-inflammatory and immunosuppressive effects of glucocorticoids (Becker, 2013), evidence also indicates that glucocorticoids exhibit proinflammatory properties (Cruz-Topete and Cidlowski, 2014). Data from literature demonstrated, for example, that glucocorticoids can activate TLR2 signaling, facilitate NLRP3 induction and inflammasome formation, and active P2Y2 receptor (ATP receptor which stimulates p38 activation, which leads to the expression of IL-6) (Busillo et al., 2011; Cruz-Topete and Cidlowski, 2014; Ding et al., 2010; Hermoso et al., 2004).

A good deal of data also has established that inflammation is detrimental to neurogenesis and neuroplasticity, thus influencing several situations in the brain, including behavior (Ekdahl et al., 2009, 2003; Kohman and Rhodes, 2013). Indeed, proinflammatory cytokine receptors are highly aggregated in regions involved in neurogenesis and MDD, such as the hippocampus (Arisi, 2014; Rehman et al., 2019). Moreover, several preclinical data highlighted that LPS attenuates adult hippocampal neurogenesis (Fujioka and Akema, 2010; Monje et al., 2003; Valero et al., 2014) and decreases protein expression (Fang et al., 2020; Zhang et al., 2014) or levels (Guan and Fang, 2006) of BDNF (an



essential neurotrophic factor for neurogenesis) in several regions of the brain, including the hippocampus and prefrontal cortex. Similarly, other studies also showed that IFN- $\alpha$  decreases neurogenesis in the hippocampus (Kaneko et al., 2006; Zheng et al., 2015, 2014) and BDNF levels in the hippocampus and prefrontal cortex (Callaghan et al., 2017). Moreover, TNF- $\alpha$  (Keohane et al., 2010), IL-6 (Pollak, 2014), and IL-1 $\beta$  (Ja and Duman, 2008; Wu et al., 2012) are also capable of affecting neurogenesis, decreasing the proliferation and differentiation of neurons. Interestingly, BDNF can be responsible for the detrimental effect of proinflammatory cytokines on neuroplasticity (Calabrese et al., 2014). Indeed, an acute and higher inflammatory insult induced by LPS or ibuprofen impairs the neuronal plasticity of adult mice (Golia et al., 2019). Finally, it is worth noting that MDD can be associated with disrupting mechanisms involved in cell survival and neuroplasticity. On the other hand, antidepressants can modulate the signaling pathways involved in cell survival and neuroplasticity (D'Sa and Duman, 2002; Pittenger and Duman, 2008).

Another related pathway to the effects of inflammation on behavior includes the capacity of cytokines and inflammatory mediators to increase glutamate release. It decreased glutamate reuptake, contributing to glutamatergic excitotoxicity and decreased BDNF (Miller et al., 2009). Indeed, a preclinical study highlighted that LPS increases glutamate in the hypothalamus (Huang et al., 2008). Moreover, in vitro, studies also showed that LPS reduced glutamate uptake in astrocytes and induced astroglial neurotoxicity (Zhou et al., 2006).

## 8. Therapeutic opportunities for MDD in the neuroinflammatory context

Based on all of the above, decreasing inflammation could be a way to assist in the treatment of MDD. In addition, targeting inflammation can be a practical approach to preventing MDD (Chu et al., 2021).

In this context, several strategies can be interesting. First, to bring something with clinical application, we will mention therapeutic opportunities that fall under two factors: have a meta-analysis of clinical studies showing that it improves depressive symptoms and have clinical and preclinical studies that show its anti-inflammatory effects.

### 8.1. Pharmacological therapeutic opportunities

As mentioned, several meta-analyses highlighted that antidepressant treatment, especially SSRIs, decreases peripheral levels of inflammatory markers in MDD subjects (Köhler et al., 2017; Wang et al., 2019; Więdocha et al., 2018). In addition, a preclinical study also demonstrated that SSRI attenuated the LPS-induced increase of TNF- $\alpha$  (Ohgi et al., 2013). In line with this, in vitro studies revealed that SSRI suppressed microglial responses to an inflammatory stimulus induced by LPS (Liu et al., 2011; Tynan et al., 2012).

Ketamine's efficacy in depressive disorders has been established in several controlled trials. Indeed, several meta-analyses showed that ketamine decreased depressive symptoms in MDD subjects (Fond et al., 2014; Xu et al., 2016). Although several mechanisms of action are proposed (Amidfar et al., 2019; Zanos and Gould, 2018), ketamine has also been shown to have an anti-inflammatory effect (Kopra et al., 2021). In individuals with treatment-resistant depression, low-dose ketamine infusion modulated peripheral proinflammatory cytokine (Chen et al., 2018). A preclinical study also revealed that ketamine could alleviate LPS-induced depressive-like behavior and that this effect was associated with decreased inflammatory cytokine expression in the prefrontal cortex (Yang et al., 2013). Consistent with this, it was also demonstrated that ketamine inhibited LPS-mediated microglial inflammation (Lu et al., 2020).

Throughout this conceptual framework, emphasis can also be made on anti-inflammatory drugs. It is worth noting that some drugs are used to treat a broad range of immune-mediated inflammatory diseases, including medications with key extracellular, membrane, or

intracellular targets (McInnes and Gravalles, 2021). Interestingly, a meta-analysis that investigated the efficacy of anti-inflammatory drugs (including nonsteroidal anti-inflammatory, cytokine inhibitors, minocycline, pioglitazone, and glucocorticoids) on MDD concluded that anti-inflammatory agents improved depressive symptoms compared to placebo as an add-on in patients with MDD and as monotherapy (Köhler-Forsberg et al., 2019). Moreover, anti-inflammatory add-on improved response and remission (Köhler-Forsberg et al., 2019). In line with this, another meta-analysis revealed that anti-inflammatory treatment had a beneficial effect on depressive symptoms (Köhler et al., 2014). In addition, sub-analyses of celecoxib (a selective cyclooxygenase 2 inhibitor) showed improved antidepressant responses with little heterogeneity, particularly with add-on treatment (Köhler et al., 2014). Two other meta-analyses evaluating only adjunctive celecoxib treatment showed an effective add-on treatment for MDD individuals (Faridhosseini et al., 2014; Na et al., 2014). Wittenberg et al. (2020), in a mega-analysis of randomized, placebo-controlled clinical trials in inflammatory disorders, demonstrated that immunomodulatory drug treatment (mainly IL-6 antibodies and an anti-IL-12/23 antibody) improved depressive symptoms, specifically in the subgroups of patients with high depressive symptoms at baseline.

The minocycline (a drug with anti-infectious activity) has shown promise in neurology (Yong et al., 2004) and psychiatry (Dean et al., 2012). The same meta-analyses demonstrated that minocycline decreased depressive symptoms in MDD individuals (Cai et al., 2020; Rosenblat and McIntyre, 2018). A meta-analysis of preclinical studies also showed that minocycline reduced depressive-like behavior in disease, but not healthy, animal models (Reis et al., 2019). Noteworthy, a recent clinical trial revealed that augmentation therapy with minocycline in treatment-resistant MDD patients is effective only in individuals with low-grade inflammation (defined as CRP  $\geq 3$  mg/L) (Nettis et al., 2021). Moreover, the same study showed that responders to minocycline had higher baseline IL-6 concentrations than non-responders and that IFN- $\gamma$  was significantly reduced after treatment with minocycline compared with placebo (Nettis et al., 2021). In line with this, minocycline treatment ameliorates IFN- $\alpha$  induced neurogenic defects and depression-like behaviors in mice (Zheng et al., 2015).

Statins, a drug traditionally used to control cholesterol, also seem helpful for psychiatric disorders (Kim et al., 2019). Two meta-analyses of clinical trials showed that statins, as an add-on to treatment as usual, improved depressive symptoms (Giorgi et al., 2021; Salagre et al., 2016). In addition, a clinical study revealed that statin treatment decreases chronic inflammation (Sørensen et al., 2019). Two meta-analyses also showed that statin decreases CRP in individuals with heart failure (Bonsu et al., 2015) and systemic lupus erythematosus (Sahebkar et al., 2016). A preclinical study also demonstrated that atorvastatin prevented LPS-induced depressive-like behavior and increased TNF- $\alpha$  (Taniguti et al., 2019).

It is important to note that no clear recommendations can result in anti-inflammatory treatment for MDD due to several critical aspects such as heterogeneity, diversity of patient populations, treatment regimen, outcomes, and limited scientific quality, as recently reported by Simon et al. (2023). Furthermore, a prediction of treatment outcome according to immune considerations is a potential prosperous of clinical research, and the use of immunomodulation clasps the promise to improve the efficacy of the next generation of antidepressant drugs. However, considering the patient's baseline immune condition, a personalized medicine approach appears reasonable for designing successful interventions (Branchi et al., 2021).

### 8.2. Nutritional therapeutic opportunities

The first nutritional therapeutic opportunity is omega-3 polyunsaturated fatty acids. Several meta-analyses show that omega-3 supplementation improves depressive symptoms in subjects with MDD (Grosso et al., 2014; Liao et al., 2019; Mocking et al., 2016).



Noteworthy, the type of omega-3 used can influence this clinical effect. The use of mainly eicosapentaenoic acid (EPA) within the preparation (EPA  $\geq$ 60%, in a dose range of 200–2200 mg EPA) rather than docosahexaenoic acid (DHA) is more effective (Grosso et al., 2014; Liao et al., 2019; Mocking et al., 2016). EPA and DHA are capable of partly inhibiting many aspects of inflammation, including modulation of cells involved in the immune system, decreasing eicosanoids like prostaglandins and leukotrienes from the arachidonic acid (omega 6 fatty acid), nuclear factor-kappa B (NF $\kappa$ B), and inflammatory cytokines (Calder, 2017; Simopoulos, 2002). In the CNS, omega-3 also regulates the molecular signaling of microglia, especially in the context of neuroinflammation and behavior (Layé et al., 2018). Interestingly, a clinical study showed that inflammation (high levels of CRP and IL-1) was a predictive biomarker for response to omega-3 fatty acids in MDD subjects (Rapaport et al., 2016). The preclinical studies also demonstrated that omega-3 prevented LPS-induced depressive-like behavior by inhibiting neuroinflammation (Shi et al., 2017). Although several other mechanisms of action have been proposed for omega-3, it is worth noting that this type of fatty acid can also modulate the gut microbiota and reduce gut inflammation, which may influence the microbiota-gut-brain axis (Barbalho et al., 2016; Costantini et al., 2017).

A continually emerging body of evidence supports the immunomodulating effects of vitamin D, which can be beneficial in autoimmune and inflammatory diseases (Aranow, 2011; Guillot et al., 2010; Sassi et al., 2018). A meta-analysis of clinical trials, for example, showed that vitamin D decreased CRP and TNF- $\alpha$  in individuals with type 2 diabetes, suggesting an improvement in chronic low-grade inflammation (Mousa et al., 2018). In line with this, a preclinical study demonstrated that vitamin D decreased NF $\kappa$ B in the hippocampus and reversed BBB permeability changes (Hajiluan et al., 2017). Another study revealed that vitamin D reversed unpredictable chronic mild stress-induced depressive-like behavior via reducing oxidative stress and inhibiting

neuroinflammation (Bakhtiari-Dovvombaygi et al., 2021). In humans, several meta-analyses showed that vitamin D reduced depressive symptoms and CRP in MDD subjects (Jamilian et al., 2019; Vellekkatt and Menon, 2019). Moreover, a meta-analysis of observational studies and randomized controlled trials revealed that low vitamin D levels were associated with an increased risk ratio of MDD (Anglin et al., 2013).

Another nutrient essential to the immune system is zinc. This mineral has a recognized role in decreasing inflammation, mainly by reducing the expression of NF $\kappa$ B and, consequently, decreasing the production of inflammatory cytokines (Prasad, 2014, 2008). A preclinical study also showed that zinc prevented depressive-like behavior induced by LPS and inhibited IFN- $\gamma$  (in the plasma) and astrogliosis in rats (Kirsten et al., 2020). *In vitro*, study also revealed that zinc inhibited LPS-induced inflammation in BV2 microglial cells (Hongxia et al., 2019). In humans, a recent meta-analysis showed that zinc supplementation (usually 25 mg/day) reduced depressive symptoms in individuals diagnosed with MDD who were treated with antidepressant therapy (da Silva et al., 2021). Other meta-analyses revealed that MDD is associated with a lower zinc concentration in peripheral blood (Swardfager et al., 2013) and lower dietary zinc intake (Li et al., 2017).

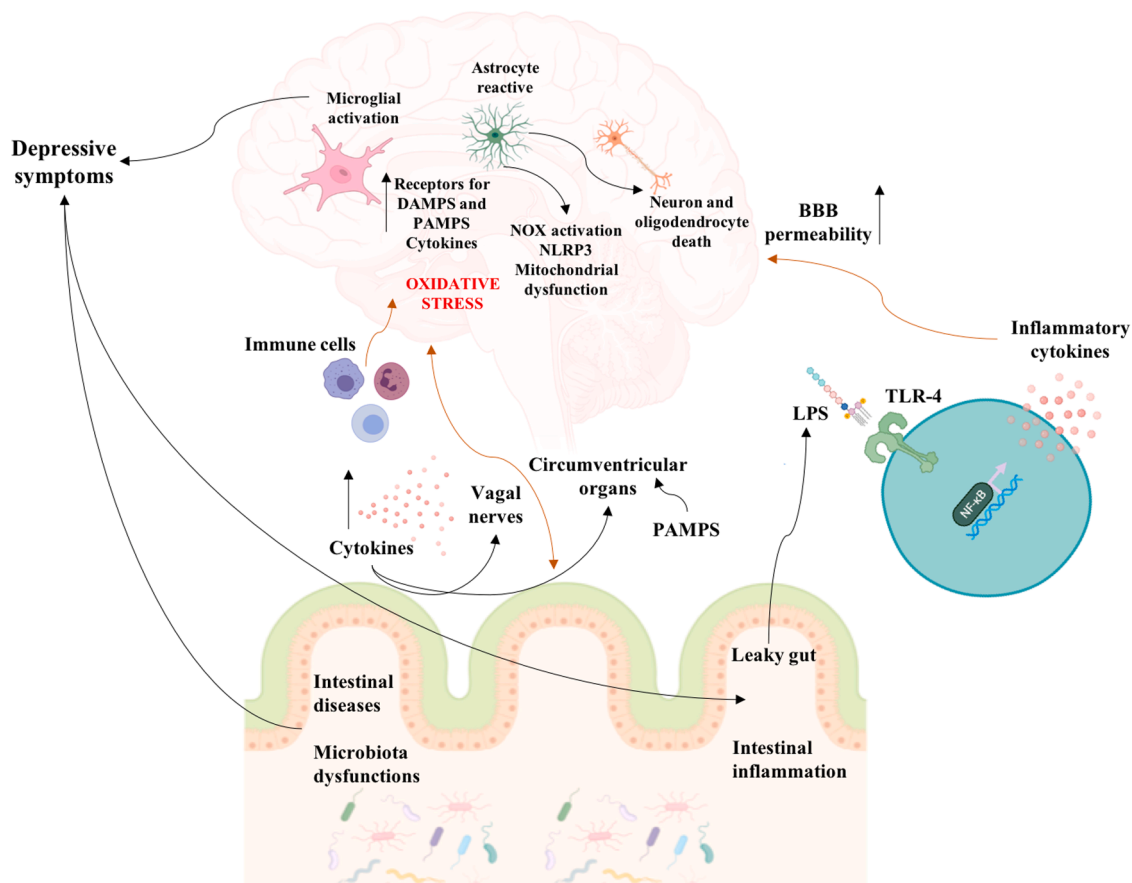
Acetyl-L-carnitine is another supplementation that can improve depressive symptoms. Indeed, a meta-analysis revealed that acetyl-L-carnitine reduced depressive symptoms compared with the control group, and it has similar effectiveness compared with antidepressants in reducing depressive symptoms (Veronese et al., 2018). This effect can be, at least in part, involved with inflammation. Preclinical studies also showed that acetyl-L-carnitine conferred neuroprotection against LPS-induced neuroinflammation (Jamali-Raeufy et al., 2021; Kazak and Yarim, 2017) by suppressing the NF $\kappa$ B pathway (Jamali-Raeufy et al., 2021). Interestingly, a meta-analysis also demonstrated that L-carnitine supplementation was associated with lower CRP, IL-6, and TNF- $\alpha$  levels

**Table 1**

Meta-analyses (of clinical trials) that investigated the antidepressant effect of anti-inflammatory drugs or nutritional strategies.

Drug / Nutritional strategy	Trials included in meta-analysis	Number of individuals included in meta-analysis	Main outcomes	Reference
Anti-inflammatory drugs	N = 36 - NSAIDs (N = 13) - Cytokine inhibitors (N = 9) - Statins (N = 7) - Minocycline (N = 3) - Pioglitazone (N = 2) - Glucocorticoids (N = 2)	N = 9422 - NSAIDs (N = 4214) - Cytokine inhibitors (N = 3345) - Statins (n = 1576) - Minocycline (N = 151) - Pioglitazone (N = 77) - Glucocorticoids (N = 59)	Improve depressive symptoms compared to placebo as an add-on and monotherapy.	(Köhler-Forsberg et al., 2019)
Anti-inflammatory drugs	N = 14 - NSAIDs (N = 10) - Cytokine inhibitors (N = 4)	N = 6262 - NSAIDs (N = 4258) - Cytokine inhibitors (N = 2004)	- Anti-inflammatory and celecoxib as an add-on to antidepressants improve depressive symptoms.	(Köhler et al., 2014)
Minocycline	N = 3	N = 158	- Minocycline had antidepressant effect.	(Rosenblat and McIntyre, 2018)
Minocycline	N = 3	N = 151	- Minocycline had antidepressant effect superior to placebo.	(Cai et al., 2020)
Statin	N = 7	N = 403	- Statins, as add-on to treatment, improved depressive symptoms.	(Salagre et al., 2016; Giorgi et al., 2021)
Omega-3 polyunsaturated fatty acids	N = 50	N = 3868	- Omega-3 improved depressive symptoms.	(Liao et al., 2019; Mocking et al., 2016; Grosso et al., 2014)
Vitamin D	N = 13	N = 2296	- Vitamin D decreased depressive symptoms.	(Jamilian et al., 2019; Vellekkatt and Menon, 2019)
Zinc	N = 5	N = 192	- Zinc supplementation reduced depressive symptoms.	(da Silva et al., 2021)
Acetyl-L-carnitine	N = 12	N = 791	- Acetyl-L-carnitine reduced depressive and had similar effectiveness compared with antidepressants.	(Veronese et al., 2018)
Curcumin	N = 10	N = 532	- Curcumin improves depressive symptoms.	(Fusar-Poli et al., 2019)
<i>Crocus sativus</i>	N = 49	N = 1459	- <i>Crocus sativus</i> improves depressive symptoms.	(Dai et al., 2020; Tóth et al., 2019; Yang et al., 2018)
Probiotics	N = 30	N = 2720	- Probiotics improve depressive symptoms.	(Sanada et al., 2020; Liu et al., 2019)

Abbreviation: MDD – major depressive disorders; NSAIDs - nonsteroidal anti-inflammatory drugs.



**Fig. 1.** Sources of inflammation associated with major depressive disorder. Changes in the intestine or gut microbiota could translocate lipopolysaccharide (LPS) to the periphery, activating the toll-like receptor 4 (TLR-4) and, by nuclear factor-kappa B (NF-κB), increasing inflammatory cytokines expression. Other periphery inflammatory changes also can induce blood-brain barrier (BBB) change or are mediated by vagal nerves or circumventricular organs that access the brain. In the brain, microglia and astrocytes exacerbate immune response by damage-associated molecular patterns (DAMP), pathogen-associated molecular patterns (PAMP), NOD-like receptor pyrin domain-containing 3 (NLRP3), and cytokines receptors, leading to oxidative stress, mitochondrial impairment, and neuron and oligodendrocyte death. These changes are associated with depressive symptoms.

than controls (Haghighatdoost et al., 2019).

*Curcuma longa* (turmeric) is traditionally used in Ayurvedic medicine to treat inflammatory diseases. The main constituent of turmeric is polyphenol curcumin. Reviews of clinical and preclinical studies highlight the anti-inflammatory role of curcumin, primarily through its ability to inhibit cyclooxygenase-2 (COX-2), lipoxygenase (LOX), and inducible nitric oxide synthase (iNOS), which are essential enzymes that mediate inflammatory processes (Jurenka, 2009; Menon and Sudheer, 2007). Curcumin also suppresses NF-κB and inhibits the production of inflammatory cytokines, such as TNF-α, IFN-α, IL-1, and IL-6 (Jurenka, 2009; Menon and Sudheer, 2007). Moreover, preclinical studies showed curcumin prevented BBB damage (Jiang et al., 2007; Wu et al., 2021). In cultured astrocytes, curcumin inhibited iNOS expression induced by LPS/TNF-α (Jiang et al., 2007). Interestingly, a preclinical study reported that curcumin attenuated collagen-induced inflammatory response through the gut-brain axis (Dou et al., 2018). Another study demonstrated that curcumin relieved depressive-like behavior via inhibition of the NLRP3 inflammasome and kynurenine pathway in rats with chronic unpredictable mild stress (Zhang et al., 2019). Allied with this context, a meta-analysis of clinical trials also demonstrated that curcumin (significantly 1000 mg/day for four-six weeks) improved depressive symptoms in MDD individuals (Fusar-Poli et al., 2019).

Another therapeutic opportunity is the *Crocus sativus* (Saffron). Some meta-analyses showed that *Crocus sativus* improves depressive symptoms in individuals with mild to moderate MDD (Dai et al., 2020; Tóth et al., 2019; Yang et al., 2018). Interestingly, the same meta-analyses

also revealed that *Crocus sativus* had comparable efficacy to synthetic antidepressants (Dai et al., 2020; Tóth et al., 2019; Yang et al., 2018). Although several mechanisms of action have been proposed, it is worth noting that *Crocus sativus* has been well-recognized for its immunomodulatory and anti-inflammatory effects (Shafiee et al., 2018; Zeinali et al., 2019). In addition, it is worth noting that crocin (an active constituent of *Crocus sativus*) attenuated LPS-induced depressive-like behavior by suppressing NF-κB and NLRP3 signaling pathways (Zhang et al., 2018).

There is also a growing interest in using probiotics to improve depressive symptoms. Some meta-analyses showed that probiotics significantly decreased the depression scale score of both the healthy population and MDD subjects (Liu et al., 2019; Sanada et al., 2020). Interestingly, a meta-analysis with 42 randomized clinical trials found a significant reduction in serum CRP, TNF-α, IL-6, IL-12, and IL-4 concentrations after probiotic supplementation (Milajerdi et al., 2020). In line with this, a preclinical study revealed that *Lactobacillus plantarum* ameliorated depression-like behaviors, attenuated hippocampal neuronal damage, and decreased inflammation in a corticosterone-induced rat model of chronic stress (Zhao et al., 2020). Another study showed that *Lactobacillus farciminis* prevented stress-induced gut hyperpermeability, LPS upload in the portal blood, and mRNA expression of pro-inflammatory cytokines in the hypothalamus (Ait-Belgnaoui et al., 2012).

Finally, it is worth mentioning that, considering MDD prevention, a meta-analysis highlighted that an anti-inflammatory diet might be an

effective intervention for reducing depression risk and symptoms (Tolkien et al., 2019).

It is important to note that clinician guidelines for the treatment of psychiatric disorders with nutraceuticals and phytochemicals published by Sarris et al. (2022) demonstrated that for adjunctive omega-3 fatty acids were recommended, vitamin D, methyl folate, and adjunctive s-adenosyl methionine (SAME) were weakly recommended; and adjunctive probiotics and zinc were provisionally recommended. On the other hand, monotherapy omega-3, folic acid, vitamin C, tryptophan, creatine, inositol, magnesium, n-acetyl cysteine, and SAME were not supported for unipolar depression use (Sarris et al., 2022), suggesting concerned advised caution when prescribing or recommending nutraceuticals and/or phytochemicals as monotherapies in cases of severe psychiatric disorders.

Table 1 shows the number of trials and the number of subjects evaluated in meta-analyses (of clinical trials) that investigated the antidepressant effect of anti-inflammatory drugs or nutritional strategies.

## 9. Conclusions

MDD is a complex disorder, and several factors are involved in the pathophysiology of this psychiatric condition. The CNS itself can be disturbed by stress situations and trigger a source of inflammation, affecting glial cells and implicating neural circuits involved in mood regulation. In addition, peripheral immune disturbances may be a source of central inflammation, leading to the activation of signaling pathways involved with increased oxidative stress and gene expression of inflammatory cytokines. Microglia, astrocytes, and oligodendrocytes are affected by inflammation, change their physiological responses, and may impair neurotransmission. The gut can also be an essential source of inflammation that, in different ways, including vagal nerves and damage to the BBB, leads to brain changes (Fig. 1). These changes could be associated to depressive symptoms. Although it is impossible to say that all individuals with MDD have inflammation, being aware of this may help in treatment. Evaluating possibilities for reducing inflammation through healthy eating, physical activity practice, or drugs or supplements can be an adjunct strategy assisting in response to treatment. Future studies are encouraged to investigate the biological effects of neuroinflammation and MDD, but many gaps remain to be investigated. Not all inflamed individuals have MDD, just as not all individuals with MDD have inflammation. Individual differences, as well as gender differences, need to be taken into account. Furthermore, further studies are needed for treatment options based on decreasing inflammation.

## Data Availability

No data was used for the research described in the article.

## Acknowledgements

Its research is supported by grants from CNPq (GZR), FAPESC (GZR), and UNESC (GZR). GZR is a 2 CNPq Research Fellow.

## References

- Abou Abbas, L., Salameh, P., Nasser, W., Nasser, Z., Godin, I., 2015. Obesity and symptoms of depression among adults in selected countries of the Middle East: a systematic review and meta-analysis. *Clin. Obes.* 5, 2–11. <https://doi.org/10.1111/cob.12082>.
- Adzic, M., Djordjevic, J., Mitic, M., Brkic, Z., Lukic, I., Radojic, M., 2015. The contribution of hypothalamic neuroendocrine, neuroplastic and neuroinflammatory processes to lipopolysaccharide-induced depressive-like behaviour in female and male rats: Involvement of glucocorticoid receptor and C/EBP- $\beta$ . *Behav. Brain Res.* 291, 130–139. <https://doi.org/10.1016/j.bbr.2015.05.029>.
- Ait-Belgnaoui, A., Durand, H., Cartier, C., Chaumaz, G., Eutamene, H., Ferrier, L., Houdeau, E., Fioramonti, J., Bueno, L., Theodorou, V., 2012. Prevention of gut leakiness by a probiotic treatment leads to attenuated HPA response to an acute psychological stress in rats. *Psychoneuroendocrinology* 37, 1885–1895. <https://doi.org/10.1016/j.psyneuen.2012.03.024>.
- Alfonso-Loeches, S., Ureña-Peralta, J.R., Morillo-Bargues, M.J., La Cruz, J.O., De, Guerri, C., 2014. Role of mitochondria ROS generation in ethanol-induced NLRP3 inflammasome activation and cell death in astroglial cells. *Front Cell Neurosci.* 8, 216. <https://doi.org/10.3389/fncel.2014.00216>.
- American Psychiatric Association, 2013. *Diagnostic and Statistical Manual of Mental Disorders: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5)*. Arlington.
- Amidfar, M., Woelfer, M., Réus, G.Z., Quevedo, J., Walter, M., Kim, Y.K., 2019. The role of NMDA receptor in neurobiology and treatment of major depressive disorder: evidence from translational research. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 94, 109668. <https://doi.org/10.1016/j.pnpbp.2019.109668>.
- Anders, S., Tanaka, M., Kinney, D.K., 2013. Depression as an evolutionary strategy for defense against infection. *Brain Behav. Immun.* 31, 9–22. <https://doi.org/10.1016/j.bbi.2012.12.002>.
- Anderson, R.J., Freedland, K.E., Clouse, R.E., Lustman, P.J., 2001. The prevalence of comorbid depression in adults with diabetes: a meta-analysis. *Diabetes Care* 24, 1069–1078. <https://doi.org/10.2337/diacare.24.6.1069>.
- Anglin, R.E.S., Samaan, Z., Walter, S.D., Sarah, D.M., 2013. Vitamin D deficiency and depression in adults: systematic review and meta-analysis. *Br. J. Psychiatry* 202, 100–107. <https://doi.org/10.1192/bjp.bp.111.106666>.
- Aranow, C., 2011. Vitamin D and the immune system. *J. Invest. Med* 59, 881–886. <https://doi.org/10.2310/JIM.0b013e31821b8755>.
- Arisi, G.M., 2014. Nervous and immune systems signals and connections: cytokines in hippocampus physiology and pathology. *Epilepsy Behav.* 38, 43–47. <https://doi.org/10.1016/j.yebeh.2014.01.017>.
- Arnone, D., Saraykar, S., Salem, H., Teixeira, A.L., Dantzer, R., Selvaraj, S., 2018. Role of Kynurenine pathway and its metabolites in mood disorders: a systematic review and meta-analysis of clinical studies. *Neurosci. Biobehav. Rev.* 92, 477–485. <https://doi.org/10.1016/j.neubiorev.2018.05.031>.
- Arora, P., Sagar, R., Mehta, M., Pallavi, P., Sharma, S., Mukhopadhyay, A.K., 2019. Serum S100B levels in patients with depression. *Indian J. Psychiatry* 61, 70–76. [https://doi.org/10.4103/psychiatry.IndianJPsychiatry\\_391\\_16](https://doi.org/10.4103/psychiatry.IndianJPsychiatry_391_16).
- Arseniou, S., Arvaniti, A., Samakouri, M., 2014. HIV infection and depression. *Psychiatry Clin. Neurosci.* 68, 96–109. <https://doi.org/10.1111/pcn.12097>.
- Aston, C., Jiang, L., Sokolov, B.P., 2005. Transcriptional profiling reveals evidence for signaling and oligodendroglial abnormalities in the temporal cortex from patients with major depressive disorder. *Mol. Psychiatry* 10, 309–322. <https://doi.org/10.1038/sj.mp.4001565>.
- de Azevedo Cardoso, T., Silva, R.H., Fernandes, J.L., Arent, C.O., Amboni, G., Borba, L. A., Padilha, A.P.Z., Botelho, M.E.M., Maciel, A.L., Barichello, F., Morales, R., Soares, S.J.B., Bagatini, M.D., Dallagnol, C., Brighenti, M.E., Ignácio, Z.M., Quevedo, J., Ceretta, L.B., Réus, G.Z., 2023. Stress levels, psychological symptoms, and C-reactive protein levels in COVID-19: a cross-sectional study. *J. Affect Disord.* 330, 216–226. <https://doi.org/10.1016/j.jad.2023.03.019>.
- Bajetto, A., Bonavia, R., Barbero, S., Schettini, G., 2002. Characterization of chemokines and their receptors in the central nervous system: physiopathological implications. *J. Neurochem* 82, 1311–1329. <https://doi.org/10.1046/j.1471-4159.2002.01091.x>.
- Bakhtiar-Dovvombaygi, H., Izadi, S., Zare Moghaddam, M., Hashemzahi, M., Hosseini, M., Azhdari-Zarmehri, H., Dinpanah, H., Beheshti, B., 2021. Beneficial effects of vitamin D on anxiety and depression-like behaviors induced by unpredictable chronic mild stress by suppression of brain oxidative stress and neuroinflammation in rats. *Naunyn Schmiede Arch. Pharm.* 394, 655–667. <https://doi.org/10.1007/s00210-020-02002-0>.
- Barbalho, S.M., Goulart, R.D.A., Quesada, K., Bechara, M.D., De Carvalho, A.D.C.A., 2016. Inflammatory bowel disease: can omega-3 fatty acids really help? *Ann. Gastroenterol.* 29, 37–43.
- Becher, B., Spath, S., Goverman, J., 2017. Cytokine networks in neuroinflammation. *Nat. Rev. Immunol.* 17, 49–59. <https://doi.org/10.1038/nri.2016.123>.
- Becker, D.E., 2013. Basic and clinical pharmacology of glucocorticosteroids. *Anesth. Prog.* 60, 25–32. <https://doi.org/10.2344/0003-3006-60.1.25>.
- Berk, M., Williams, L.J., Jacka, F.N., O'Neil, A., Pasco, J.A., Moylan, S., Allen, N.B., Stuart, A.L., Hayley, A.C., Byrne, M.L., Maes, M., 2013. So depression is an inflammatory disease, but where does the inflammation come from. *BMC Med* 11, 200. <https://doi.org/10.1186/1741-7015-11-200>.
- Black, C., Miller, B.J., 2015. Meta-analysis of cytokines and chemokines in suicidality: distinguishing suicidal versus nonsuicidal patients. *Biol. Psychiatry* 78, 28–37. <https://doi.org/10.1016/j.biopsych.2014.10.014>.
- Bonaccorso, S., Marino, V., Biondi, M., Grimaldi, F., Ippoliti, F., Maes, M., 2002. Depression induced by treatment with interferon-alpha in patients affected by hepatitis C virus. *J. Affect Disord.* 72, 237–241. [https://doi.org/10.1016/S0165-0327\(02\)00264-1](https://doi.org/10.1016/S0165-0327(02)00264-1).
- Bonsu, K.O., Reidpath, D.D., Kadirvelu, A., 2015. Effects of statin treatment on inflammation and cardiac function in heart failure: an adjusted indirect comparison meta-analysis of randomized trials. *Cardiovasc Ther.* 33, 338–346. <https://doi.org/10.1111/1755-5922.12150>.
- Branchi, I., Poggini, S., Capuron, L., Benedetti, F., Poletti, S., Tamouza, R., Drexhage, H. A., Penninx, B.W.J.H., Pariante, C.M., 2021. European college of neuropsychopharmacology (ECNP) immunoneuropsychiatry thematic working group and marion leboyer. Brain-immune crosstalk in the treatment of major depressive disorder. *Eur. Neuropsychopharmacol.* 45, 89–107. <https://doi.org/10.1016/j.euroneuro.2020.11.016>.
- Braniste, V., Al-Asmakh, M., Kowal, C., Anuar, F., Abbaspour, A., Tóth, M., Korecka, A., Bakocevic, N., Guan, N.L., Kundu, P., Gulyás, B., Halldin, C., Hultenby, K., Nilsson, H., Hebert, H., Volpe, B.T., Diamond, B., Pettersson, S., 2014. The gut microbiota influences blood-brain barrier permeability in mice. *Sci. Transl. Med.* <https://doi.org/10.1126/scitranslmed.3009759>, 6, 263ra158–263ra158.



- Brás, J.P., Bravo, J., Freitas, J., Barbosa, M.A., Santos, S.G., Summavielle, T., Almeida, M. I., 2020. TNF- $\alpha$ -induced microglia activation requires miR-342: impact on NF- $\kappa$ B signaling and neurotoxicity. *Cell Death Dis.* 11, 415. <https://doi.org/10.1038/s41419-020-2626-6>.
- Burguillos, M.A., Deierborg, T., Kavanagh, E., Persson, A., Hajji, N., Garcia-Quintanilla, A., Cano, J., Brundin, P., Englund, E., Venero, J.L., Joseph, B., 2011. Caspase signalling controls microglia activation and neurotoxicity. *Nature* 472, 319–324. <https://doi.org/10.1038/nature09788>.
- Busillo, J.M., Azzams, K.M., Cidlowski, J.A., 2011. Glucocorticoids sensitize the innate immune system through regulation of the NLRP3 inflammasome. *J. Biol. Chem.* 286, 38703–38713. <https://doi.org/10.1074/jbc.M111.275370>.
- Cai, D., Bin, Zheng, W., Zhang, Q.E., Ng, C.H., Ungvari, G.S., Huang, X., Xiang, Y.T., 2020. Minocycline for depressive symptoms: a meta-analysis of randomized, double-blind, placebo-controlled trials. *Psychiatr. Q* 91, 451–461. <https://doi.org/10.1007/s1126-019-09707-3>.
- Calabrese, F., Rossetti, A.C., Racagni, G., Gass, P., Riva, M.A., Molteni, R., 2014. Brain-derived neurotrophic factor: a bridge between inflammation and neuroplasticity. *Front Cell Neurosci.* 8, 430. <https://doi.org/10.3389/fncel.2014.00430>.
- Calder, P.C., 2017. Omega-3 fatty acids and inflammatory processes: from molecules to man. *Biochem Soc. Trans.* 45, 1105–1115. <https://doi.org/10.1042/BST20160474>.
- Callaghan, C.K., Rouine, J., O'Mara, S.M., 2017. Exercise prevents IFN- $\alpha$ -induced mood and cognitive dysfunction and increases BDNF expression in the rat. *Physiol. Behav.* 179, 377–383. <https://doi.org/10.1016/j.physbeh.2017.07.018>.
- Camilleri, M., 2019. Leaky gut: mechanisms, measurement and clinical implications in humans. *Gut* 68, 1516–1526. <https://doi.org/10.1136/gutjnl-2019-318427>.
- Capuron, L., Raison, C.L., Musselman, D.L., Lawson, D.H., Nemeroff, C.B., Miller, A.H., 2003. Association of exaggerated HPA axis response to the initial injection of interferon- $\alpha$  with development of depression during interferon- $\alpha$  therapy. *Am. J. Psychiatry* 160, 1342–1345. <https://doi.org/10.1176/appi.ajp.160.7.1342>.
- Carless, A.S., Borba, L.A., Zugno, A.I., Quevedo, J., Réus, G.Z., 2021. Gut microbiota-brain axis in depression: the role of neuroinflammation. *Eur. J. Neurosci.* 53 (1), 222–235. <https://doi.org/10.1111/ejn.14631>.
- Cavaleri, D., Bartoli, F., Capogrosso, C.A., Guzzi, P., Moretti, F., Riboldi, I., Misiak, B., Kishi, T., Rubin, R.T., Fuchs, D., Crocarno, C., Carrà, G., 2023. Blood concentrations of neopterin and biopterin in subjects with depression: a systematic review and meta-analysis. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 120, 110633. <https://doi.org/10.1016/j.pnpbp.2022.110633>.
- Chassaing, B., Kumar, M., Baker, M.T., Singh, V., Vijay-Kumar, M., 2014. Mammalian gut immunity. *Biomed. J.* 37, 246–258. <https://doi.org/10.4103/2319-4170.130922>.
- Chen, M.H., Li, C.T., Lin, W.C., Hong, C.J., Tu, P.C., Bai, Y.M., Cheng, C.M., Su, T.P., 2018. Rapid inflammation modulation and antidepressant efficacy of a low-dose ketamine infusion in treatment-resistant depression: a randomized, double-blind control study. *Psychiatry Res* 269, 207–211. <https://doi.org/10.1016/j.psychres.2018.08.078>.
- Chen, Y., Qin, C., Huang, J., Tang, X., Liu, C., Huang, K., Xu, J., Guo, G., Tong, A., Zhou, L., 2020. The role of astrocytes in oxidative stress of central nervous system: a mixed blessing. *Cell Prolif.* 53, e12781. <https://doi.org/10.1111/cpr.12781>.
- Chiu, Y.C., Yang, B.H., Yang, K.C., Liu, M.N., Hu, L.Y., Liou, Y.J., Chan, L.Y., Chou, Y.H., 2021. A study of tryptophan, kynurenine and serotonin transporter in first-episode drug-naïve major depressive disorder. *Psychiatry Res Neuroimaging* 312, 111296. <https://doi.org/10.1016/j.pscychres.2021.111296>.
- Choi, S.H., Aid, S., Kim, H.W., Jackson, S.H., Bosetti, F., 2012. Inhibition of NADPH oxidase promotes alternative and anti-inflammatory microglial activation during neuroinflammation. *J. Neurochem* 120, 292–301. <https://doi.org/10.1111/j.1471-4159.2011.07572.x>.
- Chu, A.L., Hickman, M., Steel, N., Jones, P.B., Davey Smith, G., Khandaker, G.M., 2021. Inflammation and depression: a public health perspective. *Brain Behav. Immun.* 95, 1–3. <https://doi.org/10.1016/j.bbi.2021.04.015>.
- Ciocan, D., Cassard, A.M., Bequemont, L., Verstuyft, C., Voican, C.S., El, Asmar, K., Colle, R., David, D., Trabado, S., Feve, B., Chanson, P., Perlemuter, G., Corruble, E., 2021. Blood microbiota and metabolic signature of major depression before and after antidepressant treatment: a prospective case-control study. *J. Psychiatry Neurosci.* 46 (3), E358–E368. <https://doi.org/10.1503/jpn.200159>.
- Colasanto, M., Madigan, S., Korczak, D.J., 2020. Depression and inflammation among children and adolescents: a meta-analysis. *J. Affect Disord.* 277, 940–948. <https://doi.org/10.1016/j.jad.2020.09.025>.
- Colombo, E., Farina, C., 2016. Astrocytes: key regulators of neuroinflammation. *Trends Immunol.* 37, 608–620. <https://doi.org/10.1016/j.it.2016.06.006>.
- Costantini, L., Molinari, R., Farinon, B., Merendino, N., 2017. Impact of omega-3 fatty acids on the gut microbiota. *Int. J. Mol. Sci.* 18, 2645. <https://doi.org/10.3390/ijms18122645>.
- Cotter, D., Mackay, D., Chana, G., Beasley, C., Landau, S., Everall, I.P., 2002. Reduced neuronal size and glial cell density in area 9 of the dorsolateral prefrontal cortex in subjects with major depressive disorder. *Cereb. Cortex* 12, 386–394. <https://doi.org/10.1093/cercor/12.4.386>.
- Cox, A.J., West, N.P., Cripps, A.W., 2015. Obesity, inflammation, and the gut microbiota. *Lancet Diabetes Endocrinol.* 3, 207–215. [https://doi.org/10.1016/S2213-8587\(14\)70134-2](https://doi.org/10.1016/S2213-8587(14)70134-2).
- Cruz-Topete, D., Cidlowski, J.A., 2014. One hormone, two actions: anti- and pro-inflammatory effects of glucocorticoids. *Neuroimmunomodulation* 22, 20–32. <https://doi.org/10.1159/000362724>.
- Cryan, J.F., O'Riordan, K.J., Cowan, C.S.M., Sandhu, K.V., Bastiaansen, T.F.S., Boehme, M., Codagnone, M.G., Cusotto, S., Fulling, C., Golubeva, A.V., Guzzetta, K.E., Jaggar, M., Long-Smith, C.M., Lyte, J.M., Martin, J.A., Molinero-Perez, A., Moloney, G., Morelli, E., Morillas, E., O'Connor, R., Cruz-Pereira, J.S., Peterson, V.L., Rea, K., Ritz, N.L., Sherwin, E., Spichak, S., Teichman, E.M., van de Wouw, M., Ventura-Silva, A.P., Wallace-Fitzsimons, S.E., Hyland, N., Clarke, G., Dinan, T.G., 2019. The microbiota-gut-brain axis. *Physiol. Rev.* 99, 1877–2013. <https://doi.org/10.1152/physrev.00018.2018>.
- D'Sa, C., Duman, R.S., 2002. Antidepressants and neuroplasticity. *Bipolar Disord.* 4, 183–194. <https://doi.org/10.1034/j.1399-5618.2002.01203.x>.
- Dai, L., Chen, L., Wang, W., 2020. Safety and efficacy of saffron (*Crocus sativus* L.) for treating mild to moderate depression: a systematic review and meta-analysis. *J. Nerv. Ment. Dis.* 208, 269–276. <https://doi.org/10.1097/NMD.0000000000001118>.
- Dantzer, R., O'Connor, J.C., Freund, G.G., Johnson, R.W., Kelley, K.W., 2008. From inflammation to sickness and depression: when the immune system subjugates the brain. *Nat. Rev. Neurosci.* 9, 46–56. <https://doi.org/10.1038/nrn2297>.
- Dantzer, R., O'Connor, J.C., Lawson, M.A., Kelley, K.W., 2011. Inflammation-associated depression: From serotonin to kynurenine. *Psychoneuroendocrinology* 36, 426–436. <https://doi.org/10.1016/j.psyneuen.2010.09.012>.
- Dean, O.M., Data-Franco, J., Giorlando, F., Berk, M., 2012. Minocycline: therapeutic potential in psychiatry. *CNS Drugs* 26, 391–401. <https://doi.org/10.2165/11632000-000000000-00000>.
- Deng, J., Zhou, F., Hou, W., Silver, Z., Wong, C.Y., Chang, O., Huang, E., Zuo, Q.K., 2021. The prevalence of depression, anxiety, and sleep disturbances in COVID-19 patients: a meta-analysis. *Ann. N. Y. Acad. Sci.* 1486, 90–111. <https://doi.org/10.1111/nyas.14506>.
- Ding, Y., Gao, Z.G., Jacobson, K.A., Suffredini, A.F., 2010. Dexamethasone enhances ATP-induced inflammatory responses in endothelial cells. *J. Pharm. Exp. Ther.* 335, 693–702. <https://doi.org/10.1124/jpet.110.171975>.
- Dou, Y., Luo, J., Wu, X., Wei, Z., Tong, B., Yu, J., Wang, T., Zhang, X., Yang, Y., Yuan, X., Zhao, P., Xia, Y., Hu, H., Dai, Y., 2018. Curcumin attenuates collagen-induced inflammatory response through the "gut-brain axis. *J. Neuroinflamm.* 15, 1–15. <https://doi.org/10.1186/s12974-017-1047-7>.
- Dowlati, Y., Herrmann, N., Swardfager, W., Liu, H., Sham, L., Reim, E.K., Lancôt, K.L., 2010. A meta-analysis of cytokines in major depression. *Biol. Psychiatry* 67, 446–457. <https://doi.org/10.1016/j.biopsych.2009.09.033>.
- Ekdahl, C.T., Claassen, J.H., Bonde, S., Kokaia, Z., Lindvall, O., 2003. Inflammation is detrimental for neurogenesis in adult brain. *Proc. Natl. Acad. Sci. USA* 100, 13632–13637. <https://doi.org/10.1073/pnas.2234031100>.
- Ekdahl, C.T., Kokaia, Z., Lindvall, O., 2009. Brain inflammation and adult neurogenesis: the dual role of microglia. *Neuroscience* 158, 1021–1029. <https://doi.org/10.1016/j.neuroscience.2008.06.052>.
- Enache, D., Pariante, C.M., Mondelli, V., 2019. Markers of central inflammation in major depressive disorder: a systematic review and meta-analysis of studies examining cerebrospinal fluid, positron emission tomography and post-mortem brain tissue. *Brain Behav. Immun.* 81, 24–40. <https://doi.org/10.1016/j.bbi.2019.06.015>.
- Fang, K., Li, H.R., Chen, X.X., Gao, X.R., Huang, L.L., Du, A.Q., Jiang, C., Li, H., Ge, J.F., 2020. Quercetin alleviates LPS-induced depression-like behavior in rats via regulating BDNF-related imbalance of copine 6 and TREM1/2 in the hippocampus and PFC. *Front Pharm.* 10, 1544. <https://doi.org/10.3389/fphar.2019.01544/FULL>.
- Faridhosseini, F., Sadeghi, R., Farid, L., Pourgholami, M., 2014. Celecoxib: a new augmentation strategy for depressive mood episodes. A systematic review and meta-analysis of randomized placebo-controlled trials. *Hum. Psychopharmacol.* 29, 216–223. <https://doi.org/10.1002/hup.2401>.
- Fond, G., Loundou, A., Rabu, C., Macgregor, A., Lançon, C., Brittner, M., Micoulaud-Franchi, J.A., Richieri, R., Courtet, P., Abbar, M., Roger, M., Leboyer, M., Boyer, L., 2014. Ketamine administration in depressive disorders: a systematic review and meta-analysis. *Psychopharmacol. (Berl.)* 231, 3663–3676. <https://doi.org/10.1007/s00213-014-3664-5>.
- Foster, J.A., McVey Neufeld, K.A., 2013. Gut-brain axis: how the microbiome influences anxiety and depression. *Trends Neurosci.* 36, 305–312. <https://doi.org/10.1016/j.tins.2013.01.005>.
- Fujioka, H., Akema, T., 2010. Lipopolysaccharide acutely inhibits proliferation of neural precursor cells in the dentate gyrus in adult rats. *Brain Res* 1352, 35–42. <https://doi.org/10.1016/j.brainres.2010.07.032>.
- Furman, D., Campisi, J., Verdín, E., Carrera-Bastos, P., Targ, S., Franceschi, C., Ferrucci, L., Gilroy, D.W., Fasano, A., Miller, G.W., Miller, A.H., Mantovani, A., Weyand, C.M., Barzilai, N., Goronzy, J.J., Rando, T.A., Effros, R.B., Lucia, A., Kleinstreuer, N., Slavich, G.M., 2019. Chronic inflammation in the etiology of disease across the life span. *Nat. Med.* 25, 1822–1832. <https://doi.org/10.1038/s41591-019-0675-0>.
- Fusar-Poli, L., Vozza, L., Gabbadini, A., Vanella, A., Concas, I., Tinacci, S., Petralia, A., Signorelli, M.S., Aguglia, E., 2019. Curcumin for depression: a meta-analysis. *Crit. Rev. Food Sci. Nutr.* 60, 2643–2653. <https://doi.org/10.1080/10408398.2019.1653260>.
- Ginhoux, F., Greter, M., Leboeuf, M., Nandi, S., See, P., Gokhan, S., Mehler, M.F., Conway, S.J., Ng, L.G., Stanley, E.R., Samokhvalov, I.M., Merad, M., 2010. Fate mapping analysis reveals that adult microglia derive from primitive macrophages. *Science* 330, 841–845. <https://doi.org/10.1126/science.1194637>.
- Giorgi, R., De Crescenzo, F., De Pesci, N.R., Martens, M., Howard, W., Cowen, P.J., Harmer, C.J., 2021. Statins for major depressive disorder: a systematic review and meta-analysis of randomized controlled trials. *PLoS One* 16, e0249409. <https://doi.org/10.1371/journal.pone.0249409>.
- Gisterà, A., Hansson, G.K., 2017. The immunology of atherosclerosis. *Nat. Rev. Nephrol.* 13, 368–380. <https://doi.org/10.1038/nrneph.2017.51>.
- Gittins, R.A., Harrison, P.J., 2011. A morphometric study of glia and neurons in the anterior cingulate cortex in mood disorder. *J. Affect Disord.* 133, 328–332. <https://doi.org/10.1016/j.jad.2011.03.042>.



- Golia, M.T., Poggini, S., Alboni, S., Garofalo, S., Ciano Albanese, N., Viglione, A., Ajmone-Cat, M.A., St-Pierre, A., Brunello, N., Limatola, C., Branchi, I., Maggi, L., 2019. Interplay between inflammation and neural plasticity: both immune activation and suppression impair LTP and BDNF expression. *Brain Behav. Immun.* 81, 484–494. <https://doi.org/10.1016/j.bbi.2019.07.003>.
- Goodhand, J.R., Wahed, M., Mawdsley, J.E., Farmer, A.D., Aziz, Q., Rampton, D.S., 2012. Mood disorders in inflammatory bowel disease: Relation to diagnosis, disease activity, perceived stress, and other factors. *Inflamm. Bowel Dis.* 18, 2301–2309. <https://doi.org/10.1002/ibd.22916>.
- Grosso, G., Pajak, A., Marventano, S., Castellano, S., Galvano, F., Bucolo, C., Drago, F., Caraci, F., 2014. Role of omega-3 fatty acids in the treatment of depressive disorders: a comprehensive meta-analysis of randomized clinical trials. *PLoS One* 9, e96905. <https://doi.org/10.1371/journal.pone.0096905>.
- Guan, Z., Fang, J., 2006. Peripheral immune activation by lipopolysaccharide decreases neurotrophins in the cortex and hippocampus in rats. *Brain Behav. Immun.* 20, 64–71. <https://doi.org/10.1016/j.bbi.2005.04.005>.
- Guillemin, G.J., Smythe, G., Takikawa, O., Brew, B.J., 2005. Expression of indoleamine 2,3-dioxygenase and production of quinolinic acid by human micro. *Glia, Astro, Neurons* 49, 15–23. <https://doi.org/10.1002/glia.20090>.
- Guillot, X., Semerano, L., Saldenber-Kermanac'h, N., Falgarone, G., Boissier, M.C., 2010. Vitamin D and inflammation. *Jt Bone Spine* 77, 552–557. <https://doi.org/10.1016/j.jbspin.2010.09.018>.
- Haghighatdoost, F., Jabbari, M., Hariri, M., 2019. The effect of L-carnitine on inflammatory mediators: a systematic review and meta-analysis of randomized clinical trials. *Eur. J. Clin. Pharm.* 75, 1037–1046. <https://doi.org/10.1007/s00228-019-02666-5>.
- Hajiluan, G., Nameni, G., Shahabi, P., Mesgari-Abbasi, M., Sadigh-Eteghad, S., Farhang, M.A., 2017. Vitamin D administration, cognitive function, BBB permeability and neuroinflammatory factors in high-fat diet-induced obese rats. *Int J. Obes.* 41, 639–644. <https://doi.org/10.1038/ijo.2017.10>.
- Haroon, E., Welle, J.R., Woolwine, B.J., Goldsmith, D.R., Baer, W., Patel, T., Felger, J.C., Miller, A.H., 2020. Associations among peripheral and central kynurenine pathway metabolites and inflammation in depression. *Neuropsychopharmacology* 45, 998–1007. <https://doi.org/10.1038/s41386-020-0607-1>.
- Hashimoto, K., 2023. Neuroinflammation through the vagus nerve-dependent gut-microbiota-brain axis in treatment-resistant depression. *Prog. Brain Res* 278, 61–77. <https://doi.org/10.1016/bs.pbr.2023.01.003>.
- Hauser, P., Khosla, J., Aurora, H., Laurin, J., Kling, M.A., Hill, J., Gulati, M., Thornton, A. J., Schultz, R.L., Valentine, A.D., Meyers, C.A., Howell, C.D., 2002. A prospective study of the incidence and open-label treatment of interferon-induced major depressive disorder in patients with hepatitis C. *Mol. Psychiatry* 7, 942–947. <https://doi.org/10.1038/sj.mp.4001119>.
- Häuser, W., Janke, K.H., Klump, B., Hinz, A., 2011. Anxiety and depression in patients with inflammatory bowel disease: comparisons with chronic liver disease patients and the general population. *Inflamm. Bowel Dis.* 17, 621–632. <https://doi.org/10.1002/ibd.21346>.
- Hermoso, M.A., Matsuguchi, T., Smoak, K., Cidlowski, J.A., 2004. Glucocorticoids and tumor necrosis factor alpha cooperatively regulate toll-like receptor 2 gene expression. *Mol. Cell Biol.* 24, 4743–4756. <https://doi.org/10.1128/mcb.24.11.4743-4756.2004>.
- Hongxia, L., Yuxiao, T., Zhilei, S., Yan, S., Yicui, Q., Jiamin, S., Xin, X., Jianxin, Y., Fengling, M., Hui, S., 2019. Zinc inhibited LPS-induced inflammatory responses by upregulating A20 expression in microglia BV2 cells. *J. Affect Disord.* 249, 136–142. <https://doi.org/10.1016/j.jad.2019.02.041>.
- Hoofnagle, J.H., 2002. Course and outcome of hepatitis C. *Hepatology* 36, s21–s29. <https://doi.org/10.1053/jhep.2002.36227>.
- Horikawa, N., Yamazaki, T., Izumi, N., Uchihara, M., 2003. Incidence and clinical course of major depression in patients with chronic hepatitis type C undergoing interferon-alpha therapy: a prospective study. *Gen. Hosp. Psychiatry* 25, 34–38. [https://doi.org/10.1016/S0163-8343\(02\)00239-6](https://doi.org/10.1016/S0163-8343(02)00239-6).
- Howren, M.B., Lamkin, D.M., Suls, J., 2009. Associations of depression with c-reactive protein, IL-1, and IL-6: a meta-analysis. *Psychosom. Med* 71, 171–186. <https://doi.org/10.1097/PSY.0b013e3181907c1b>.
- Huang, W.T., Niu, K.C., Chang, C.K., Lin, M.T., Chang, C.P., 2008. Curcumin inhibits the increase of glutamate, hydroxyl radicals and PGE2 in the hypothalamus and reduces fever during LPS-induced systemic inflammation in rabbits. *Eur. J. Pharm.* 593, 105–111. <https://doi.org/10.1016/j.ejphar.2008.07.017>.
- Hunt, C., Macedo e Cordeiro, T., Suchting, R., de Dios, C., Cuellar Leal, V.A., Soares, J.C., Dantzer, R., Teixeira, A.L., Selvaraj, S., 2020. Effect of immune activation on the kynurenine pathway and depression symptoms - a systematic review and meta-analysis. *Neurosci. Biobehav. Rev.* 118, 514–523. <https://doi.org/10.1016/j.neubiorev.2020.08.010>.
- Ja, W.K., Duman, R.S., 2008. IL-1 $\beta$  is an essential mediator of the antineurogenic and anhedonic effects of stress. *Proc. Natl. Acad. Sci. USA* 105, 751–756. <https://doi.org/10.1073/pnas.0708092105>.
- Jack, C.S., Arbour, N., Manusow, J., Montgrain, V., Blain, M., McCrea, E., Shapiro, A., Antel, J.P., 2005. TLR signaling tailors innate immune responses in human microglia and astrocytes. *J. Immunol.* 175, 4320–4330. <https://doi.org/10.4049/jimmunol.175.7.4320>.
- Jamali-Raeufy, N., Alizadeh, F., Mehrabi, Z., Mehrabi, S., Goudarzi, M., 2021. Acetyl-L-carnitine confers neuroprotection against lipopolysaccharide (LPS)-induced neuroinflammation by targeting TLR4/NF $\kappa$ B, autophagy, inflammation and oxidative stress. *Metab. Brain Dis.* 1–11. <https://doi.org/10.1007/s11011-021-00715-6>.
- Jamilian, H., Amirani, E., Milajerdi, A., Kolahdooz, F., Mirzaei, H., Zaroudi, M., Ghaderi, A., Asemi, Z., 2019. The effects of vitamin D supplementation on mental health, and biomarkers of inflammation and oxidative stress in patients with psychiatric disorders: A systematic review and meta-analysis of randomized controlled trials. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 94, 109651. <https://doi.org/10.1016/j.pnpbp.2019.109651>.
- Jiang, H., Ling, Z., Zhang, Y., Mao, H., Ma, Z., Yin, Y., Wang, W., Tang, W., Tan, Z., Shi, J., Li, L., Ruan, B., 2015. Altered fecal microbiota composition in patients with major depressive disorder. *Brain Behav. Immun.* 48, 186–194. <https://doi.org/10.1016/j.bbi.2015.03.016>.
- Jiang, J., Wang, W., Sun, Y.J., Hu, M., Li, F., Zhu, D.Y., 2007. Neuroprotective effect of curcumin on focal cerebral ischemic rats by preventing blood-brain barrier damage. *Eur. J. Pharm.* 561, 54–62. <https://doi.org/10.1016/j.ejphar.2006.12.028>.
- Jurenka, J.S., 2009. Anti-inflammatory properties of curcumin, a major constituent of *Curcuma longa*: a review of preclinical and clinical research - PubMed. *Alter. Med Rev.* 14, 141–153.
- Kamada, N., Seo, S.U., Chen, G.Y., Núñez, G., 2013. Role of the gut microbiota in immunity and inflammatory disease. *Nat. Rev. Immunol.* 13, 321–335. <https://doi.org/10.1038/nri3430>.
- Kaneko, N., Kudo, K., Mabuchi, T., Takemoto, K., Fujimaki, K., Wati, H., Iguchi, H., Tezuka, H., Kanba, S., 2006. Suppression of cell proliferation by interferon-alpha through interleukin-1 production in adult rat dentate gyrus. *Neuropsychopharmacology* 31, 2619–2626. <https://doi.org/10.1038/sj.npp.1301137>.
- Kassmann, C.M., Lappe-Siefke, C., Baes, M., Brügger, B., Mildner, A., Werner, H.B., Natt, O., Michaelis, T., Prinz, M., Frahm, J., Nave, K.A., 2007. Axonal loss and neuroinflammation caused by peroxisome-deficient oligodendrocytes. *Nat. Genet* 39, 969–976. <https://doi.org/10.1038/ng2070>.
- Kaufmann, F.N., Costa, A.P., Ghisleni, G., Diaz, A.P., Rodrigues, A.L.S., Peluffo, H., Kaster, M.P., 2017. NLRP3 inflammasome-driven pathways in depression: clinical and preclinical findings. *Brain Behav. Immun.* 64, 367–383. <https://doi.org/10.1016/j.bbi.2017.03.002>.
- Kavanagh, K., Hsu, F.C., Davis, A.T., Kritchevsky, S.B., Rejeski, W.J., Kim, S., 2019. Biomarkers of leaky gut are related to inflammation and reduced physical function in older adults with cardiometabolic disease and mobility limitations. *Geroscience* 41, 923–933. <https://doi.org/10.1007/s11357-019-00112-z>.
- Kawai, T., Akira, S., 2010. The role of pattern-recognition receptors in innate immunity: update on toll-like receptors. *Nat. Immunol.* 11, 373–384. <https://doi.org/10.1038/ni.1863>.
- Kazak, F., Yarim, G.F., 2017. Neuroprotective effects of acetyl-L-carnitine on lipopolysaccharide-induced neuroinflammation in mice: Involvement of brain-derived neurotrophic factor. *Neurosci. Lett.* 658, 32–36. <https://doi.org/10.1016/j.neulet.2017.07.059>.
- Kelley, K.W., Bluthé, R.M., Dantzer, R., Zhou, J.H., Shen, W.H., Johnson, R.W., Broussard, S.R., 2003. Cytokine-induced sickness behavior. *Brain Behav. Immun.* 17, 112–118. [https://doi.org/10.1016/S0889-1591\(02\)00077-6](https://doi.org/10.1016/S0889-1591(02)00077-6).
- Kelly, J.R., Borre, Y., O' Brien, C., Patterson, A.E., El Aidy, S., Deane, J., Kennedy, P.J., Beers, S., Scott, K., Moloney, G., Hoban, A.E., Scott, L., Fitzgerald, P., Ross, P., Stanton, C., Clarke, G., Cryan, J.F., Dinan, T.G., 2016. Transferring the blues: depression-associated gut microbiota induces neurobehavioural changes in the rat. *J. Psychiatr. Res.* 82, 109–118. <https://doi.org/10.1016/j.jpsychires.2016.07.019>.
- Keohane, A., Ryan, S., Maloney, E., Sullivan, A.M., Nolan, Y.M., 2010. Tumour necrosis factor- $\alpha$  impairs neuronal differentiation but not proliferation of hippocampal neural precursor cells: role of Hes1. *Mol. Cell Neurosci.* 43, 127–135. <https://doi.org/10.1016/j.mcn.2009.10.003>.
- Kiecolt-Glaser, J.K., Wilson, S.J., Bailey, M.L., Andridge, R., Peng, J., Jaremka, L.M., Fagundes, C.P., Malarkey, W.B., Laskowski, B., Belury, M.A., 2018. Marital distress, depression, and a leaky gut: translocation of bacterial endotoxin as a pathway to inflammation. *Psychoneuroendocrinology* 98, 52–60. <https://doi.org/10.1016/j.psyneuen.2018.08.007>.
- Kim, S.W., Kang, H.J., Jhon, M., Kim, J.W., Lee, J.Y., Walker, A.J., Agustini, B., Kim, J. M., Berk, M., 2019. Statins and inflammation: new therapeutic opportunities in psychiatry. *Front Psychiatry* 10, 103. <https://doi.org/10.3389/fpsy.2019.00103>.
- Kim, Y.K., Na, K.S., Myint, A.M., Leonard, B.E., 2016. The role of pro-inflammatory cytokines in neuroinflammation, neurogenesis and the neuroendocrine system in major depression. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 64, 277–284. <https://doi.org/10.1016/j.pnpbp.2015.06.008>.
- Kirsten, T.B., Cabral, D., Galvão, M.C., Monteiro, R., Bondan, E.F., Bernardi, M.M., 2020. Zinc, but not paracetamol, prevents depressive-like behavior and sickness behavior, and inhibits interferon-gamma and astrogliosis in rats. *Brain Behav. Immun.* 87, 489–497. <https://doi.org/10.1016/j.bbi.2020.01.019>.
- Klawonn, A.M., Fritz, M., Castany, S., Pignatelli, M., Canal, C., Similä, F., Tejada, H.A., Levinsson, J., Jaarola, M., Jakobsson, J., Hidalgo, J., Heilig, M., Bonci, A., Engblom, D., 2021. Microglial activation elicits a negative affective state through prostaglandin-mediated modulation of striatal neurons. *e6 Immunity* 54, 225–234. <https://doi.org/10.1016/j.immuni.2020.12.016>.
- Kochar, B., Barnes, E.L., Long, M.D., Cushing, K.C., Galanko, J., Martin, C.F., Raffals, L. E., Sandler, R.S., 2018. Depression is associated with more aggressive inflammatory bowel disease. *Am. J. Gastroenterol.* 113, 80–85. <https://doi.org/10.1038/ajg.2017.423>.
- Köhler, C.A., Freitas, T.H., Maes, M., de Andrade, N.Q., Liu, C.S., Fernandes, B.S., Stubbs, B., Solmi, M., Veronese, N., Herrmann, N., Raison, C.L., Miller, B.J., Lancôt, K.L., Carvalho, A.F., 2017. Peripheral cytokine and chemokine alterations in depression: a meta-analysis of 82 studies. *Acta Psychiatr. Scand.* 135, 373–387. <https://doi.org/10.1111/acps.12698>.
- Köhler, C.A., Freitas, T.H., Stubbs, B., Maes, M., Solmi, M., Veronese, N., de Andrade, N. Q., Morris, G., Fernandes, B.S., Brunoni, A.R., Herrmann, N., Raison, C.L., Miller, B. J., Lancôt, K.L., Carvalho, A.F., 2018. Peripheral alterations in cytokine and

- chemokine levels after antidepressant drug treatment for major depressive disorder: systematic review and meta-analysis. *Mol. Neurobiol.* 55, 4195–4206. <https://doi.org/10.1007/s12035-017-0632-1>.
- Köhler, O., Benros, E., Nordentoft, M., Farkouh, M., Iyengar, M.E., Mors, R.L., Krogh, J., O., 2014. Effect of anti-inflammatory treatment on depression, depressive symptoms, and adverse effects: a systematic review and meta-analysis of randomized clinical trials. *JAMA Psychiatry* 71, 1381–1391. <https://doi.org/10.1001/jamapsychiatry.2014.1611>.
- Köhler-Forsberg, O., Lydholm, N., Hjorthøj, C., Nordentoft, C., Mors, M., Benros, M.E., O., 2019. Efficacy of anti-inflammatory treatment on major depressive disorder or depressive symptoms: meta-analysis of clinical trials. *Acta Psychiatr. Scand.* 139, 404–419. <https://doi.org/10.1111/acps.13016>.
- Kohman, R.A., Rhodes, J.S., 2013. Neurogenesis, inflammation and behavior. *Brain Behav. Immun.* 27, 22–32. <https://doi.org/10.1016/j.bbi.2012.09.003>.
- Kopra, E., Mondelli, V., Pariante, C., Nikkheslat, N., 2021. Ketamine's effect on inflammation and kynurenine pathway in depression: a systematic review. *J. Psychopharmacol.* 35, 934–945. <https://doi.org/10.1177/02698811211026426>.
- Kurina, L.M., Goldacre, M.J., Yeates, D., Gill, L.E., 2001. Depression and anxiety in people with inflammatory bowel disease. *J. Epidemiol. Community Health* 55 (1978), 716–720. <https://doi.org/10.1136/jech.55.10.716>.
- Layé, S., Nadjar, A., Joffre, C., Bazinet, R.P., 2018. Anti-inflammatory effects of omega-3 fatty acids in the brain: physiological mechanisms and relevance to pharmacology. *Pharm. Rev.* 70, 12–38. <https://doi.org/10.1124/pr.117.014092>.
- Leboyer, M., Berk, M., Yolken, R.H., Tamouza, R., Kupfer, D., Groc, L., 2016. Immunopsychiatry: an agenda for clinical practice and innovative research. *BMC Med* 14 (1), 173. <https://doi.org/10.1186/s12916-016-0712-5>.
- Leng, F., Edison, P., 2021. Neuroinflammation and microglial activation in Alzheimer disease: where do we go from here. *Nat. Rev. Neurol.* 17, 157–172. <https://doi.org/10.1038/s41582-020-00435-y>.
- Li, C., Zhao, B., Lin, C., Gong, Z., An, X., 2019. TREM2 inhibits inflammatory responses in mouse microglia by suppressing the PI3K/NF- $\kappa$ B signaling. *Cell Biol. Int* 43, 360–372. <https://doi.org/10.1002/cbin.10975>.
- Li, N., Wang, Q., Wang, Y., Sun, A., Lin, Y., Jin, Y., Li, X., 2019. Fecal microbiota transplantation from chronic unpredictable mild stress mice donors affects anxiety-like and depression-like behavior in recipient mice via the gut microbiota-inflammation-brain axis. *Stress* 22, 592–602. <https://doi.org/10.1080/10253890.2019.1617267>.
- Li, Z., Li, B., Song, X., Zhang, D., 2017. Dietary zinc and iron intake and risk of depression: a meta-analysis. *Psychiatry Res* 251, 41–47. <https://doi.org/10.1016/j.psychres.2017.02.006>.
- Liao, Y., Xie, B., Zhang, H., He, Q., Guo, L., Subramaniapillai, M., Fan, B., Lu, C., McIntyer, R.S., 2019. Efficacy of omega-3 PUFAs in depression: a meta-analysis. *Transl. Psychiatry* 9, 1–9. <https://doi.org/10.1038/s41398-019-0515-5>.
- Liddelow, S.A., Guttenplan, K.A., Clarke, L.E., Bennett, F.C., Bohlen, C.J., Schirmer, L., Bennett, M.L., Münch, A.E., Chung, W.S., Peterson, T.C., Wilton, D.K., Frouin, A., Napier, B.A., Panicker, N., Kumar, M., Buckwalter, M.S., Rowitch, D.H., Dawson, V. L., Dawson, T.M., Stevens, B., Barres, B.A., 2017. Neurotoxic reactive astrocytes are induced by activated microglia. *Nature* 541, 481–487. <https://doi.org/10.1038/nature21029>.
- Liu, D., Wang, Z., Liu, S., Wang, F., Zhao, S., Hao, A., 2011. Anti-inflammatory effects of fluoxetine in lipopolysaccharide(LPS)-stimulated microglial cells. *Neuropharmacology* 61, 592–599. <https://doi.org/10.1016/j.neuropharm.2011.04.033>.
- Liu, J.J., Wei, Y., Bin, Strawbridge, R., Bao, Y., Chang, S., Shi, L., Que, J., Gadad, B.S., Trivedi, M.H., Kelsoe, J.R., Lu, L., 2020. Peripheral cytokine levels and response to antidepressant treatment in depression: a systematic review and meta-analysis. *Mol. Psychiatry* 25, 339–350. <https://doi.org/10.1038/s41380-019-0474-5>.
- Liu, R.T., Walsh, R.F.L., Sheehan, A.E., 2019. Prebiotics and probiotics for depression and anxiety: a systematic review and meta-analysis of controlled clinical trials. *Neurosci. Biobehav. Rev.* 102, 13–23. <https://doi.org/10.1016/j.neubiorev.2019.03.023>.
- Lu, Y., Ding, X., Wu, X., Huang, S., 2020. Ketamine inhibits LPS-mediated BV2 microglial inflammation via NMDA receptor blockage. *Fundam. Clin. Pharm.* 34, 229–237. <https://doi.org/10.1111/fcp.12508>.
- Luppino, F.S., De Wit, L.M., Bouvy, P.F., Stijnen, T., Cuijpers, P., Penninx, B.W.J.H., Zitman, F.G., 2010. Overweight, obesity, and depression: a systematic review and meta-analysis of longitudinal studies. *Arch. Gen. Psychiatry* 67, 220–229. <https://doi.org/10.1001/archgenpsychiatry.2010.2>.
- Lynall, M.E., Turner, L., Bhatti, J., Cavanagh, J., de Boer, P., Mondelli, V., Jones, D., Dreves, W.C., Cowen, P., Harrison, N.A., Pariante, C.M., Pointon, L., Clatworthy, M. R., Bullmore, E., 2020. Peripheral blood cell-stratified subgroups of inflamed depression. *Biol. Psychiatry* 88, 185–196. <https://doi.org/10.1016/j.biopsych.2019.11.017>.
- Maes, M., Kubera, M., Leunis, J.-C., 2008. The gut-brain barrier in major depression: Intestinal mucosal dysfunction with an increased translocation of LPS from gram negative enterobacteria (leaky gut) plays a role in the inflammatory pathophysiology of depression. *Neuro Endocrinol. Lett.* 29, 117–124.
- Malhi, G.S., Mann, J.J., 2018. Depression. *Lancet* 392, 2299–2312. [https://doi.org/10.1016/S0140-6736\(18\)31948-2](https://doi.org/10.1016/S0140-6736(18)31948-2).
- Mannan, M., Mamun, A., Doi, S., Clavarino, A., 2016. Is there a bi-directional relationship between depression and obesity among adult men and women? systematic review and bias-adjusted meta analysis. *Asian J. Psychiatr.* 21, 51–66. <https://doi.org/10.1016/j.ajp.2015.12.008>.
- Maslowski, K.M., Vieira, A.T., Ng, A., Kranich, J., Sierro, F., Di, Yu, Schilter, H.C., Rolph, M.S., MacKay, F., Artis, D., Xavier, R.J., Teixeira, M.M., MacKay, C.R., 2009. Regulation of inflammatory responses by gut microbiota and chemoattractant receptor GPR43. *Nature* 461, 1282–1286. <https://doi.org/10.1038/nature08530>.
- Masuda, T., Sankowski, R., Staszewski, O., Böttcher, C., Amann, L., Sagar, Scheiwe, C., Nessler, S., Kunz, P., van Loo, G., Coenen, V.A., Reinacher, P.C., Michel, A., Sure, U., Gold, R., Grün, D., Priller, J., Stadelmann, C., Prinz, M., 2019. Spatial and temporal heterogeneity of mouse and human microglia at single-cell resolution. *Nature* 566, 388–392. <https://doi.org/10.1038/s41586-019-0924-x>.
- Mazza, M.G., Lucchi, S., Tringali, A.G.M., Rossetti, A., Botti, E.R., Clerici, M., 2018. Neutrophil/lymphocyte ratio and platelet/lymphocyte ratio in mood disorders: a meta-analysis. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 84, 229–236. <https://doi.org/10.1016/j.pnpbp.2018.03.012>.
- McInnes, I.B., Gravalles, E.M., 2021. Immune-mediated inflammatory disease therapeutics: past, present and future. *Nat. Rev. Immunol.* 21, 680–686. <https://doi.org/10.1038/s41577-021-00603-1>.
- Menon, V.P., Sudheer, A.R., 2007. Antioxidant and anti-inflammatory properties of curcumin. *Adv. Exp. Med Biol.* 595, 105–125. [https://doi.org/10.1007/978-0-387-46401-5\\_3](https://doi.org/10.1007/978-0-387-46401-5_3).
- Messaoud, A., Mensi, R., Douki, W., Neffati, F., Najjar, M.F., Gobbi, G., Valtorta, F., Gaha, L., Comai, S., 2019. Reduced peripheral availability of tryptophan and increased activation of the kynurenine pathway and cortisol correlate with major depression and suicide. *World J. Biol. Psychiatry* 20, 703–711. <https://doi.org/10.1080/15622975.2018.1468031>.
- Meyer, J.H., Cervenkova, S., Kim, M.J., Kreis, W.C., Henter, I.D., Innis, R.B., 2020. Neuroinflammation in psychiatric disorders: PET imaging and promising new targets. *Lancet Psychiatry* 7, 1064–1074. [https://doi.org/10.1016/S2215-0366\(20\)30255-8](https://doi.org/10.1016/S2215-0366(20)30255-8).
- Miguel-Hidalgo, J.J., Waltzer, R., Whittom, A.A., Austin, M.C., Rajkowska, G., Stockmeier, C.A., 2010. Glial and glutamatergic markers in depression, alcoholism, and their comorbidity. *J. Affect Disord.* 127, 230–240. <https://doi.org/10.1016/j.jad.2010.06.003>.
- Milajerdi, A., Mousavi, S.M., Sadeghi, A., Salari-Moghaddam, A., Parohan, M., Larijani, B., Esmailzadeh, A., 2020. The effect of probiotics on inflammatory biomarkers: a meta-analysis of randomized clinical trials. *Eur. J. Nutr.* 59, 633–649. <https://doi.org/10.1007/s00394-019-01931-8>.
- Milaneschi, Y., Lamers, F., Berk, M., Penninx, B.W.J.H., 2020. Depression heterogeneity and its biological underpinnings: toward immunometabolic depression. *Biol. Psychiatry* 88, 369–380. <https://doi.org/10.1016/j.biopsych.2020.01.014>.
- Miller, A.H., Raison, C.L., 2016. The role of inflammation in depression: From evolutionary imperative to modern treatment target. *Nat. Rev. Immunol.* 16, 22–34. <https://doi.org/10.1038/nri.2015.5>.
- Miller, A.H., Maletic, V., Raison, C.L., 2009. Inflammation and its discontents: the role of cytokines in the pathophysiology of major depression. *Biol. Psychiatry* 65, 732–741. <https://doi.org/10.1016/j.biopsych.2008.11.029>.
- Mittermaier, C., Dejaco, C., Waldhoer, T., Oefflerbauer-Ernst, A., Michlsler, W., Beier, M., Tillinger, W., Gangl, A., Moser, G., 2004. Impact of depressive mood on relapse in patients with inflammatory bowel disease: a prospective 18-month follow-up study. *Psychosom. Med* 66, 79–84. <https://doi.org/10.1097/01.PSY.0000106907.24881.F2>.
- Mocking, R., Harmsen, I., Assies, J., Koeter, M., Ruhé, H., Schene, A., 2016. Meta-analysis and meta-regression of omega-3 polyunsaturated fatty acid supplementation for major depressive disorder. *Transl. Psychiatry* 6, e756. <https://doi.org/10.1038/tp.2016.29>.
- Mohammad-Zadeh, L.F., Moses, L., Gwaltney-Brant, S.M., 2008. Serotonin: a review. *J. Vet. Pharm. Ther.* 31, 187–199. <https://doi.org/10.1111/j.1365-2885.2008.00944.x>.
- Monje, M.L., Toda, H., Palmer, T.D., 2003. Inflammatory blockade restores adult hippocampal neurogenesis. *Science* 302, 1760–1765. <https://doi.org/10.1126/science.1088417>.
- Moore, S., Meschkat, M., Ruhwedel, T., Trevisiol, A., Tzvetanova, I.D., Battfeld, A., Kusch, K., Kole, M.H.P., Strenke, N., Möbius, W., de Hoz, L., Nave, K.A., 2020. A role of oligodendrocytes in information processing. *Nat. Commun.* 11, 5497. <https://doi.org/10.1038/s41467-020-19152-7>.
- Moras, L.H., Schreiber, H.L., Mazmanian, S.K., 2020. The gut microbiota–brain axis in behaviour and brain disorders. *Nat. Rev. Microbiol.* 19, 241–255. <https://doi.org/10.1038/s41579-020-00460-0>.
- Morris, G., Puri, B.K., Maes, M., Olive, L., Berk, M., Carvalho, A.F., 2020. The role of microglia in neurodegenerative disorders: mechanisms and possible neurotherapeutic effects of induced ketosis. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 99, 109858. <https://doi.org/10.1016/j.pnpbp.2020.109858>.
- Mousa, A., Naderpoor, N., Teede, H., Scragg, R., de Courten, B., 2018. Vitamin D supplementation for improvement of chronic low-grade inflammation in patients with type 2 diabetes: a systematic review and meta-analysis of randomized controlled trials. *Nutr. Rev.* 76, 380–394. <https://doi.org/10.1093/nutri/nux077>.
- Murray, K.O., Resnick, M., Miller, V., 2007. Depression after infection with West Nile virus. *Emerg. Infect. Dis.* 13, 479–481. <https://doi.org/10.3201/eid1303.060602>.
- Myint, A.M., Kim, Y.K., Verkerk, R., Scharpé, S., Steinbusch, H., Leonard, B., 2007. Kynurenine pathway in major depression: evidence of impaired neuroprotection. *J. Affect Disord.* 98, 143–151. <https://doi.org/10.1016/j.jad.2006.07.013>.
- Na, K.S., Lee, K.J., Lee, J.S., Cho, Y.S., Jung, H.Y., 2014. Efficacy of adjunctive celecoxib treatment for patients with major depressive disorder: a meta-analysis. *Neuropsychopharmacol. Biol. Psychiatry* 48, 79–85. <https://doi.org/10.1016/j.pnpbp.2013.09.006>.
- Nave, K.A., Werner, H.B., 2014. Myelination of the nervous system: mechanisms and functions. *Annu. Rev. Cell Dev. Biol.* 30, 503–533. <https://doi.org/10.1146/annurev-cellbio-100913-013101>.
- Nettis, M.A., Lombardo, G., Hastings, C., Zajkowska, Z., Mariani, N., Nikkheslat, N., Worrell, C., Enache, D., McLaughlin, A., Kose, M., Sforzini, L., Bogdanova, A., Cleare, A., Young, A.H., Pariante, C.M., Mondelli, V., 2021. Augmentation therapy

- with minocycline in treatment-resistant depression patients with low-grade peripheral inflammation: results from a double-blind randomised clinical trial. *Neuropsychopharmacology* 46, 939–948. <https://doi.org/10.1038/s41386-020-00948-6>.
- O'Connor, J.C., André, C., Wang, Y., Lawson, M.A., Szegedi, S.S., Lestage, J., Castanon, N., Kelley, K.W., Dantzer, R., 2009a. Interferon- $\gamma$  and tumor necrosis factor- $\alpha$  mediate the upregulation of indoleamine 2,3-dioxygenase and the induction of depressive-like behavior in mice in response to bacillus calmette-guérin. *J. Neurosci.* 29, 4200–4209. <https://doi.org/10.1523/JNEUROSCI.5032-08.2009>.
- O'Connor, J.C., Lawson, M.A., André, C., Moreau, M., Lestage, J., Castanon, N., Kelley, K.W., Dantzer, R., 2009b. Lipopolysaccharide-induced depressive-like behavior is mediated by indoleamine 2,3-dioxygenase activation in mice. *Mol. Psychiatry* 14, 511–522. <https://doi.org/10.1038/sj.mp.4002148>.
- Obrenovich, M., 2018. Leaky Gut, Leaky Brain. *Microorganisms* 6, 107. <https://doi.org/10.3390/microorganisms6040107>.
- Odenwald, M.A., Turner, J.R., 2017. The intestinal epithelial barrier: a therapeutic target. *Nat. Rev. Gastroenterol. Hepatol.* 14, 1–21. <https://doi.org/10.1038/nrgastro.2016.169>.
- Ohgi, Y., Futamura, T., Kikuchi, T., Hashimoto, K., 2013. Effects of antidepressants on alterations in serum cytokines and depressive-like behavior in mice after lipopolysaccharide administration. *Pharm. Biochem. Behav.* 103, 853–859. <https://doi.org/10.1016/j.pbb.2012.12.003>.
- Osimo, E.F., Baxter, L.J., Lewis, G., Jones, P.B., Khandaker, G.M., 2019. Prevalence of low-grade inflammation in depression: a systematic review and meta-analysis of CRP levels. *Psychol. Med* 49, 1958–1970. <https://doi.org/10.1017/S0033291719001454>.
- Pace, T.W.W., Miller, A.H., 2009. Cytokines and glucocorticoid receptor signaling: relevance to major depression. *Ann. N. Y. Acad. Sci.* 1179, 86–105. <https://doi.org/10.1111/j.1749-6632.2009.04984.x>.
- Pace, T.W.W., Hu, F., Miller, A.H., 2007. Cytokine-effects on glucocorticoid receptor function: relevance to glucocorticoid resistance and the pathophysiology and treatment of major depression. *Brain Behav. Immun.* 21, 9–19. <https://doi.org/10.1016/j.bbi.2006.08.009>.
- Pariante, C.M., Miller, A.H., 2001. Glucocorticoid receptors in major depression: relevance to pathophysiology and treatment. *Biol. Psychiatry* 49, 391–404. [https://doi.org/10.1016/S0006-3223\(00\)01088-X](https://doi.org/10.1016/S0006-3223(00)01088-X).
- Park, C., Lee, S., Cho, I.H., Lee, H.K., Kim, D., Choi, S.Y., Oh, S.B., Park, K., Kim, J.S., Lee, S.J., 2006. TLR3-mediated signal induces proinflammatory cytokine and chemokine gene expression in astrocytes: differential signaling mechanisms of TLR3-induced IP-10 and IL-8 gene expression. *Glia* 53, 248–256. <https://doi.org/10.1002/glia.20278>.
- Pashaki, M.S., Mezel, J.A., Mokhtari, Z., Gheshlagh, R.G., Hesabi, P.S., Nematifard, T., Khaki, S., 2019. The prevalence of comorbid depression in patients with diabetes: a meta-analysis of observational studies. *Diabetes Metab. Syndr.* 13, 3113–3119. <https://doi.org/10.1016/j.dsx.2019.11.003>.
- Pereira-Miranda, E., Costa, P.R.F., Queiroz, V.A.O., Pereira-Santos, M., Santana, M.L.P., 2017. Overweight and obesity associated with higher depression prevalence in adults: a systematic review and meta-analysis. *J. Am. Coll. Nutr.* 36, 223–233. <https://doi.org/10.1080/07315724.2016.1261053>.
- Persoons, P., Vermeire, S., Demyttenaere, K., Fischler, B., Vandenbergh, J., Van Oudenhove, L., Pierik, M., Hlavaty, T., Van Assche, G., Noman, M., Rutgeerts, P., 2005. The impact of major depressive disorder on the short- and long-term outcome of Crohn's disease treatment with infliximab. *Aliment. Pharm. Ther.* 22, 101–110. <https://doi.org/10.1111/j.1365-2036.2005.02535.x>.
- Pittenger, C., Duman, R.S., 2008. Stress, depression, and neuroplasticity: a convergence of mechanisms. *Neuropsychopharmacology* 33, 88–109. <https://doi.org/10.1038/sj.npp.1301574>.
- Pollak, D.D., 2014. Long-term effects of maternal immune activation on depression-like behavior in the mouse. *Transl. Psychiatry* 4, e363. <https://doi.org/10.1038/tp.2013.132>.
- Powell, N., Walker, M.M., Talley, N.J., 2017. The mucosal immune system: master regulator of bidirectional gut-brain communications. *Nat. Rev. Gastroenterol. Hepatol.* 14, 143–159. <https://doi.org/10.1038/nrgastro.2016.191>.
- Prasad, A.S., 2008. Clinical, immunological, anti-inflammatory and antioxidant roles of zinc. *Exp. Gerontol.* 43, 370–377. <https://doi.org/10.1016/j.exger.2007.10.013>.
- Prasad, A.S., 2014. Zinc is an antioxidant and anti-inflammatory agent: its role in human health. *Front Nutr.* 1, 14. <https://doi.org/10.3389/fnut.2014.00014>.
- Prinz, M., Jung, S., Priller, J., 2019. Microglia biology: one century of evolving concepts. *Cell* 179, 292–311. <https://doi.org/10.1016/j.cell.2019.08.053>.
- Pu, J., Liu, Y., Zhang, H., Tian, L., Gui, S., Yu, Y., Chen, X., Chen, Y., Yang, L., Ran, Y., Zhong, X., Xu, S., Song, X., Liu, L., Zheng, P., Wang, H., Xie, P., 2021. An integrated meta-analysis of peripheral blood metabolites and biological functions in major depressive disorder. *Mol. Psychiatry* 26 (8), 4265–4276. <https://doi.org/10.1038/s41380-020-0645-4>.
- Raddatz, D., Toth, S., Schwörer, H., Ramadori, G., 2001. Glucocorticoid receptor signaling in the intestinal epithelial cell lines IEC-6 and Caco-2: evidence of inhibition by interleukin-1 $\beta$ . *Int. J. Colorectal Dis.* 16, 377–383. <https://doi.org/10.1007/s003840100331>.
- Raison, C.L., Capuron, L., Miller, A.H., 2006. Cytokines sing the blues: inflammation and the pathogenesis of depression. *Trends Immunol.* 27, 24–31. <https://doi.org/10.1016/j.it.2005.11.006>.
- Raison, C.L., Borisov, A.S., Majer, M., Drake, D.F., Pagnoni, G., Woolwine, B.J., Vogt, G. J., Massung, B., Miller, A.H., 2009. Activation of central nervous system inflammatory pathways by interferon-alpha: relationship to monoamines and depression. *Biol. Psychiatry* 65, 296–303. <https://doi.org/10.1016/j.biopsych.2008.08.010>.
- Raison, C.L., Dantzer, R., Kelley, K.W., Lawson, M.A., Woolwine, B.J., Vogt, G., Spivey, J. R., Saito, K., Miller, A.H., 2010. CSF concentrations of brain tryptophan and kynurenes during immune stimulation with IFN- $\alpha$ : relationship to CNS immune responses and depression. *Mol. Psychiatry* 15, 393–403. <https://doi.org/10.1038/mp.2009.116>.
- Rajkowska, G., Miguel-Hidalgo, J.J., Wei, J., Dilley, G., Pittman, S.D., Meltzer, H.Y., Overholser, J.C., Roth, B.L., Stockmeier, C.A., 1999. Morphometric evidence for neuronal and glial prefrontal cell pathology in major depression. *Biol. Psychiatry* 45, 1085–1098. [https://doi.org/10.1016/S0006-3223\(99\)00041-4](https://doi.org/10.1016/S0006-3223(99)00041-4).
- Rapaport, M.H., Nierenberg, A.A., Schettler, P.J., Kinkead, B., Caroods, A., Walker, R., Mischoulon, D., 2016. Inflammation as a predictive biomarker for response to omega-3 fatty acids in major depressive disorder: a proof-of-concept study. *Mol. Psychiatry* 21, 71–79. <https://doi.org/10.1038/mp.2015.22>.
- Rehman, S.U., Ali, T., Alam, S.I., Ullah, R., Zeb, A., Lee, K.W., Rutten, B.P.F., Kim, M.O., 2019. Ferulic acid rescues LPS-induced neurotoxicity via modulation of the tlr4 receptor in the mouse hippocampus. *Mol. Neurobiol.* 56, 2774–2790. <https://doi.org/10.1007/s12035-018-1280-9>.
- Reis, D.J., Casteen, E.J., Iardi, S.S., 2019. The antidepressant impact of minocycline in rodents: a systematic review and meta-analysis. *Sci. Rep.* 9, 261. <https://doi.org/10.1038/s41598-018-36507-9>.
- Rios, C., Santamaria, A., 1991. Quinolinic acid is a potent lipid peroxidant in rat brain homogenates. *Neurochem Res* 16, 1139–1143. <https://doi.org/10.1007/BF00966592>.
- Rosenblat, J.D., McIntyre, R.S., 2018. Efficacy and tolerability of minocycline for depression: a systematic review and meta-analysis of clinical trials. *J. Affect Disord.* 227, 219–225. <https://doi.org/10.1016/j.jad.2017.10.042>.
- Rush, A.J., Trivedi, M.H., Wisniewski, S.R., Nierenberg, A.A., Stewart, J.W., Warden, D., Nederehe, G., Thase, M.E., Lavori, P.W., Lebowitz, B.D., McGrath, P.J., Rosenbaum, J.F., Sackeim, H.A., Kupfer, D.J., Luther, J., Fava, M., 2006. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR\*D report. *Am. J. Psychiatry* 163, 1905–1917. <https://doi.org/10.1176/AJP.2006.163.11.1905>.
- Sacchet, M.D., Gotlib, I.H., 2017. Myelination of the brain in major depressive disorder: an in vivo quantitative magnetic resonance imaging study. *Sci. Rep.* 7, 2200. <https://doi.org/10.1038/s41598-017-02062-y>.
- Sahebkar, A., Rathouska, J., Derosa, G., Maffioli, P., Nachtigal, P., 2016. Statin impact on disease activity and C-reactive protein concentrations in systemic lupus erythematosus patients: a systematic review and meta-analysis of controlled trials. *Autoimmun. Rev.* 15, 344–353. <https://doi.org/10.1016/j.autrev.2015.12.007>.
- Salagre, E., Fernandes, B.S., Dodd, S., Brownstein, D.J., Berk, M., 2016. Statins for the treatment of depression: a meta-analysis of randomized, double-blind, placebo-controlled trials. *J. Affect Disord.* 200, 235–242. <https://doi.org/10.1016/j.jad.2016.04.047>.
- Sanada, K., Nakajima, S., Kurokawa, S., Barceló-Soler, A., Ikuse, D., Hirata, A., Yoshizawa, A., Tomizawa, Y., Salas-Valero, M., Noda, Y., Mimura, M., Iwanami, A., Kishimoto, T., 2020. Gut microbiota and major depressive disorder: a systematic review and meta-analysis. *J. Affect Disord.* 266, 1–13. <https://doi.org/10.1016/j.jad.2020.01.102>.
- Sarris, J., Ravindran, A., Yatham, L.N., Marx, W., Rucklidge, J.J., McIntyre, R.S., Akhondzadeh, S., Benedetti, F., Caneo, C., Cramer, H., Cribb, L., de Manincor, M., Dean, O., Deslandes, A.C., Freeman, M.P., Gangadhar, B., Harvey, B.H., Kasper, S., Lake, J., Lopresti, A., Lu, L., Metri, N.J., Mischoulon, D., Ng, C.H., Nishi, D., Rahimi, R., Seedat, S., Sinclair, J., Su, K.P., Zhang, Z.J., Berk, M., 2022. Clinician guidelines for the treatment of psychiatric disorders with nutraceuticals and phytochemicals: the world federation of societies of biological psychiatry (WFSBP) and Canadian network for mood and anxiety treatments (CANMAT) taskforce. *World J. Biol. Psychiatry* 23 (6), 424–455. <https://doi.org/10.1080/15622975.2021.2013041>.
- Sassi, F., Tamone, C., D'Amelio, P., 2018. Vitamin D: nutrient, hormone, and immunomodulator. *Nutrients* 10, 1656. <https://doi.org/10.3390/nu10111656>.
- Savage, C.D., Lopez-Castejon, G., Denes, A., Brough, D., 2012. NLRP3-inflammasome activating DAMPs stimulate an inflammatory response in glia in the absence of priming which contributes to brain inflammation after injury. *Front Immunol.* 3, 288. <https://doi.org/10.3389/fimmu.2012.00288>.
- Schwarz, R., Pellicciari, R., 2002. Manipulation of brain kynurenes: glial targets, neuronal effects, and clinical opportunities. *J. Pharm. Exp. Ther.* 303, 1–10. <https://doi.org/10.1124/jpet.102.034439>.
- Shafiee, A., Arekhi, S., Omranzadeh, A., Sahebkar, A., 2018. Saffron in the treatment of depression, anxiety and other mental disorders: Current evidence and potential mechanisms of action. *J. Affect Disord.* 227, 330–337. <https://doi.org/10.1016/j.jad.2017.11.020>.
- Shi, Z., Ren, H., Huang, Z., Peng, Y., He, B., Yao, X., Yuan, T.F., Su, H., 2017. Fish oil prevents lipopolysaccharide-induced depressive-like behavior by inhibiting neuroinflammation. *Mol. Neurobiol.* 54, 7327–7334. <https://doi.org/10.1007/s12035-016-0212-9>.
- Si, X., Miguel-Hidalgo, J.J., O'Dwyer, G., Stockmeier, C.A., Rajkowska, G., 2004. Age-dependent reductions in the level of glial fibrillary acidic protein in the prefrontal cortex in major depression. *Neuropsychopharmacology* 29, 2088–2096. <https://doi.org/10.1038/sj.npp.1300525>.
- da Silva, L.E.M., de Santana, M.L.P., Costa, P.R., de F., Pereira, E.M., Nepomuceno, C.M. M., Queiroz, V.A., de O., de Oliveira, L.P.M., Machado, M.E.P. da C., de Sena, E.P., 2021. Zinc supplementation combined with antidepressant drugs for treatment of patients with depression: a systematic review and meta-analysis. *Nutr. Rev.* 79, 1–12. <https://doi.org/10.1093/NUTRIT/NUA0039>.
- Simon, M.S., Arteaga-Henríquez, G., Fouad, Algendy, A., Siepmann, T., Illigens, B.M.W., 2023. Anti-inflammatory treatment efficacy in major depressive disorder: a



- systematic review of meta-analyses. *Neuropsychiatr. Dis. Treat.* 19, 1–25. <https://doi.org/10.2147/NDT.S385117>.
- Simopoulos, A.P., 2002. Omega-3 fatty acids in inflammation and autoimmune diseases. *J. Am. Coll. Nutr.* 21, 495–505. <https://doi.org/10.1080/07315724.2002.10719248>.
- Simpson, D.S.A., Oliver, P.L., 2020. Ros generation in microglia: understanding oxidative stress and inflammation in neurodegenerative disease. *Antioxidants* 9, 743. <https://doi.org/10.3390/antiox9080743>.
- Smith, K.J., Au, B., Ollis, L., Schmitz, N., 2018. The association between C-reactive protein, Interleukin-6 and depression among older adults in the community: a systematic review and meta-analysis. *Exp. Gerontol.* 102, 109–132. <https://doi.org/10.1016/j.exger.2017.12.005>.
- Sorensen, A.L., Hasselbalch, H.C., Nielsen, C.H., Poulsen, H.E., Ellervik, C., 2019. Statin treatment, oxidative stress and inflammation in a Danish population. *Redox Biol.* 21, 101088. <https://doi.org/10.1016/j.redox.2018.101088>.
- Stahl, S.M., 1998. Mechanism of action of serotonin selective reuptake inhibitors. *Serotonin Recept. Pathw. Mediat. Ther. Eff. Side-. Eff. J. Affect Disord.* 51, 215–235. [https://doi.org/10.1016/S0165-0327\(98\)00221-3](https://doi.org/10.1016/S0165-0327(98)00221-3).
- Steiner, J., Bielau, H., Brisch, R., Danos, P., Ullrich, O., Mawrin, C., Bernstein, H.G., Bogerts, B., 2008. Immunological aspects in the neurobiology of suicide: elevated microglial density in schizophrenia and depression is associated with suicide. *J. Psychiatr. Res.* 42, 151–157. <https://doi.org/10.1016/j.jpsychires.2006.10.013>.
- Stilling, R.M., Dinan, T.G., Cryan, J.F., 2014. Microbial genes, brain & behaviour - epigenetic regulation of the gut-brain axis. *Genes Brain Behav.* 13, 69–86. <https://doi.org/10.1111/gbb.12109>.
- Stratoulis, V., Venero, J.L., Tremblay, M., Joseph, B., 2019. Microglial subtypes: diversity within the microglial community. *EMBO J.* 38, e101997. <https://doi.org/10.15252/emboj.2019191097>.
- Strawbridge, R., Arnone, D., Danese, A., Papadopoulos, A., Herane Vives, A., Cleare, A.J., 2015. Inflammation and clinical response to treatment in depression: a meta-analysis. *Eur. Neuropsychopharmacol.* 25, 1532–1543. <https://doi.org/10.1016/j.euroneuro.2015.06.007>.
- Swardfager, W., Herrmann, N., Mazereeuw, G., Goldberger, K., Harimoto, T., Lancôt, K. L., 2013. Zinc in depression: a meta-analysis. *Biol. Psychiatry* 74, 872–878. <https://doi.org/10.1016/j.biopsych.2013.05.008>.
- Tan, Y.-L., Yuan, Y., Tian, L., 2020. Microglial regional heterogeneity and its role in the brain. *Mol. Psychiatry* 25, 351–367. <https://doi.org/10.1038/s41380-019-0609-8>.
- Taniguchi, K., Karin, M., 2018. NF- $\kappa$ B, inflammation, immunity and cancer: coming of age. *Nat. Rev. Immunol.* 18, 309–324. <https://doi.org/10.1038/nri.2017.142>.
- Taniguti, E.H., Ferreira, Y.S., Stupp, I.J.V., Fraga-Junior, E.B., Doneda, D.L., Lopes, L., Rios-Santos, F., Lima, E., Buss, Z.S., Viola, G.G., Vandresen-Filho, S., 2019. Atorvastatin prevents lipopolysaccharide-induced depressive-like behaviour in mice. *Brain Res Bull.* 146, 279–286. <https://doi.org/10.1016/j.brainresbull.2019.01.018>.
- Tolkien, K., Bradburn, S., Murgatroyd, C., 2019. An anti-inflammatory diet as a potential intervention for depressive disorders: a systematic review and meta-analysis. *Clin. Nutr.* 38, 2045–2052. <https://doi.org/10.1016/j.clnu.2018.11.007>.
- Tóth, B., Hegyi, P., Lantos, T., Szakács, Z., Kerémi, B., Varga, G., Tenk, J., Pétervári, E., Balaskó, M., Rumbus, Z., Rakonczay, Z., Bálint, E.R., Kiss, T., Csopor, D., 2019. The efficacy of saffron in the treatment of mild to moderate depression: a meta-analysis. *Planta Med* 85, 24–31. <https://doi.org/10.1055/a-0660-9565>.
- Tsao, C.W., Lin, Y.S., Chen, C.C., Bai, C.H., Wu, S.R., 2006. Cytokines and serotonin transporter in patients with major depression. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 30, 899–905. <https://doi.org/10.1016/j.pnpb.2006.01.029>.
- Turner, J.R., 2009. Intestinal mucosal barrier function in health and disease. *Nat. Rev. Immunol.* 9, 799–809. <https://doi.org/10.1038/nri2653>.
- Tynan, R.J., Weidenhofer, J., Hinwood, M., Cairns, M.J., Day, T.A., Walker, F.R., 2012. A comparative examination of the anti-inflammatory effects of SSRI and SNRI antidepressants on LPS stimulated microglia. *Brain Behav. Immun.* 26, 469–479. <https://doi.org/10.1016/j.bbi.2011.12.011>.
- Udina, M., Hidalgo, D., Navinés, R., Fornis, X., Solà, R., Farré, M., Capuron, L., Vieta, E., Martín-Santos, R., 2014. Prophylactic antidepressant treatment of interferon-induced depression in chronic hepatitis C: a systematic review and meta-analysis. *J. Clin. Psychiatry* 75, e1113–e1121. <https://doi.org/10.4088/JCP.13r08800>.
- Uranova, N.A., Vostrikov, V.M., Orlovskaya, D.D., Rachmanova, V.I., 2004. Oligodendroglial density in the prefrontal cortex in schizophrenia and mood disorders: a study from the stanley neuropathology consortium. *Schizophr. Res* 67, 269–275. [https://doi.org/10.1016/S0920-9964\(03\)00181-6](https://doi.org/10.1016/S0920-9964(03)00181-6).
- Valero, J., Mastrella, G., Neiva, I., Sánchez, S., Malva, J.O., 2014. Long-term effects of an acute and systemic administration of LPS on adult neurogenesis and spatial memory. *Front Neurosci.* 8, 83. <https://doi.org/10.3389/fnins.2014.00083>.
- Vellekkatt, F., Menon, V., 2019. Efficacy of Vitamin D supplementation in major depression: a meta-analysis of randomized controlled trials. *J. Post. Med* 65, 74–80. [https://doi.org/10.4103/jpgm.JPGM\\_571\\_17](https://doi.org/10.4103/jpgm.JPGM_571_17).
- Veronese, N., Stubbs, B., Solmi, M., Ajnakina, O., Carvalho, A.F., Maggi, S., 2018. Acetyl-L-carnitine supplementation and the treatment of depressive symptoms: a systematic review and meta-analysis. *Psychosom. Med* 80, 154–159. <https://doi.org/10.1097/PSY.0000000000000537>.
- Wang, J., Luo, Y., Tang, J., Liang, X., Huang, C., Gao, Y., Qi, Y., Yang, C., Chao, F.L., Zhang, Y., Tang, Y., 2020. The effects of fluoxetine on oligodendrocytes in the hippocampus of chronic unpredictable stress-induced depressed model rats. *J. Comp. Neurol.* 528, 2583–2594. <https://doi.org/10.1002/cne.24914>.
- Wang, L., Wang, R., Liu, L., Qiao, D., Baldwin, D.S., Hou, R., 2019. Effects of SSRIs on peripheral inflammatory markers in patients with major depressive disorder: A systematic review and meta-analysis. *Brain Behav. Immun.* 79, 24–38. <https://doi.org/10.1016/j.bbi.2019.02.021>.
- Wellen, K.E., Hotamisligil, G.S., 2005. Inflammation, stress, and diabetes. *J. Clin. Invest.* 115, 1111–1119. <https://doi.org/10.1172/jci25102>.
- WHO, 2020. Depression [WWW Document]. URL [https://www.who.int/health-topics/depression#tab=tab\\_1](https://www.who.int/health-topics/depression#tab=tab_1) (accessed 9.27.20).
- Więdołcha, M., Marciniowicz, P., Krupa, R., Janoska-Jazdzik, M., Janus, M., Dębowska, W., Mosiolek, A., Waszkiewicz, N., Szulc, A., 2018. Effect of antidepressant treatment on peripheral inflammation markers – a meta-analysis. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 80, 217–226. <https://doi.org/10.1016/j.pnpb.2017.04.026>.
- de Wit, L., Luppino, F., van Straten, A., Penninx, B., Zitman, F., Cuijpers, P., 2010. Depression and obesity: a meta-analysis of community-based studies. *Psychiatry Res* 178, 230–235. <https://doi.org/10.1016/j.psychres.2009.04.015>.
- Wittenberg, G.M., Stylianou, A., Zhang, Y., Sun, Y., Gupta, A., Jagannatha, P.S., Wang, D., Hsu, B., Curran, M.E., Khan, S., ImmunoPsychiatry Consortium, M.R.C., Chen, G., Bullmore, E.T., Drevets, W.C., 2020. Effects of immunomodulatory drugs on depressive symptoms: a mega-analysis of randomized, placebo-controlled clinical trials in inflammatory disorders. *Mol. Psychiatry* 25 (6), 1275–1285. <https://doi.org/10.1038/s41380-019-0471-8>.
- Wu, M.D., Hein, A.M., Moravan, H.J., Shaftel, S.S., Olschowka, J.A., O'Banion, M.K., 2012. Adult murine hippocampal neurogenesis is inhibited by sustained IL-1 $\beta$  and not rescued by voluntary running. *Brain Behav. Immun.* 26, 292–300. <https://doi.org/10.1016/j.bbi.2011.09.012>.
- Wu, S., Guo, T., Qi, W., Li, Y., Gu, J., Liu, C., Sha, Y., Yang, B., Hu, S., Zong, X., 2021. Curcumin ameliorates ischemic stroke injury in rats by protecting the integrity of the blood brain barrier. *Exp. Ther. Med* 22, 783. <https://doi.org/10.3892/etm.2021.10215>.
- Xu, Y., Hackett, M., Carter, G., Loo, C., Gálvez, V., Glozier, N., Glue, P., Lapidus, K., McGirr, A., Somogyi, A.A., Mitchell, P.B., Rodgers, A., 2016. Effects of low-dose and very low-dose ketamine among patients with major depression: a systematic review and meta-analysis. *Int J. Neuropsychopharmacol.* 19, pyv124. <https://doi.org/10.1093/ijnp/pyv124>.
- Yang, C., Shen, J., Hong, T., Hu, T.T., Li, Z.J., Zhang, H.T., Zhang, Y.J., Zhou, Z.Q., Yang, J.J., 2013. Effects of ketamine on lipopolysaccharide-induced depressive-like behavior and the expression of inflammatory cytokines in the rat prefrontal cortex. *Mol. Med Rep.* 8, 887–890. <https://doi.org/10.3892/mmr.2013.1600>.
- Yang, C.C., Hsiao, L., Der, Tseng, H.C., Kuo, C.M., Yang, C.M., 2020. Pristimerin inhibits MMP-9 expression and cell migration through attenuating NOX/ROS-dependent NF- $\kappa$ B activation in rat brain astrocytes challenged with LPS. *J. Inflamm. Res* 13, 325–341. <https://doi.org/10.2147/JIR.S252659>.
- Yang, X., Chen, X., Fu, Y., Luo, Q., Du, L., Qiu, H., Qiu, T., Zhang, L., Meng, H., 2018. Comparative efficacy and safety of Crocus sativus L. For treating mild to moderate major depressive disorder in adults: a meta-analysis of randomized controlled trials. *Neuropsychiatr. Dis. Treat.* 14, 1297–1305. <https://doi.org/10.2147/NDT.S157550>.
- Yong, V.W., Wells, J., Giuliani, F., Casha, S., Power, C., Metz, L.M., 2004. The promise of minocycline in neurology. *Lancet Neurol.* 3, 744–751. [https://doi.org/10.1016/S1474-4422\(04\)00937-8](https://doi.org/10.1016/S1474-4422(04)00937-8).
- Yu, X., Zhang, F., Shi, J., 2019. Effect of sevoflurane treatment on microglia activation, NF- $\kappa$ B and MAPK activities. *Immunobiology* 224, 638–644. <https://doi.org/10.1016/j.imbio.2019.07.004>.
- Zanos, P., Gould, T.D., 2018. Mechanisms of ketamine action as an antidepressant. *Mol. Psychiatry* 23, 801–811. <https://doi.org/10.1038/mp.2017.255>.
- Zeinali, M., Zarak, M.R., Rezaee, S.A., Karimi, G., Hosseinzadeh, H., 2019. Immunoregulatory and anti-inflammatory properties of Crocus sativus (Saffron) and its main active constituents: a review. *Iran. J. Basic Med Sci.* 22, 334–344. <https://doi.org/10.22038/ijbms.2019.34365.8158>.
- Zhang, J.C., Wu, J., Fujita, Y., Yao, W., Ren, Q., Yang, C., Li, S.X., Shirayama, Y., Hashimoto, K., 2014. Antidepressant effects of TrkB ligands on depression-like behavior and dendritic changes in mice after inflammation. *Int J. Neuropsychopharmacol.* 18 pyu077. <https://doi.org/10.1093/ijnp/pyu077>.
- Zhang, L., Previn, R., Lu, L., Liao, R.F., Jin, Y., Wang, R.K., 2018. Crocin, a natural product attenuates lipopolysaccharide-induced anxiety and depressive-like behaviors through suppressing NF- $\kappa$ B and NLRP3 signaling pathway. *Brain Res Bull.* 142, 352–359. <https://doi.org/10.1016/j.brainresbull.2018.08.021>.
- Zhang, Q., Sun, Y., He, Z., Xu, Y., Li, X., Ding, J., Lu, M., Hu, G., 2020. Kynurenine regulates NLRP2 inflammasome in astrocytes and its implications in depression. *Brain Behav. Immun.* 88, 471–481. <https://doi.org/10.1016/j.bbi.2020.04.016>.
- Zhang, W., Guo, Y., Han, W., Yang, M., Wen, L., Wang, K., Jiang, P., 2019. Curcumin relieves depressive-like behaviors via inhibition of the NLRP3 inflammasome and kynurenine pathway in rats suffering from chronic unpredictable mild stress. *Int Immunopharmacol.* 67, 138–144. <https://doi.org/10.1016/j.intimp.2018.12.012>.
- Zhao, Y., Yang, G., Zhao, Z., Wang, C., Duan, C., Gao, L., Li, S., 2020. Antidepressant-like effects of Lactobacillus plantarum DP189 in a corticosterone-induced rat model of chronic stress. *Behav. Brain Res* 395, 112853. <https://doi.org/10.1016/j.bbr.2020.112853>.
- Zheng, L.S., Hitoshi, S., Kaneko, N., Takao, K., Miyakawa, T., Tanaka, Y., Xia, H., Kalinke, U., Kudo, K., Kanba, S., Ikenaka, K., Sawamoto, K., 2014. Mechanisms for interferon- $\alpha$ -induced depression and neural stem cell dysfunction. *Stem Cell Rep.* 3, 73–84. <https://doi.org/10.1016/j.stemcr.2014.05.015>.
- Zheng, L.S., Kaneko, N., Sawamoto, K., 2015. Minocycline treatment ameliorates interferon-alpha-induced neurogenic defects and depression-like behaviors in mice. *Front Cell Neurosci.* 9, 5. <https://doi.org/10.3389/fncel.2015.00005>.
- Zhou, F., Yao, H.H., Wu, J.Y., Yang, Y.J., Ding, J.H., Zhang, J., Hu, G., 2006. Activation of group II/III metabotropic glutamate receptors attenuates LPS-induced astroglial



- neurotoxicity via promoting glutamate uptake. *J. Neurosci. Res* 84, 268–277. <https://doi.org/10.1002/jnr.20897>.
- Zhu, C.Bin, Blakely, R.D., Hewlett, W.A., 2006. The proinflammatory cytokines interleukin-1beta and tumor necrosis factor-alpha activate serotonin transporters. *Neuropsychopharmacology* 31, 2121–2131. <https://doi.org/10.1038/sj.npp.1301029>.
- Zhu, J.H., Bo, H.H., Liu, B.P., Jia, C.X., 2023. The associations between DNA methylation and depression: a systematic review and meta-analysis. *J. Affect Disord.* 327, 439–450. <https://doi.org/10.1016/j.jad.2023.01.079>.

Review

# The neuronal cilium – a highly diverse and dynamic organelle involved in sensory detection and neuromodulation

Nathalie Jurisch-Yaksi <sup>1,\*</sup> Dagmar Wachten,<sup>2</sup> and Jay Gopalakrishnan<sup>3,4</sup>

**Cilia are fascinating organelles that act as cellular antennae, sensing the cellular environment. Cilia gained significant attention in the late 1990s after their dysfunction was linked to genetic diseases known as ciliopathies. Since then, several breakthrough discoveries have uncovered the mechanisms underlying cilia biogenesis and function. Like most cells in the animal kingdom, neurons also harbor cilia, which are enriched in neuromodulatory receptors. Yet, how neuronal cilia modulate neuronal physiology and animal behavior remains poorly understood. By comparing ciliary biology between the sensory and central nervous systems (CNS), we provide new perspectives on the functions of cilia in brain physiology.**

## Primary cilia as cellular antennae

Cilia are conserved antenna-like organelles extending from the cell surface and consist of a microtubule core, the axoneme, covered by the plasma membrane [1,2]. While motile cilia are involved in cell motility and fluid pumping [3,4], primary cilia play sensory and structural roles and are located on most vertebrate cells, including neurons [1,5]. Cilia do not have local mRNA translation mechanisms, and thus depend on an active intraflagellar transport machinery for protein import and export [6]. A diffusion barrier at the ciliary base, called the transition zone, regulates protein trafficking to the cilium [2,6,7]. Through these mechanisms, the molecular composition of cilia differs from the rest of the cytoplasm.

The primary cilia (hereafter referred to as cilia) have uniquely evolved to sense and transmit extracellular signals. They are enriched with receptors and signaling cascades, and transduce signals distinct from the cell soma [1,2]. The second messenger cAMP and its downstream signaling is a prime example of this compartmentalization [8,9].

Cilia play critical roles in brain development. They regulate the self-renewing and differentiating capacities of neural progenitors [10,11], and the migration and synapse formation of newly generated neurons. This is achieved through control of both cell cycle progression and signaling, for example, Hedgehog signaling (Boxes 1 and 2), which have been comprehensively reviewed elsewhere [1,2,12–15]. This review focuses on recent findings on cilia in neurons of the mature nervous system. We first discuss ciliary function in sensory neurons, including chemosensing in olfactory sensory neurons, photosensing in photoreceptors, and mechanosensing in auditory cells, before moving to the neurons of the CNS. One of our goals is to draw parallels between sensory and CNS neurons to deepen our understanding of the functions of cilia in brain physiology.

## Cilia are the core of most sensory systems

Cilia are involved in many sensory systems, with perhaps the best example being the olfactory system. In olfactory sensory neurons (Figure 1A), cilia extend into the olfactory cavity, where

## Highlights

Cilia are critical for detecting chemo-, photo-, and mechanosensory stimuli in the sensory systems.

Most neurons in the central nervous system harbor cilia enriched in neuromodulatory receptors.

Neuronal cilia are structurally and molecularly diverse, and play various roles in the central nervous system. The functionality of cilia is determined by their structure, the presence of receptors and ion channels, and their specific context.

Neuronal cilia are found adjacent to chemical synapses and axonal segments, and can form axociliary synapses.

<sup>1</sup>Department of Clinical and Molecular Medicine (IKOM), Faculty of Medicine and Health Science, Norwegian University of Science and Technology, Erling Skalgssons gate 1, 7491 Trondheim, Norway

<sup>2</sup>Department of Biophysical Imaging, Institute of Innate Immunity, Medical Faculty, University of Bonn, Venusberg-Campus 1, 53127 Bonn, Germany

<sup>3</sup>Institute of Human Genetics, University Hospital, Heinrich-Heine-Universität, 40225 Düsseldorf, Germany

<sup>4</sup>Institute for Human Genetics, Universitätsklinikum Jena, Friedrich-Schiller-Universität Jena, 07740 Jena, Germany

\*Correspondence: [nathalie.jurisch-yaksi@ntnu.no](mailto:nathalie.jurisch-yaksi@ntnu.no) (N. Jurisch-Yaksi).



### Box 1. Primary ciliary dynamics in neuronal progenitors

Since the presence of cilia is incompatible with mitosis, primary cilia in cycling neuronal precursor cells are highly dynamic in terms of cilium assembly and disassembly. Progenitor cells assemble cilia when exiting mitosis and disassemble them at cell cycle re-entry, that is, disassembly starts at the G1–S transition and is completed at the G2–M transition [1]. Cells utilize different ciliary disassembly pathways, which are discussed in detail in other reviews [1,121]. This process is tightly controlled, as delayed ciliary disassembly hinders cells from mitotic entry. This is particularly important in neuronal progenitors because their abundance and diversity expand exponentially during development to generate brains with various neuronal populations [122,123]. Accordingly, defects in the temporal kinetics of cilium disassembly disrupt progenitor cell diversity and severely impact brain development [10,11].

Since differentiated neurons do not re-enter cell cycle under physiological conditions, they usually do not fully disassemble their cilia. An exception to this rule are cerebellar granule cells, which undergo ciliary disassembly through mechanisms distinct from the ones in progenitors [73]. In this context, future work should investigate the precise structural and molecular organization at the site where cilia emerge from the cell surface, such as the ciliary pocket and pit, since these have often been seen in dividing cells and are indicative of vesicular transport. Recent analyses of large electron microscopy datasets have shown that ciliary pockets are more frequently observed in proliferating progenitors than differentiated neurons [49,50], supporting a faster remodeling of cilia in progenitors than in post-mitotic neurons.

they sense chemical cues [16,17]. Olfactory cilia contain all the necessary machinery for chemosensation, including the odorant G protein-coupled receptors (GPCRs) and the channels triggering membrane depolarization [18]. Ciliary function in the olfactory system is widely conserved from nematodes to mammals and necessary for olfaction [17,19,20].

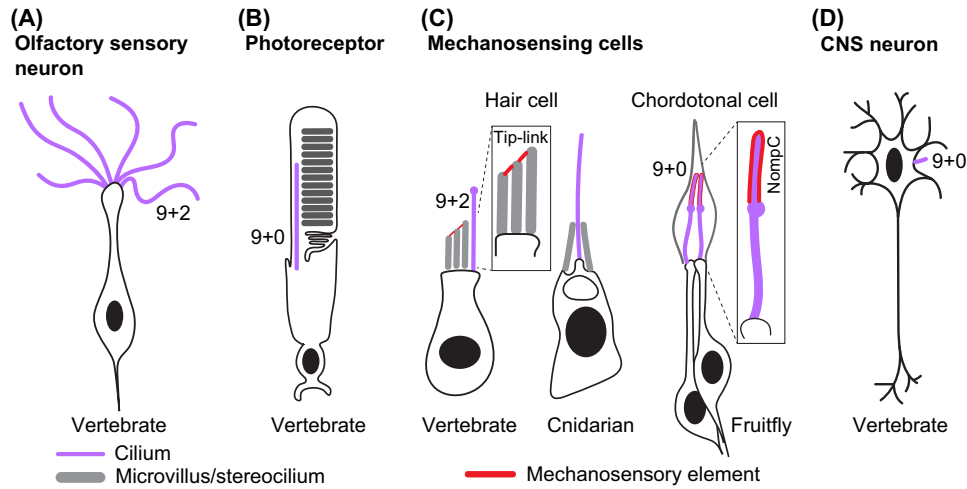
In the retina, photoreceptors harbor a modified connecting cilium (Figure 1B) that connects the cell soma to the photosensory outer segments (OSs). The OSs contain membrane rafts enriched with rhodopsin, which, together with retinal cofactors, form a photopigment that initiates the phototransduction cascade [21]. Since the connecting cilia of the photoreceptors are involved in transporting molecules from the cell soma to the OS, ciliary dysfunction leads to defects in protein trafficking and OS structural integrity, ultimately causing photoreceptor degeneration and blindness [21,22].

In the vertebrate inner ear, hair cell cilia are not directly involved in mechanosensing. Instead, cilia are essential for assembling the microvilli-based stereocilia, which detect mechanical vibrations through the tip links [23]. Cilia are commonly involved in mechanosensing (Figure 1C) in various tissues and species [24] (e.g., the chordotonal organs in *Drosophila* [25], or in vertebrate hair cells during development when the stereocilia are not fully mature [26]).

### Box 2. Primary cilia-dependent Hedgehog signaling during brain development

Primary cilia serve as signaling hubs during neuronal development by transducing various signaling pathways, including Hedgehog, Notch, Wnt, GPCR-mediated, and PDGFRA-mediated signaling [1,2,124]. Among these, canonical Hedgehog signaling is critical at the early stage of brain development in vertebrate neural tube patterning [1,125]. Hedgehog ligands bind their ciliary receptor Patched 1 (PTCH1), which initiates the signaling cascade and, in turn, regulates gene expression [124,125]. Notably, Hedgehog signaling components dynamically localize in and out of the cilium, depending on whether the signaling pathway is turned ON or OFF. In the OFF state, PTCH1 is located at the ciliary membrane, inhibiting the atypical GPCR Smoothened (SMO). Concomitantly, ciliary SMO levels are maintained at a low level due to its ubiquitination and removal from the cilium. Upon binding of the Hedgehog ligands, PTCH1 is inactivated and exits from the cilium. In turn, SMO accumulates in the cilium. Following its activation, SMO changes its conformation, which blocks the phosphorylation of GLI2/3 by protein kinase A (PKA) and, thereby, activates GLI2/3, which traffic to the nucleus to drive the expression of Hedgehog target genes [1,124,125].

Since ciliary assembly and disassembly are dynamic processes, signaling events, including Hedgehog, should be in tight coherence with the ciliary dynamics. In other words, the relative amount of Hedgehog signal titrated and transduced in neuronal progenitors depends on the ciliary status. Thus, the cilia are dynamic, as well as their associated signaling. Increasing evidence shows that ciliary dynamics also impact Hedgehog signaling in differentiating and fully mature neurons [68,72,112,116].



Trends in Neurosciences

**Figure 1. Ciliary diversity in sensory cells and neurons.** (A) Vertebrate olfactory sensory neurons harbor multiple immotile cilia with a 9+2 ultrastructure. Note that the cilia of olfactory neurons can be long. (B) In vertebrate photoreceptors, the 9+0 connecting cilium connects the membrane-rich outer segment with the cell soma. (C) Mechanosensory cells have different morphologies. Hair cells in vertebrates and cnidarians, which use stereocilia (gray) for mechanosensation, are shown here. Note that the tip link connecting stereocilia (red) is crucial for mechanosensation in vertebrate hair cells. In chordotonal cells of the fruitfly, the mechanosensory channel NompC (red), which detects cilia bending, is expressed in the distal part of the cilium after the ciliary dilatation. Cilia in mechanosensory cells can have a 9+0 (chordotonal cells) or 9+2 (vertebrate hair cells) ultrastructure. (D) Cilia of central nervous system neurons have a 9+0 structure and usually emanate from the cell body. The structure of cnidarian hair cells is modified from [126], and chordotonal cells from [32].

### Ciliary diversity in sensory systems

Cilia in sensory cell types are highly specialized to accomplish their distinct function (Figure 1). For example, olfactory cilia are surprisingly long, with more than 100  $\mu\text{m}$  lengths in mice [27] and frogs [28]. Olfactory cilia also display highly elaborated arborization, as seen in *Caenorhabditis elegans* [19,29], providing a larger membrane surface for enhanced detection of olfactory cues. For cilia-dependent mechanosensing, cilia bending is commonly detected through specialized structures within the cilia or at the apical part of the cell [25]. For instance, microvilli located at the base of the sensory cilium and connected by tip-link-like fibers, similar to hair cells, have been described in mechanosensory cells; for example, in *Platynereis* and cnidarians [24,25,30] (Figure 1C). Alternatively, mechanosensory channels may be directly localized in the cilia. This is the case for chordotonal cells in *Drosophila*, where the mechanosensor NompC is enriched at the distal part of the cilia [31,32] (Figure 1C).

Besides exhibiting various morphologies, cilia in sensory systems may also display an altered axonemal ultrastructure. For instance, olfactory and mechanosensory cilia commonly harbor an extra pair of microtubules at their core, known as the central pair (referred to as 9+2 cilia) [17,30], which is the hallmark of motile cilia [33,34] and not commonly seen in primary cilia (referred to as a 9+0 cilia). However, the cilia of olfactory sensory neurons and inner ear hair cells are usually immotile, except in some amphibians, like *Rana pipiens*, which harbor motile cilia on their olfactory neurons [35]. The central pair is suggested to provide the extra stability needed for their function [36]. In other systems, including photoreceptors, the base of the cilium, including the transition zone, has undergone some specialization. In photoreceptors, the transition zone is exceptionally long, to support its role in protein trafficking from the cell soma to the OS [37].



In the sensory system, ciliary diversity is well described. However, the genetic programs and cellular mechanisms leading to the specialization of ciliary structure are only starting to be uncovered. For instance, ciliated cells can alter ciliogenesis pathways to generate cilia with diverse ultrastructure. This is the case for olfactory sensory neurons and hair cells, which express the master transcriptional regulator of motile ciliogenesis Foxj1, despite not having motile cilia [35,38–40]. These cells repurposed the Foxj1-dependent transcription program to generate cilia with motile cilia-like features [35]. Divergent tubulin isotypes have been associated with ciliary specialization in *C. elegans* [41] and mammals [42]. Given that cilia are highly diverse across species and cell types [24,43–47], unraveling mechanisms underlying ciliary diversity throughout evolution could give new insights into the context-specific functions of cilia.

### Cilia in CNS neurons are diverse and dynamic

Besides the sensory organs, cilia are found in nearly all types of neurons of the vertebrate CNS [48–50]. Cilia were also described in some ganglionic neurons of cnidarians [51–53], in CNS neurons of the chordate *Ciona intestinalis* [54,55], but not in *Drosophila* and *C. elegans* [56,57], suggesting that some species have evolved to maintain neuronal cilia to perform specific physiological functions. Notably, *Drosophila*, in contrast to many other species, including all vertebrates and the last common ancestor of bilaterians, does not require cilia for transducing Hedgehog signaling, a key, cilia-dependent neurodevelopmental signaling pathway [43,46,57].

Cilia are typically marked using antibodies that detect microtubule modifications, such as acetylation or glutamylation [58,59]. However, the axonal microtubules are heavily acetylated in neurons, which is critical for microtubule stability and neuronal processes [58,60]. This makes it challenging to identify cilia in the brain using tubulin-based markers. Instead, alternative markers are used to label neuronal cilia, including the membrane GTPase ARL13B, the cAMP-generating enzyme adenylate cyclase 3 (AC3), and the GPCRs somatostatin receptor 3 (SSTR3), serotonin receptor 6 (HTR6), and melanin-concentrating hormone receptor (MCHR), which are all enriched in the ciliary compartment [61–64].

### Diverse content of neuronal cilia

There is a diversity in the expression of markers among neuronal cilia. While most neuronal cilia express AC3 [61,64,65], HTR6 is preferentially found in projection neurons of the hippocampus and cerebral cortex, in striatal, medium-sized spiny neurons, and in a minor population of interneurons [64]. Similarly, SSTR3 is enriched in cortical excitatory neurons, with rare expression (<10%) in the interneurons [66]. In addition, the expression of cilia markers is not uniform during the life cycle of neurons. While AC3 expression in the CA1 region of the hippocampus remains relatively stable after postnatal day 10 (P10), ARL13B reaches high levels in the stratum pyramidale layer at P10 before fading away at P40 in mice [67].

### Ciliary length, ultrastructure, and orientation are dynamic in space and time

Recent studies aimed to gain further insights on the abundance, structure, length, orientation, and dynamics of cilia in the CNS. Ciliogenesis is tightly connected to the cell cycle and occurs following cell division during neuronal differentiation [1,2] (Box 1). Ciliary length increases during postnatal development in the mouse dorsal cortex over several weeks to reach stabilized lengths at P60 [48]. For fully differentiated neurons, ciliary length can range from only a few micrometers up to 20  $\mu\text{m}$ , depending on the species, brain region, and cell type [49,50,64,68–71]. One exception is cerebellar granule cells, which are devoid of cilia [72,73].

Recent studies using volumetric electron microscopy in the human anterior temporal lobe cortex [49] and mouse primary visual cortex [50] reported that almost all excitatory and inhibitory

neurons (>99.5%) harbor cilia. Ciliary pockets, which are indicative of vesicle transport and often seen in dividing cells, such as retinal pigmented epithelial cells and neural progenitor cells [10,74], were observed in 24–56% of neurons in the human cortex [49], but not in mouse cortex [50]. The ultrastructure of neuronal cilia also varies among cells. Notably, the 9+0 axonemal organization of microtubules was not identified throughout the cilia, and microtubules extended only to 70–86% of the ciliary length in human neurons [49]. Besides, interneuronal cilia were less smooth than excitatory neurons and presented convoluted or beaded ciliary membrane sheets, suggesting increased shedding of extracellular vesicles [49].

The ciliary length of neuronal cilia is actively regulated. For example, cilia are elongated in MCH-neurons when MCHR1 signaling is inhibited. Conversely, cilia are shortened when MCH pathways are activated [75]. Likewise, cilia in the suprachiasmatic nucleus of mice, especially in neuromedin S-producing neurons, are regulated in a circadian matter: ciliary length oscillates from 1 to 6  $\mu\text{m}$ , and ciliary abundance ranges from 10% to 40% during the circadian cycle [68]. Circadian modulation of ciliary length was also observed in the nucleus accumbens, somatosensory cortex, and three hypothalamic nuclei [69]. Finally, metabolic alterations, including leptin treatment, fasting, and diet-induced obesity, modulated cilium length in hypothalamic neurons [76].

Beside signaling, differentiation processes *per se* can alter ciliary length in neurons. This was shown in granule cells in the mouse cerebellum, which lose their cilia during differentiation [73]. The deciliation process in granule cells is different from ciliary disassembly that occurs during cell division of progenitors [73] (Box 1). It involves intermediate steps, where cilia are concealed in a membrane submerged in the cytoplasm, [73] and genes important for ciliary maintenance are downregulated [72]. The mother centriole docks at the plasma membrane of fully differentiated granule cells, but remains unciliated, probably due to the recruitment of the centriolar cap protein CEP97 [72]. This study also reported that ciliary deconstruction in the granule cells correlated with a reduction in Hedgehog signaling [72] (Box 2). As ciliary length directly modulates ciliary signaling, control of ciliary length as well as presence and accessibility to extracellular milieu could directly impact neuronal function [77].

In the same way that ciliary length follows systematic patterns, ciliary orientation in neurons is not stochastic and, at least in some cases, seems to follow specific patterns [49,50,69]. In the cortex, cilia are usually oriented toward the pia in mice [50] and within the plane of the cortical wall in humans [49]. In the hippocampus, cilia are aligned to the radial axis [67]. Orientation of cilia in neurons may be a specific feature of certain brain regions and not the entire brain, since no clear cilia directionality was reported in the thalamus, amygdala, hypothalamus, or in interneurons of the hippocampus [67].

All these findings suggest that neuronal cilia are dynamic and diverse, but the underlying molecular mechanisms and the physiological relevance for this apparent diversity remain elusive. It should be noted, however, that cilia in neurons are less dynamic than in neuronal progenitors (Box 1). Future work, identifying the specific molecular content of cilia in different subpopulations of neurons at different developmental and adult stages, may reveal the importance of compositional diversity of neuronal cilia and shed light on their respective functions in the brain.

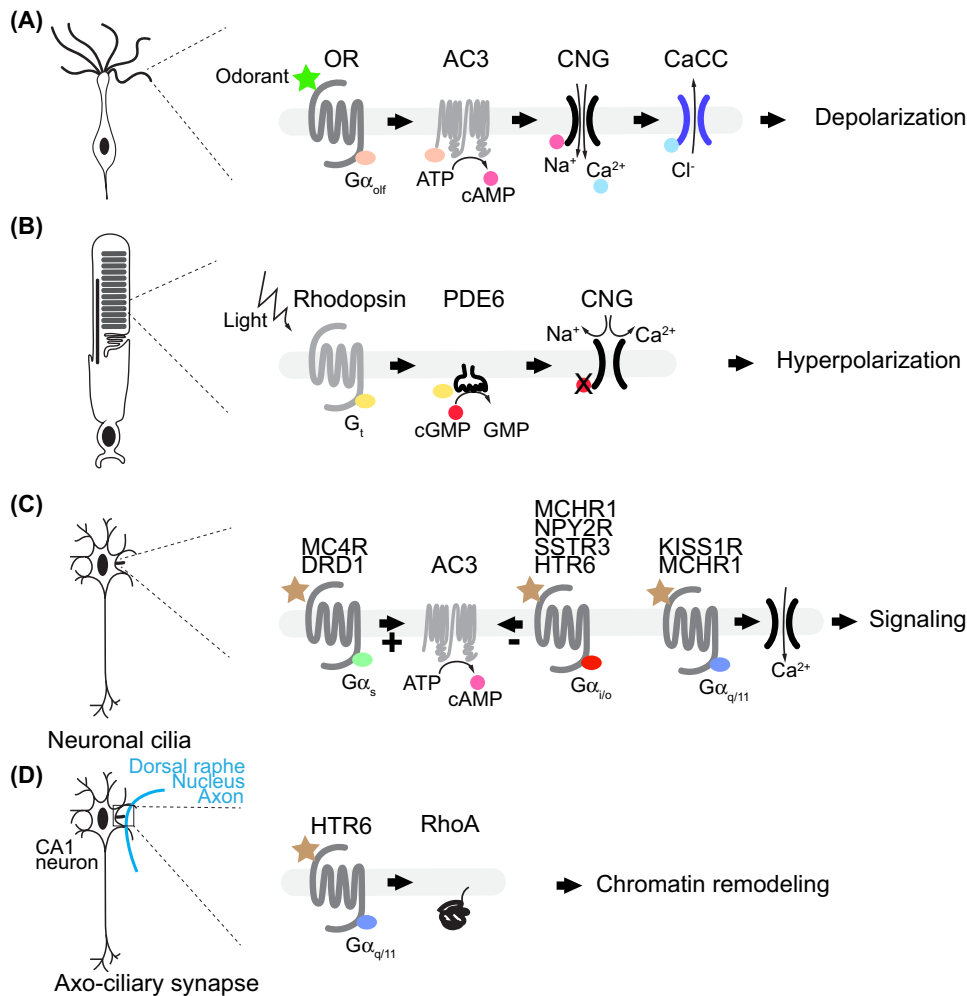
### How do neuronal cilia receive, transmit, and communicate information?

Cilia are typically enriched with receptors, especially GPCRs [78]. Examples are olfactory receptors, rhodopsin, receptors for neuropeptides (SSTR3, MCHR1, NPY2R, NPY5R, MC4R, and KISS1R), for neurotransmitters (HTR6 and DRD1) [78], and receptors involved in Hedgehog signaling (Box 2). The

list of signaling proteins in neuronal cilia keeps growing due to new techniques, including *in vivo* BiLD-based cilia proteomics in neurons [79].

### Ciliary signaling in neurons: a critical role for cAMP

Analogous to the olfactory system, where olfactory signals trigger a cAMP rise following activation of the  $G_{\alpha_{olf}}$  and adenylate cyclase [80] (Figure 2A), cilia-localized GPCRs in neurons also commonly signal through cAMP (Figure 2C). However, they can elicit either an increase of cAMP via  $G_{\alpha_s}$  (DRD1 and MC4R) or a decrease via  $G_{\alpha_{i/o}}$  (SSTR3, 5HT6, NPY2R, and MCHR1). Some



### Trends in Neurosciences

**Figure 2. Ciliary signaling cascades in sensory and central nervous system neurons.** (A) In olfactory sensory neurons, binding of an odorant molecule to its olfactory receptor (OR) triggers the activation of  $G_{\alpha_{olf}}$ , which stimulates AC3 and cAMP production. cAMP activates the CNG channel, leading to the entry of  $Ca^{2+}$ , the  $Ca^{2+}$ -dependent chloride channel (CaCC) opening, and membrane depolarization. (B) In photoreceptors, activation of rhodopsin by light stimulates the G protein transducing ( $G_t$ ), which leads to cGMP hydrolysis by PDE6. Reduction in cGMP results in closure of the CNG channel and photoreceptor hyperpolarization. (C) In neurons, most cilia-localized receptors are G protein-coupled receptors (GPCRs), associated with  $G_{\alpha_s}$  that stimulate AC3,  $G_{\alpha_{i/o}}$  that inhibits AC3 or  $G_{\alpha_{q/11}}$  that stimulate  $Ca^{2+}$  entry in the cells. All pathways lead to the activation of various signaling cascades. (D) In the context of the axo-ciliary synapse described between CA1 neurons and dorsal raphe nucleus axons, activation of HTR6 by serotonin stimulates RhoA through  $G_{\alpha_{q/11}}$  and leads to chromatin remodeling.

GPCRs do not regulate cAMP levels, but instead stimulate  $\text{Ca}^{2+}$  changes via  $G_{\alpha q/11}$  signaling (KISS1R) [78]. Notably, cAMP synthesis in the neuronal cilia is performed by the adenylate cyclase AC3, which solely localizes to the cilia [61].

However, to date, how cilia signaling modulates neuronal activity remains elusive. This contrasts to olfactory cilia (Figure 2A), where the immediate downstream effectors of cAMP signals are well characterized, and involve the sequential activation of cyclic nucleotide-gated (CNG) channels and the  $\text{Ca}^{2+}$ -activated  $\text{Cl}^-$  channel, which induces neuronal depolarization [18,80–82]. Similarly, the cascade of events in photoreceptors, following activation of rhodopsin, is well defined (Figure 2B). It involves reduction of cGMP levels by the phosphodiesterase PDE6, followed by an inactivation of the CNG channel, leading to hyperpolarization of photoreceptors [83].

CNG channels, which play a central role in converting cAMP and cGMP levels into voltage changes in sensory neurons, have not been described in the context of neuronal cilia in the CNS [18]. Thus, GPCR signaling must be decoded by different mechanisms that remain to be discovered. Besides, how ciliary signals are transmitted to the cell soma to modulate neuronal physiology remains unknown. One possibility is that the secondary messengers  $\text{Ca}^{2+}$  or cAMP leak out from the cilia and bind effectors in the soma. However, this would be slow and inefficient as the number of molecules are low due to the large difference between the ciliary and somatic volume. In agreement with this, reports have identified that ciliary  $\text{Ca}^{2+}$ , whose concentration is six- to sevenfold higher in the ciliary compartment than in the cytoplasm, is poorly coupled to somatic  $\text{Ca}^{2+}$  [84,85]. Ciliary  $\text{Ca}^{2+}$  levels are elevated due to the activity of  $\text{Ca}^{2+}$ -permeant channels, including the heteromeric TRP channel, PKD1L1–PKD2L1, which are enriched in the cilium and counterbalance steady diffusion of  $\text{Ca}^{2+}$  into the cytoplasm at its base [84,85].

How then is ciliary signaling translated into a cellular response? Advances in spatially resolved optogenetics have started to answer this long-standing question. In renal epithelial cells, the increase of cAMP in the cilium, but not in the cell body, evoked a distinct gene expression program, mediated through PKA-dependent phosphorylation of CREB in the cilia [9]. Ciliary PKA is also involved in cAMP-dependent inhibition of Hedgehog signaling in the zebrafish somites [8]. Whether similar processes and effectors are involved in neuronal physiology remains to be clarified.

Another major question that remains to be addressed is why some neuromodulatory GPCRs are in the cilia and some not, and what are the implications of ciliary localization of GPCRs on signaling. A recent study has identified that manipulating solely the ciliary localization of the extraretinal opsin GPCR alter its response kinetics in zebrafish spinal neurons [86]. This suggests that ciliary localization *per se* provides a mechanism for modulating GPCR signaling [86], which may depend on the distinct lipid content of the ciliary membrane [87,88].

#### Cilium–synapse interactions

Most ciliary GPCRs respond either to neuropeptides or to serotonin and dopamine, which are neurotransmitters usually released within the synapse [89]. Recently, it has been shown that cilia can form a synapse with serotonergic neurons in mice [90]. The so-called axociliary synapse elicited chromatin remodeling in hippocampal CA1 pyramidal neurons through a non-canonical  $G_{\alpha q}$ -RhoA pathway, distinct from serotonin sensing at the plasma membrane (Figure 2D). Not all neurons analyzed in this brain region form axociliary synapses, suggesting that some neurons receive different axonal and ciliary inputs.

Two recent electron microscopy studies in the human and mouse cortex investigated further cilium–synapse interactions [49,50]. While neither of these two studies reported an axociliary



synapse, they showed that neuronal cilia are commonly found adjacent to chemical synapses [49,50], axonal segments containing vesicles [49], and dense core vesicles [50]. In the human cortex, 51% of cilia engaged in synaptic contacts and contacted up to 47 axons [49]. Ciliary proximity to synapses appears, however, random and not enriched in the synapse-rich neuropil [50]. Even in the absence of direct axociliary synapse, cilia would still have access to synaptic spillover of neuromodulators to the same extent as adjacent astrocytes, and, thereby, are well positioned to sense synaptic activity. Future work, looking across the nervous system at cilia connectomics and their functional implications, will be needed to understand the broad implication of cilium–axon interactions.

### What is the impact of cilia on neuronal activity and animal behavior?

Ciliary function in sensory neurons is well established. Cilia detect the sensory modalities and trigger neuronal depolarization or hyperpolarization through the activity of ion channels [80,83] (Figure 2A,B). Hence, defects in ciliogenesis or ciliary dysfunction led to reduced neuronal excitability and sensory deficits, including loss of smell, retinal degeneration and blindness, and hearing impairment [14,21,22].

In contrast to sensory neurons, cilia in CNS neurons play a neuromodulatory role through controlling cellular signaling. However, studying how cilia modulate neuronal activity and animal behavior has been challenging. Given the significant role of cilia in embryonic and postnatal brain development [5,10,14,71,91,92], it is difficult to disentangle the impact of cilia on neurophysiology from a possible consequence of cilia-related developmental defects. Nevertheless, multiple groups have started to investigate the impact of cilia on neuronal physiology.

### Lessons learnt from the hypothalamus

So far, much of the knowledge acquired on ciliary function in neurons comes from studying the hypothalamus, where cilia seem to regulate food intake. The following observations have promoted this research: (i) conditional loss of cilia in adult mice results in hyperphagia and obesity; (ii) this phenotype is mirrored in mice with neuron-specific cilia loss, particularly in pro-opiomelanocortin (POMC)-expressing neurons in the hypothalamus [93,94] during development [95]; and (iii) certain ciliopathy patients, such as those with Alström or Bardet–Biedl syndrome (BBS), exhibit obesity as a major symptom [96–99]. Loss of function of the BBSome, a multimeric protein complex made of several BBS proteins, affects retrograde ciliary protein trafficking and, thereby, interferes with the ciliary localization of GPCRs, including MCHR1, SSTR3 [100], NPY2R [101,102], and MC4R [103]. MC4R expression in the paraventricular nucleus (PVN) of the hypothalamus controls food intake [104], and MC4R activation in MC4R-expressing PVN neurons is cilia-dependent [105]. This underscores that GPCR signaling in cilia controls energy homeostasis, and supports the notion that GPCR dysregulation in BBS underlies the hyperphagia [98,99]. Besides controlling GPCRs by the BBSome, the ciliary localization of ARL13B is also crucial for preventing hyperphagia and obesity development [106]. In addition, mutations in the *ADCY3* gene, encoding AC3, lead to obesity in both humans [103,107–109] and mice [110], underscoring a critical role of ciliary cAMP signaling in the hypothalamus.

High-fat diet feeding in mice reduces ciliation in POMC neurons [111], and blocking autophagy evokes a similar effect on the ciliogenesis [111], connecting high-fat diet, autophagy, ciliary dysfunction, and obesity development. Another study further supports this by indicating that cilia in the hypothalamus control lysosomal degradation of autophagic vacuoles [95].

The hypothalamus, especially the suprachiasmatic nucleus, also controls circadian rhythm. In this context, cilia were shown to maintain the coupling of cellular oscillators in the suprachiasmatic

nucleus to regulate circadian rhythms. Notably, ciliary dysfunction accelerates the adjustment to an altered light environment in a jet-lag situation through cilia-dependent Sonic Hedgehog signaling in neuromedin S-expressing neurons [68,112].

#### Control of neuronal excitability by cilia

Besides triggering a cascade of ciliary signaling, cilia were shown to directly modulate neuronal excitability by controlling synapse function in cortical neurons and striatal interneurons [66,113] and the neurophysiology of hippocampal neurons [114]. Acute disruption of ciliary function by shRNA in cultures of cortical neurons led to increased excitatory synapses, AMPAR-mediated excitatory currents, and spontaneous firing rates [66]. The application of SSTR3 agonists and antagonists also affected the levels of synaptic markers Shank3 and Vglut1 in a cilia-dependent manner [66]. Besides regulating synaptic activity, neuronal cilia were shown to regulate hippocampal excitability through the action of cilia-localized PKD2L1 channels [114]. Loss of PKD2L1 and electrical currents within cilia impaired ciliary maturation in mice, leading to autism-like behavioral features and seizure susceptibility [114].

#### Concluding remarks and future perspectives

Research over the past few decades has been instrumental in identifying cilia as evolutionary conserved organelles, playing a wide range of functions across the body and the nervous system. In turn, molecules involved in cilia biogenesis and maintenance and their association with human diseases and neurological symptoms have been identified [1,2,12,14,22]. Even though research has started to clarify how cilia control neuronal signaling, excitability, and function, a broader picture is still needed of the scale of the impact of cilia on neural circuits, animal behavior, and neurological disorders. Ciliary abnormalities are increasingly observed in neurological diseases that do not fall under classical ciliopathies, such as microcephaly, neurodegeneration, and neuropsychiatric disorders. For instance, a delay in ciliary disassembly can result in neural precursor cell depletion and microcephaly, a neurodevelopmental disorder that predominantly affects brain growth [10,11,115]. In human and mouse models of Parkinson's disease, neural progenitors and neurons display shortened cilia, accompanied by an elevated Sonic Hedgehog signal transduction [116]. Similarly, shortened cilia were identified in schizophrenia-causing PCM1 (pericentriolar material 1) and DISC1 (disrupted-in-schizophrenia-1) mutations [117,118]. These observations suggest that ciliary dysfunction may be a common feature among neurological disorders, which will require further investigation.

Even though cilia show conserved features, they are remarkably diverse and dynamic in their structure and content. However, the functional implications of cilia diversity and dynamics still need to be better understood, especially in the brain (see [Outstanding questions](#)). While parallels can be made between the mode of action of cilia in sensory systems and the CNS, compositional differences, such as the presence or absence of CNG channels, imply that ciliary signals could be interpreted in various ways in a context-dependent manner in different cell types.

The development of cutting-edge tools, including spatially resolved proteomics, optogenetics, chemogenetics, biosensors [1,79,119,120], and advances in genetic engineering are now opening new avenues for shedding light on the functional implications of ciliary diversity and signaling. Up to now, many studies have ablated cilia entirely to investigate their function. However, primarily through studying BBS mutants, it has become clear that impaired trafficking and signaling will affect ciliary function differently than changes in cilia presence or length. It is, therefore, essential to manipulate ciliary functions using different tools and approaches to generate a comprehensive understanding of context-specific ciliary functions.

#### Outstanding questions

How diverse and dynamic are cilia in neuronal cells?

What is the time scale of cilia dynamic (ranging from seconds to minutes and up to the lifetime) in an organism?

What are the molecular mechanisms underlying ciliary diversity and dynamics in neuronal populations?

What are the similarities and differences between cilia in the developing and the mature brain? What can we learn from studying developmental processes about ciliary biology in mature systems?

How is the signal, sensed by the cilia, transduced to the cell soma? Are there specific and conserved signaling avenues transducing ciliary signaling in the soma and what are their identities?

Cilia have been described near synapses in the cortex and have even been shown to form direct synapses with mouse hippocampal neurons. Are they common or rare among neurons? Are these contact points dynamic? Which brain area exhibits axon-ciliary synapses? What is the physiological role of cilium-synapse contacts in the nervous system?

Why are some neuromodulatory GPCRs in the cilia and some not?

How is the specificity achieved between ciliary and synaptic signaling? How is the information differentially encoded in the two locations?

What are the impacts of cilia on neuronal physiology? Which molecules are involved beyond GPCRs? To what extent are ciliary ion channels and membrane potential involved in neuromodulation?

How specific is ciliary signaling in given neuronal populations, and to what degree does signaling depend on ciliary dynamics?

What is the link between ciliary dysfunction and neurological manifestations in disorders that are not classical ciliopathies, such as microcephaly, neurodegeneration, and neuropsychiatric disorders?

### Acknowledgments

N.J.-Y. is funded by The Research Council of Norway (FRIPRO grant 314189) and the Deutsche Forschungsgemeinschaft (DFG, Germany Research Foundation), FOR5547 - Project-ID 503306912. D.W. is funded by DFG: SFB 1454 - Project-ID 432325352, TRR333/1 - Project-ID 450149205, FOR5547 - Project-ID 503306912, WA 3382/8-1 - Project-ID 513767027, under Germany's Excellence Strategy - EXC2151 - Project-ID 390873048, and the Else Kröner Fresenius Foundation (2021.EKFSE.53). J.G. is supported by the SPP2127-GO 2301/5-2 'Gene and Cell-Based Therapies to Counteract Neuroretinal Degeneration and DFG FOR5547 - Project-ID 503306912. We thank Dr Pawel Burkhardt and Prof Gaspar Jékely for constructive discussions on the evolutionary conservation of neuronal cilia.

### Declaration of interests

The authors declare no competing interests in relation to this work.

### References

- Gopalakrishnan, J. *et al.* (2023) Emerging principles of primary cilia dynamics in controlling tissue organization and function. *EMBO J.* e113891
- Mill, P. *et al.* (2023) Primary cilia as dynamic and diverse signalling hubs in development and disease. *Nat. Rev. Genet.* 24, 421–441
- Mitchison, H.M. and Valente, E.M. (2017) Motile and non-motile cilia in human pathology: from function to phenotypes. *J. Pathol.* 241, 294–309
- Wan, K.Y. and Jékely, G. (2020) On the unity and diversity of cilia. *Philos. Trans. R. Soc. B Biol. Sci.* 375, 20190148
- Guemez-Gamboa, A. *et al.* (2014) Primary cilia in the developing and mature brain. *Neuron* 82, 511–521
- Moran, A.L. *et al.* (2024) Transport and barrier mechanisms that regulate ciliary compartmentalization and ciliopathies. *Nat. Rev. Nephrol.* 20, 83–100
- Park, K. and Leroux, M.R. (2022) Composition, organization and mechanisms of the transition zone, a gate for the cilium. *EMBO Rep.* 23, e55420
- Truong, M.E. *et al.* (2021) Vertebrate cells differentially interpret ciliary and extraciliary cAMP. *Cell* 184, 2911–2926.e2918
- Hansen, J.N. *et al.* (2022) A cAMP signalosome in primary cilia drives gene expression and kidney cyst formation. *EMBO Rep.* 23, e54315
- Gabriel, E. *et al.* (2016) CPAP promotes timely cilium disassembly to maintain neural progenitor pool. *EMBO J.* 35, 803–819
- Zhang, W. *et al.* (2019) Modeling microcephaly with cerebral organoids reveals a WDR62-CEP170-KIF2A pathway promoting cilium disassembly in neural progenitors. *Nat. Commun.* 10, 2612
- Andreu-Cervera, A. *et al.* (2021) Cilia, ciliopathies and hedgehog-related forebrain developmental disorders. *Neurobiol. Dis.* 150, 105236
- Bear, R.M. and Caspary, T. (2023) Uncovering cilia function in glial development. *Ann. Hum. Genet.* 88, 27–44
- Suciu, S.K. and Caspary, T. (2021) Cilia, neural development and disease. *Semin. Cell Dev. Biol.* 110, 34–42
- Stoufflet, J. and Caillé, I. (2022) The primary cilium and neuronal migration. *Cells* 11, 3384
- Ache, B.W. and Young, J.M. (2005) Olfaction: diverse species, conserved principles. *Neuron* 48, 417–430
- McClintock, T.S. *et al.* (2020) Maturation of the olfactory sensory neuron and its cilia. *Chem. Senses* 45, 805–822
- Kaup, U.B. and Seifert, R. (2002) Cyclic nucleotide-gated ion channels. *Physiol. Rev.* 82, 769–824
- Mukhopadhyay, S. *et al.* (2007) Distinct IFT mechanisms contribute to the generation of ciliary structural diversity in *C. elegans*. *EMBO J.* 26, 2966–2980
- Poncelet, G. and Shimeld, S.M. (2020) The evolutionary origins of the vertebrate olfactory system. *Open Biol.* 10, 200330
- Bachmann-Gagescu, R. and Neuhauss, S.C. (2019) The photoreceptor cilium and its diseases. *Curr. Opin. Genet. Dev.* 56, 22–33
- Reiter, J.F. and Leroux, M.R. (2017) Genes and molecular pathways underpinning ciliopathies. *Nat. Rev. Mol. Cell Biol.* 18, 533–547
- Hudspeth, A.J. and Jacobs, R. (1979) Stereocilia mediate transduction in vertebrate hair cells (auditory system/cilium/vestibular system). *Proc. Natl. Acad. Sci. U. S. A.* 76, 1506–1509
- Bezares-Calderón, L.A. *et al.* (2020) Diversity of cilia-based mechanosensory systems and their functions in marine animal behaviour. *Philos. Trans. R. Soc. B Biol. Sci.* 375, 20190376
- Sun, L. *et al.* (2019) Ultrastructural organization of NompC in the mechanoreceptive organelle of *Drosophila* campaniform mechanoreceptors. *Proc. Natl. Acad. Sci. U. S. A.* 116, 7343–7352
- Kindt, K.S. *et al.* (2012) Kinocilia mediate mechanosensitivity in developing zebrafish hair cells. *Dev. Cell* 23, 329–341
- Williams, C.L. *et al.* (2014) Direct evidence for BBSome-associated intraflagellar transport reveals distinct properties of native mammalian cilia. *Nat. Commun.* 5, 5813
- Flannery, R.J. *et al.* (2006) Clustering of cyclic-nucleotide-gated channels in olfactory cilia. *Biophys. J.* 91, 179–188
- Doroquez, D.B. *et al.* (2014) A high-resolution morphological and ultrastructural map of anterior sensory cilia and glia in *Caenorhabditis elegans*. *eLife* 3, e01948
- Bezares-Calderón, L.A. *et al.* (2018) Neural circuitry of a polycystin-mediated hydrodynamic startle response for predator avoidance. *eLife* 7, e36262
- Cheng, L.E. *et al.* (2010) The role of the TRP channel NompC in *Drosophila* larval and adult locomotion. *Neuron* 67, 373–380
- Xiang, W. *et al.* (2022) The dynamics of protein localisation to restricted zones within *Drosophila* mechanosensory cilia. *Sci. Rep.* 12, 13338
- Loreng, T.D. and Smith, E.F. (2017) The central apparatus of cilia and eukaryotic flagella. *Cold Spring Harb. Perspect. Biol.* 9, a028118
- Gui, M. *et al.* (2022) Ciliary central apparatus structure reveals mechanisms of microtubule patterning. *Nat. Struct. Mol. Biol.* 29, 483–492
- Rayamajhi, D. *et al.* (2024) The forkhead transcription factor Foxj1 controls vertebrate olfactory cilia biogenesis and sensory neuron differentiation. *PLoS Biol.* 22, e3002468
- O'Donnell, J. and Zheng, J. (2022) Vestibular hair cells require CAMSAP3, a microtubule minus-end regulator, for formation of normal kinocilia. *Front. Cell. Neurosci.* 16, 876805
- Mercey, O. *et al.* (2022) The connecting cilium inner scaffold provides a structural foundation that protects against retinal degeneration. *PLoS Biol.* 20, e3001649
- Yu, X. *et al.* (2011) Cilia-driven fluid flow as an epigenetic cue for otolith biomineralization on sensory hair cells of the inner ear. *Development* 138, 487–494
- Whitfield, T.T. (2020) Cilia in the developing zebrafish ear. *Philos. Trans. R. Soc. B Biol. Sci.* 375, 20190163
- Ringers, C. *et al.* (2020) The role of motile cilia in the development and physiology of the nervous system. *Philos. Trans. R. Soc. B Biol. Sci.* 375, 20190156
- Akella, J.S. *et al.* (2023) Cell-specific  $\alpha$ -tubulin TBA-6 and pan-ciliary IFT cargo RAB-28 generate a non-canonical transition zone. *bioRxiv*, Published online November 17, 2023. <https://doi.org/10.1101/2023.11.16.567340>
- Mechaussier, S. *et al.* (2022) TUBB4B variants specifically impact ciliary function, causing a ciliopathic spectrum. *medRxiv*, Published online October 21, 2022. <https://doi.org/10.1101/2022.10.19.22280748>

What is the evolutionary basis of neuronal cilia? Why do certain species lack cilia on their neurons, while others require cilia for proper brain development and physiology?

43. Derderian, C. *et al.* (2023) Seriously cilia: a tiny organelle illuminates evolution, disease, and intercellular communication. *Dev. Cell* 58, 1333–1349
44. Musser, J.M. *et al.* (2021) Profiling cellular diversity in sponges informs animal cell type and nervous system evolution. *Science* 374, 717–723
45. Sachkova, M.Y. *et al.* (2021) Neuropeptide repertoire and 3D anatomy of the ctenophore nervous system. *Curr. Biol.* 31, 5274–5285.e5276
46. Sigg, M.A. *et al.* (2017) Evolutionary proteomics uncovers ancient associations of cilia with signaling pathways. *Dev. Cell* 43, 744–762.e711
47. Dobbelaere, J. *et al.* (2023) A phylogenetic profiling approach identifies novel cilogenesis genes in *Drosophila* and *C. elegans*. *EMBO J.* 42, e113616
48. Arellano, J.I. *et al.* (2012) Development and distribution of neuronal cilia in mouse neocortex. *J. Comp. Neurol.* 520, 848–873
49. Wu, J.Y. *et al.* (2023) Mapping of neuronal and glial primary cilia contactome and connectome in the human cerebral cortex. *Neuron* 112, 41–55.e3
50. Ott, C.M. *et al.* (2023) Nanometer-scale views of visual cortex reveal anatomical features of primary cilia poised to detect synaptic spillover. *bioRxiv*, Published online November 6, 2023. <https://doi.org/10.1101/2023.10.31.564838>
51. Kinnamon, J.C. and Westfall, J.A. (1982) Types of neurons and synaptic connections at hypostome-tentacle junctions in Hydra. *J. Morphol.* 173, 119–128
52. Saripalli, L.D. and Westfall, J.A. (1996) Classification of nerve cells dissociated from tentacles of the sea anemone *Calliactis parasitica*. *Biol. Bull.* 190, 111–124
53. Westfall, J.A. and Epp, L.G. (1985) Scanning electron microscopy of neurons isolated from the pedal disk and body column of Hydra. *Tissue Cell* 17, 161–170
54. Ryan, K. *et al.* (2016) The CNS connectome of a tadpole larva of *Ciona intestinalis* (L.) highlights sidedness in the brain of a chordate sibling. *eLife* 5, e16962
55. Ryan, K. and Meinertzhagen, I.A. (2019) Neuronal identity: the neuron types of a simple chordate sibling, the tadpole larva of *Ciona intestinalis*. *Curr. Opin. Neurobiol.* 56, 47–60
56. Nechipurenko, I.V. and Sengupta, P. (2017) The rise and fall of basal bodies in the nematode *Caenorhabditis elegans*. *Cilia* 6, 9
57. Louvi, A. and Grove, E.A. (2011) Cilia in the CNS: the quiet organelle claims center stage. *Neuron* 69, 1046–1060
58. Song, Y. and Brady, S.T. (2015) Post-translational modifications of tubulin: pathways to functional diversity of microtubules. *Trends Cell Biol.* 25, 125–136
59. Wloga, D. *et al.* (2017) Posttranslational modifications of tubulin and cilia. *Cold Spring Harb. Perspect. Biol.* 9, a028159
60. Magiera, M.M. and Janke, C. (2014) Post-translational modifications of tubulin. *Curr. Biol.* 24, R351–R354
61. Bishop, G.A. *et al.* (2007) Type III adenylyl cyclase localizes to primary cilia throughout the adult mouse brain. *J. Comp. Neurol.* 505, 562–571
62. Caspary, T. *et al.* (2007) The graded response to Sonic Hedgehog depends on cilia architecture. *Dev. Cell* 12, 767–778
63. Händel, M. *et al.* (1999) Selective targeting of somatostatin receptor 3 to neuronal cilia. *Neuroscience* 89, 909–926
64. Dupuy, V. *et al.* (2023) Spatiotemporal dynamics of 5-HT(6) receptor ciliary localization during mouse brain development. *Neurobiol. Dis.* 176, 105949
65. Sipos, É. *et al.* (2018) Quantitative comparison of primary cilia marker expression and length in the mouse brain. *J. Mol. Neurosci.* 64, 397–409
66. Tereshko, L. *et al.* (2021) Ciliary neuropeptidergic signaling dynamically regulates excitatory synapses in postnatal neocortical pyramidal neurons. *eLife* 10, e65427
67. Yang, J. *et al.* (2023) Primary cilia directionality reveals a slow reverse migration of principal neurons for postnatal positioning and lamination in the cerebral cortex. *bioRxiv*, Published online November 08, 2023. <https://doi.org/10.1101/2021.12.21.473383>
68. Tu, H.-Q. *et al.* (2023) Rhythmic cilia changes support SCN neuron coherence in circadian clock. *Science* 380, 972–979
69. Monfared, R.V. *et al.* (2023) Spatiotemporal mapping of brain cilia reveals region-specific oscillation of length and orientation. *bioRxiv*, Published online July 15, 2023. <https://doi.org/10.1101/2023.06.28.546950>
70. Macarelli, V. *et al.* Regulation of the length of neuronal primary cilia and its potential effects on signalling. *Trends Cell Biol.* 33, 979–990
71. D’Gama, P.P. *et al.* (2023) Loss of cilia after neurulation impacts brain development and neuronal activity in larval zebrafish. *bioRxiv*, Published online September 21, 2023. <https://doi.org/10.1101/2023.09.20.558654>
72. Constable, S. *et al.* (2023) Programmed withdrawal of cilia maintenance followed by centriole capping leads to permanent cilia loss during cerebellar granule cell neurogenesis. *bioRxiv*, Published online December 11, 2023. <https://doi.org/10.1101/2023.12.07.565993>
73. Ott, C.M. *et al.* (2023) Permanent deconstruction of intracellular primary cilia in differentiating granule cell neurons. *bioRxiv*, Published online December 8, 2023. <https://doi.org/10.1101/2023.12.07.565988>
74. Elke Gabriel, W.A. *et al.* (2023) Generation of iPSC-derived human forebrain organoids assembling bilateral eye primordia. *Nat. Protoc.* 18, 1893–1929
75. Alhassen, W. *et al.* (2022) Regulation of brain primary cilia length by MCH signaling: evidence from pharmacological, genetic, optogenetic, and chemogenic manipulations. *Mol. Neurobiol.* 59, 245–265
76. Han, Y.M. *et al.* (2014) Leptin-promoted cilia assembly is critical for normal energy balance. *J. Clin. Invest.* 124, 2193–2197
77. Macarelli, V. *et al.* (2023) Regulation of the length of neuronal primary cilia and its potential effects on signalling. *Trends Cell Biol.* 33, 979–990
78. Wachten, D. and Mick, D.U. (2021) Signal transduction in primary cilia – analyzing and manipulating GPCR and second messenger signaling. *Pharmacol. Ther.* 224, 107836
79. Loukil, A. *et al.* (2023) Identification of new ciliary signaling pathways in the brain and insights into neurological disorders. *bioRxiv*, Published online December 21, 2023. <https://doi.org/10.1101/2023.12.20.572700>
80. Kaupp, U.B. (2010) Olfactory signalling in vertebrates and insects: differences and commonalities. *Nat. Rev. Neurosci.* 11, 188–200
81. Nakamura, T. and Gold, G.H. (1987) A cyclic nucleotide-gated conductance in olfactory receptor cilia. *Nature* 325, 442–444
82. Kleene, S.J. (1993) Origin of the chloride current in olfactory transduction. *Neuron* 11, 123–132
83. Gulati, S. and Palczewski, K. (2023) Structural view of G protein-coupled receptor signaling in the retinal rod outer segment. *Trends Biochem. Sci.* 48, 172–186
84. DeCaen, P.G. *et al.* (2013) Direct recording and molecular identification of the calcium channel of primary cilia. *Nature* 504, 315–318
85. Delling, M. *et al.* (2013) Primary cilia are specialized calcium signalling organelles. *Nature* 504, 311–314
86. Winans, A.M. *et al.* (2023) Ciliary localization of a light-activated neuronal GPCR shapes behavior. *Proc. Natl. Acad. Sci. U. S. A.* 120, e2311131120
87. Kinnebrew, M. *et al.* (2019) Cholesterol accessibility at the ciliary membrane controls hedgehog signaling. *eLife* 8, e50051
88. Nechipurenko, I.V. (2020) The enigmatic role of lipids in cilia signaling. *Front. Cell Dev. Biol.* 8, 777
89. Peters, K.Z. *et al.* (2021) Modulating the neuromodulators: dopamine, serotonin, and the endocannabinoid system. *Trends Neurosci.* 44, 464–477
90. Sheu, S.H. *et al.* (2022) A serotonergic axon-cilium synapse drives nuclear signaling to alter chromatin accessibility. *Cell* 185, 3390–3407.e3318
91. Tereshko, L. *et al.* (2022) Primary cilia in the postnatal brain: subcellular compartments for organizing neuromodulatory signaling. *Curr. Opin. Neurobiol.* 74, 102533
92. Ramani, A. *et al.* (2023) Reliability of high-quantity human brain organoids for modeling microcephaly, glioma invasion, and drug screening. *bioRxiv*, Published online October 30, 2023. <https://doi.org/10.1101/2023.10.29.564523>
93. Davenport, J.R. *et al.* (2007) Disruption of intraflagellar transport in adult mice leads to obesity and slow-onset cystic kidney disease. *Curr. Biol.* 17, 1586–1594



94. Sun, J.S. *et al.* (2021) Ventromedial hypothalamic primary cilia control energy and skeletal homeostasis. *J. Clin. Invest.* 131, e138107
95. Lee, C.H. *et al.* (2020) Primary cilia mediate early life programming of adiposity through lysosomal regulation in the developing mouse hypothalamus. *Nat. Commun.* 11, 5772
96. Forsythe, E. and Beales, P.L. (2013) Bardet-Biedl syndrome. *Eur. J. Hum. Genet.* 21, 8–13
97. Engle, S.E. *et al.* (2021) Cilia signaling and obesity. *Semin. Cell Dev. Biol.* 110, 43–50
98. DeMars, K.M. *et al.* (2023) Neuronal primary cilia integrate peripheral signals with metabolic drives. *Front. Physiol.* 14, 1150232
99. Brewer, K.M. *et al.* (2022) Neuronal cilia in energy homeostasis. *Front. Cell Dev. Biol.* 10, 1082141
100. Berbari, N.F. *et al.* (2008) Bardet-Biedl syndrome proteins are required for the localization of G protein-coupled receptors to primary cilia. *Proc. Natl. Acad. Sci. U. S. A.* 105, 4242–4246
101. Loktev, A.V. and Jackson, P.K. (2013) Neuropeptide Y family receptors traffic via the Bardet-Biedl syndrome pathway to signal in neuronal primary cilia. *Cell Rep.* 5, 1316–1329
102. Guo, D.F. *et al.* (2019) The BBSome in POMC and AgRP neurons is necessary for body weight regulation and sorting of metabolic receptors. *Diabetes* 68, 1591–1603
103. Siljee, J.E. *et al.* (2018) Subcellular localization of MC4R with ADCY3 at neuronal primary cilia underlies a common pathway for genetic predisposition to obesity. *Nat. Genet.* 50, 180–185
104. Shah, B.P. *et al.* (2014) MC4R-expressing glutamatergic neurons in the paraventricular hypothalamus regulate feeding and are synaptically connected to the parabrachial nucleus. *Proc. Natl. Acad. Sci. U. S. A.* 111, 13193–13198
105. Wang, Y. *et al.* (2021) Melanocortin 4 receptor signals at the neuronal primary cilium to control food intake and body weight. *J. Clin. Invest.* 131, 142064
106. Terry, T.T. *et al.* (2023) Ciliary ARL13B prevents obesity in mice. *bioRxiv*, Published online August 4, 2023. <https://doi.org/10.1101/2023.08.02.551695>
107. Saeed, S. *et al.* (2018) Loss-of-function mutations in ADCY3 cause monogenic severe obesity. *Nat. Genet.* 50, 175–179
108. Barroso, I. (2018) ADCY3, neuronal primary cilia and obesity. *Nat. Genet.* 50, 166–167
109. Grarup, N. *et al.* (2018) Loss-of-function variants in ADCY3 increase risk of obesity and type 2 diabetes. *Nat. Genet.* 50, 172–174
110. Wang, Z. *et al.* (2009) Adult type 3 adenylyl cyclase-deficient mice are obese. *PLoS One* 4, e6979
111. Ávalos, Y. *et al.* (2022) Palmitic acid control of ciliogenesis modulates insulin signaling in hypothalamic neurons through an autophagy-dependent mechanism. *Cell Death Dis.* 13, 659
112. Kim, D.W. and Blackshaw, S. (2023) A super Sonic circadian synchronizer. *Science* 380, 896–897
113. Guo, J. *et al.* (2017) Primary cilia signaling shapes the development of interneuronal connectivity. *Dev. Cell* 42, 286–300.e284
114. Vien, T.N. *et al.* (2023) Primary cilia TRP channel regulates hippocampal excitability. *Proc. Natl. Acad. Sci. U. S. A.* 120, e2219686120
115. Farooq, M. *et al.* (2020) RRP7A links primary microcephaly to dysfunction of ribosome biogenesis, resorption of primary cilia, and neurogenesis. *Nat. Commun.* 11, 5816
116. Schmidt, S. *et al.* (2022) Primary cilia and SHH signaling impairments in human and mouse models of Parkinson's disease. *Nat. Commun.* 13, 4819
117. Monroe, T.O. *et al.* (2020) PCM1 is necessary for focal ciliary integrity and is a candidate for severe schizophrenia. *Nat. Commun.* 11, 5903
118. Muñoz-Estrada, J. *et al.* (2018) Primary cilia formation is diminished in schizophrenia and bipolar disorder: a possible marker for these psychiatric diseases. *Schizophr. Res.* 195, 412–420
119. Mick, D.U. *et al.* (2015) Proteomics of primary cilia by proximity labeling. *Dev. Cell* 35, 497–512
120. May, E.A. *et al.* (2021) Time-resolved proteomics profiling of the ciliary Hedgehog response. *J. Cell Biol.* 220, e202007207
121. Wang, L. and Dynlacht, B.D. (2018) The regulation of cilium assembly and disassembly in development and disease. *Development* 145, dev.151407
122. Rakic, P. (1995) A small step for the cell, a giant leap for mankind: a hypothesis of neocortical expansion during evolution. *Trends Neurosci.* 18, 383–388
123. Florio, M. and Huttner, W.B. (2014) Neural progenitors, neurogenesis and the evolution of the neocortex. *Development* 141, 2182–2194
124. Hilgendorf, K.I. *et al.* (2024) Emerging mechanistic understanding of cilia function in cellular signalling. *Nat. Rev. Mol. Cell Biol.* <https://doi.org/10.1038/s41580-023-00698-5>
125. Bangs, F. and Anderson, K.V. (2017) Primary cilia and mammalian hedgehog signaling. *Cold Spring Harb. Perspect. Biol.* 9, a028175
126. Holtmann, M. and Thurm, U. (2001) Variations of concentric hair cells in a Cnidarian sensory epithelium (*Coryne tubulosa*). *J. Comp. Neurol.* 432, 550–563

## Review

## Regulation of the length of neuronal primary cilia and its potential effects on signalling

Viviana Macarelli,<sup>1,2</sup> Eleni Leventea,<sup>3</sup> and Florian T. Merkle <sup>1,2,\*</sup>

**Primary cilia protrude from most vertebrate cell bodies and act as specialized ‘signalling antennae’ that can substantially lengthen or retract in minutes to hours in response to specific stimuli. Here, we review the conditions and mechanisms responsible for regulating primary cilia length (PCL) in mammalian nonsensory neurons, and propose four models of how they could affect ciliary signalling and alter cell state and suggest experiments to distinguish between them. These models include (i) the passive indicator model, where changes in PCL have no consequence; (ii) the rheostat model, in which a longer cilium enhances signalling; (iii) the local concentration model, where ciliary shortening increases the local protein concentration to facilitate signalling; and (iv) the altered composition model where changes in PCL skew signalling.**

### Dynamic changes in the length of primary cilia

Primary cilia are long, thin, nonmotile organelles that project from the plasma membrane of most vertebrate cell types (Figure 1). At their core is an **axoneme** (see Glossary) of microtubule doublets that arise from a centriole-derived basal body, which anchors to the plasma membrane during the cell cycle’s interphase [1]. The axoneme is enclosed in a lipid bilayer that differs from the plasma membrane in both its lipid and protein content, making the cilium a privileged cellular compartment with a distinct membrane and matrix composition [2]. The distinctive composition of the cilium is maintained by a **transition zone** complex, a specialized domain at the ciliary base [3] that prevents the free diffusion of large soluble proteins ( $\geq 100$  kDa) and vesicles containing membrane-associated proteins that require specific import mechanisms [2]. Once at the base of the cilium, an encoded ciliary localisation sequence enables these membrane proteins to couple to the **intraflagellar transport (IFT)** machinery aided by the tubby family proteins TUB and TULP3 [4–6]. The IFT proteins form two main complexes, IFT-A and IFT-B, which assemble into long polymeric structures known as ‘trains’ coupled to Kinesin-2 and Dynein-2 motor proteins and their cargoes (Figure 1). IFT-B associated Kinesin-2 moves the train in the anterograde direction toward the ciliary tip, where molecular rearrangements allow the activation of Dynein-2 and the inhibition of Kinesin-2, resulting in retrograde transport back to the ciliary base with the help of the **BBSome** [2,7]. IFT plays a central role in ciliogenesis and maintenance of the cilium’s composition, including the composition of the membrane proteins involved in signalling [2]. Indeed, as discussed below, primary cilia play a crucial role in cell signalling, since genetic mutations that disrupt the function of primary cilia in the brain can have profound effects on the development, physiology, and behaviour of the organism, and can lead to a class of diseases collectively referred to as **primary ciliopathies** [8].

Here, we consider another mechanism that may affect ciliary signalling: dynamic alterations in **PCL**. The length of primary cilia can vary over a wide range across different cell types, organisms, and stages of the cell cycle, since cilia are disassembled during cell division and then later reassembled, as reviewed elsewhere [1]. For clarity, here, we consider the mechanisms that affect the PCL in postmitotic **nonsensory neurons** of the healthy mammalian central nervous

### Highlights

Most vertebrate cells, including mammalian postmitotic neurons, have fingerlike primary cilia that protrude from the cell body and act as specialized ‘signalling antennae’ enriched with specific receptors.

Primary cilia length (PCL) can substantially change within minutes to hours in response to physiological changes such as circadian rhythms and metabolic state, and secreted factors such as neurotransmitters and hormones.

The functional consequences of changes in PCL are still poorly understood, but are likely to affect cellular responses to exogenous signals across many cell types and processes.

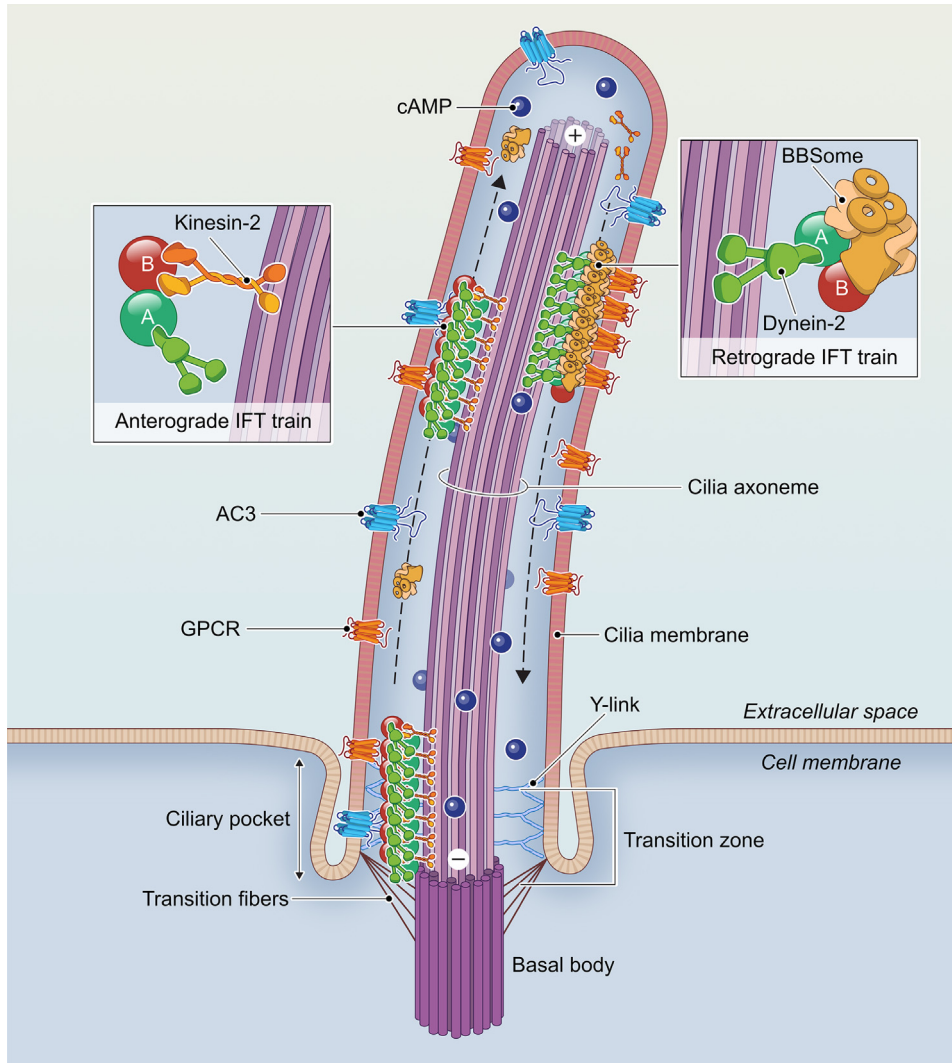
<sup>1</sup>Metabolic Research Laboratories, Wellcome Trust – Medical Research Council Institute of Metabolic Science, University of Cambridge, Cambridge CB2 0QQ, UK

<sup>2</sup>Wellcome Trust – Medical Research Council Cambridge Stem Cell Institute, University of Cambridge, Cambridge CB2 0AW, UK

<sup>3</sup>Wolfson Diabetes and Endocrine Clinic, Cambridge University Hospitals NHS Foundation Trust, Cambridge CB2 0QQ, UK

\*Correspondence: [fm436@medschl.cam.ac.uk](mailto:fm436@medschl.cam.ac.uk) (F.T. Merkle).





Trends in Cell Biology

**Figure 1. Structure of neuronal primary cilia.** The primary cilium projects from the cell body and is built around a microtubule-based axoneme templated from the basal body at the ciliary base. Here, certain membrane proteins engage with adaptor proteins to couple to IFT trains that transport them through the transition zone and along the axoneme toward the cilium's tip. This active process gives the ciliary membrane a protein and phospholipid composition that is distinct from that of the plasma membrane, indicated by different shading. Abbreviations: AC3, adenylate cyclase 3; cAMP, cyclic AMP; GPCR, G-protein coupled receptor; IFT, intraflagellar transport.

system, in which primary cilia are relatively stably established, and where the signalling role of primary cilia is likely to contribute to normal brain function [9]. We further define changes in PCL as alterations in the mean PCL of a given cell population, since the absolute length varies within a given cell type. We do not consider mutations that significantly perturb the cilium's structure or function, but instead focus on the responses of primary cilia to physiological changes. Finally, although changes in PCL are also associated with aging [10] and pathological conditions such as Alzheimer's disease [11] and Parkinson's disease [12], the long timeframe of these conditions complicates the discovery of causal relationships, so we focus here on the changes in PCL that occur within the timeframe of minutes to hours. We propose that these stimulus-induced

### Glossary

**Actin stress fibres:** parallel bundles of actin filaments found in nonmuscle cells that are anchored at the cell membrane, and provide mechanical support and contractile force to cells. The formation of stress fibres is regulated by the Rho family of GTPases.

**Adenylate cyclase 3 (AC3):** an enzyme that catalyzes the conversion of ATP to cyclic AMP (cAMP) upon activation by Gα<sub>s</sub>-coupled G-protein coupled receptors (GPCRs). It localizes to primary cilia in several cell types, including many neuron types in the mammalian brain.

**Axoneme:** in nonmotile mammalian primary cilia, the axoneme consists of nine microtubule doublets regularly positioned along the circumference of a circle in a so-called '9+0' configuration, and it serves as the structural core of the primary cilium.

**BBSome:** a complex of eight proteins associated with Bardet-Biedl syndrome (BBS) that are mainly involved in regulating the exit of membrane proteins from the primary cilium, and also the ciliary entry of certain GPCRs by coupling them to the intraflagellar transport (IFT) machinery.

**Ectosomes:** vesicles released from the plasma membrane of cells. They can contain lipids, proteins, and nucleic acids, and are involved in a variety of cellular processes. Ciliary ectosomes are released from the tip of primary cilia and can be enriched in specific ciliary proteins such as GPCRs. Ectosomes are not to be confused with exosomes, which are small endosomally derived membrane microvesicles.

**Intraflagellar transport (IFT):** the process by which molecules and proteins are actively transported along the primary cilia axoneme toward the tip by kinesin motor proteins and toward the ciliary base by dynein motor proteins.

**Nonsensory neurons:** neurons of the central nervous system that are not directly involved in sensing inputs from the environment and that have more typical primary cilia, as opposed to sensory neurons such as photoreceptor neurons, which respond to photons, or olfactory neurons, which respond to olfactory molecules via highly specialized primary cilia.

**Primary ciliopathies:** a group of genetic disorders caused by defects in the structure or function of primary cilia that affects their signalling role. This

changes in PCL in mammalian neurons may alter the concentration and composition of ciliary membrane proteins to affect the function of primary cilia at the level of an organelle, which, in turn, could affect cellular behaviour, ultimately leading to behavioural and/or physiological changes at the level of the organism.

In mammalian neurons, primary cilia range in length from approximately 2  $\mu\text{m}$  [13–16] to 12  $\mu\text{m}$  [11,12,17]. Primary cilia conservatively have 100-fold less surface area and 2500-fold less volume than a cell body, assuming a soma 10  $\mu\text{m}$  in diameter and a primary cilium 5  $\mu\text{m}$  in length and 0.2  $\mu\text{m}$  in diameter. In reality, these differences are likely to be much more pronounced since a neuron's axon and dendrites contain the overwhelming majority of its volume and surface area. Thus, ciliary localisation enables a substantially higher local concentration of ciliary membrane proteins to facilitate efficient signal detection and transduction, leading to the widespread recognition of primary specialized signalling hub or 'sensory antenna' [3]. Indeed, primary cilia play a pivotal role in olfactory sensory neurons, which have multiple cilia that are highly enriched in odorant-sensitive G-protein coupled receptors (GPCRs) [3,18], and photoreceptor neurons in which the outer segments are specialized primary cilia that house the light-sensitive GPCR rhodopsin [19]. They are also important mediators of numerous signalling pathways including the Sonic Hedgehog (SHH), Wntless-related integration site (WNT), and Transforming Growth Factor  $\beta$  and Bone Morphogenetic Protein (TGF $\beta$ /BMP) pathways, and thereby broadly affect cell specification, tissue development, and tissue homeostasis [18].

The essential role of primary cilia in mediating signal transduction during development is clearly demonstrated by the severe defects that can emerge when cilia are functionally perturbed in primary ciliopathies [8]. These diseases can arise from disruptions in the cilia's structure, as seen with the loss of IFT proteins [8], and from disruptions in the cilia's function, for example, due to an altered ciliary protein composition, as seen with mutations of the genes associated with Bardet-Biedl syndrome (BBS) [8] that perturb the normal ciliary localisation of GPCRs, including SSTR3, MCHR1, and NPY2R [6,20]. Each of these GPCRs is involved in regulating food intake and body weight, suggesting that their disrupted signalling may contribute to the hyperphagia and obesity in patients with BBS mutations. In addition to providing a static environment where receptors, transducers, and second messengers are brought close together, primary cilia can also mediate signalling by dynamically changing the localisation of ciliary proteins. For example, SHH signalling is regulated by the trafficking of the receptors Patched 1 (PTCH1) and GPR161 out of the cilium upon ligand binding, and the ciliary entry of Smoothened (SMO), which then activates the downstream mediators [21].

### Physiological drivers of changes in neuronal PCL

Neuronal primary cilia can substantially shrink or grow in response to either physiological or pathological processes [11,22], and exhibit significant changes in length within hours [23,24]. For example, in the hippocampal neurons of mice, acutely administered serotonin increases PCL, partly through activation of the ciliary serotonin receptor HTR6 [11]. Indeed, the overexpression of HTR6 elongates cilia in the cortical neurons of mice [17], while HTR6 antagonists reduce the length of cilia in the primary striatal neurons of mice [23]. Similarly, agonists of the G $\alpha_s$ -coupled ciliary dopamine receptor DRD1 induce elongation of cilia [25], whereas agonists of the G $\alpha_i$ -coupled ciliary DRD2 cause shortening of the cilia [12], suggesting these divergent effects are tied to ciliary production of cyclic adenosine monophosphate (cAMP) [26,27] (Figure 2A).

Alterations in neuronal PCL are also seen in response to diverse physiological changes. Some of the most compelling evidence comes from studies of the hypothalamus, a brain structure containing neuron populations that regulate many fundamental homeostatic and behavioural processes. For example, a recent preprint suggested that neurons in the suprachiasmatic nucleus

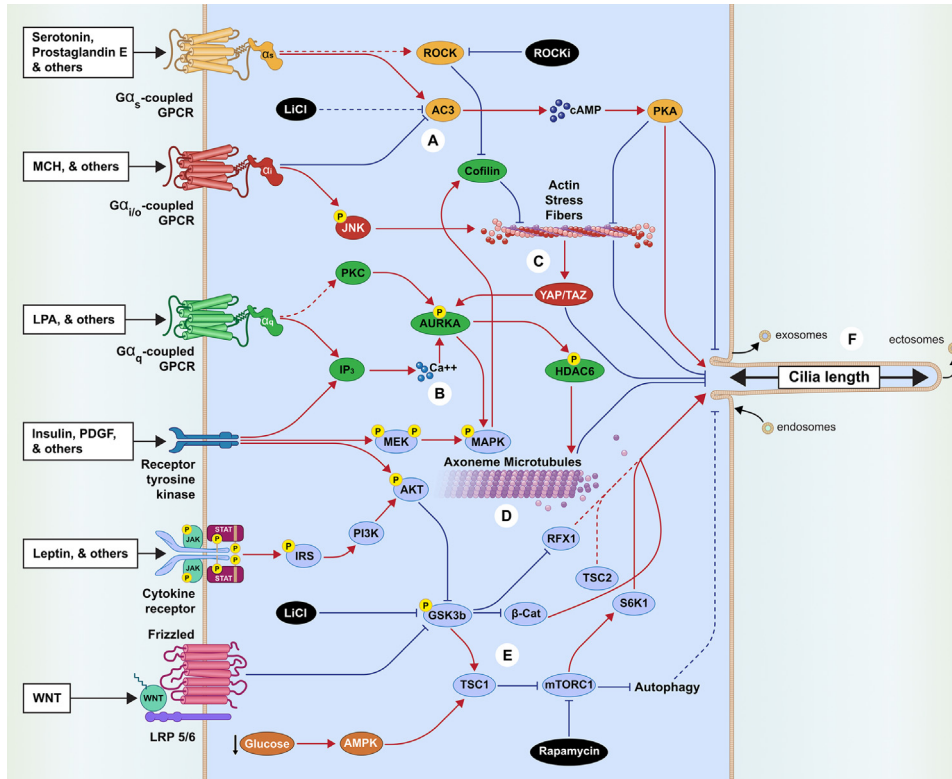
diverse group of disorders can affect a wide range of organ systems, including the eyes, kidneys, brain, and skeletal system.

**Primary cilia length (PCL):** here, we defined PCL as the length of primary cilia from the base to the tip as measured in different studies, largely by fluorescence imaging of primary cilia stained for ciliary proteins or of primary cilia labelled with a reporter gene.

**Rho-associated kinase (ROCK):** a serine/threonine kinase activated by the small GTPase RhoA. ROCK plays a role in regulating of the actin cytoskeleton and therefore acts on PCL.

**Transition zone:** a specific subregion of primary cilia that acts as a gate to selectively permit the entrance of proteins, helping make the cilium a separate cellular compartment.





Trends in Cell Biology

**Figure 2. Overview of the signalling pathways controlling changes in PCL.** This schematic represents a generic cell, showing the receptors at the plasma membrane (or primary cilium) to the left, their downstream signalling pathways within the cell (light blue), and their effects on PCL to the right. The involvement of each indicated pathway may differ depending on the cell type or cellular state, and are explained in detail in the main text. Critical signalling components are indicated in coloured ovals and are coloured by pathway. Black ovals indicate drugs acting on these pathways, yellow circles with a central 'P' indicate protein phosphorylation, and dotted lines indicate controversial or unclear evidence. Abbreviations: AC3, adenylyl cyclase 3; GPCR, G-protein coupled receptor; IP<sub>3</sub>, inositol 1,4,5- trisphosphate; JNK, Jun N-terminal kinase; LiCl, lithium chloride; LPA, lysophosphatidic acid; MCH, melanin concentrating hormone; MEK, mitogen-activated protein kinase kinase; PCL, primary cilia length; PDGF, platelet-derived growth factor; PI3K, phosphatidylinositol 3-kinase; PKA, protein kinase A; RFX1, regulatory factor X; S6K1, ribosomal protein S6 kinase 1; WNT, Wingless-related integration site.

(SCN) of mice, which help regulate circadian rhythms, underwent cyclic changes in PCL, ranging in length from ~1 μm at the onset of the dark cycle to ~6 μm at the onset of the light cycle. The same group also suggested that the rhythmicity of these changes in PCL was disrupted by mutations in the core circadian clock genes, and perturbations in the cilia, in turn, disrupted the robustness of suprachiasmatic nucleus-generated circadian rhythms [28]. Although the functional consequences for the cell's circuitry are still unclear, metabolic signals can also induce significant changes in PCL in hypothalamic neuron populations that regulate food intake [29,30]. Many GPCRs that respond to metabolic signals are localized to the cilium [30], and acute fasting is associated with shorter hypothalamic cilia, whereas cilia lengthen upon refeeding [14]. However, prolonged exposure to a calorie-dense diet is associated with significantly shorter hypothalamic cilia [14]. At the molecular level, PCL is affected by exposure to dietary palmitic acid, the adipocyte-derived hormone leptin, the pancreatic hormone insulin [30,31], and the orexigenic (appetite-promoting) neuropeptide MCH [16]. Specifically, the acute administration of leptin leads to a 50% increase in PCL in leptin-sensitive regions of the hypothalamus, whereas the loss of leptin or its receptor is associated with significantly shorter primary cilia [14]. Similarly, acute insulin treatment increases the length of the cilia in hypothalamic neurons by

up to 25%, probably through activation of the phosphoinositide 3 kinase (PI3K) pathway [14] (Figure 2). Glucose deprivation plays a dual role in regulating PCL by both inducing the formation of cilia to increase the fraction of ciliated cells and, at the same time, inducing ciliary shortening via a mechanism that may involve the activation of AMP-activated protein kinase (AMPK) [32] (Figure 2E). Furthermore, MCHR1 localizes to primary cilia in some appetite-regulatory hypothalamic neurons, where its activation reduces PCL in mice [16,24], potentially by inhibiting production of cAMP [33,34] and promoting **actin stress fibre** production [16] (Figure 2A). More broadly, hypothalamic PCL is affected by posttranslational modifications that reflect the organism's metabolic state such as the O-linked  $\beta$ -*N*-acetylglucosamine (O-GlcNAc) modification of serine and threonine residues. Higher levels of this widespread modification reflect higher internal cellular nutrient levels [35], and inhibition of O-GlcNAcylation is associated with elongated cilia [36]. Indeed, O-GlcNAcylation of ciliary  $\alpha$ -tubulin and HDAC6 induces shortening of the cilia, probably via disassembly of the ciliary microtubules [36] (Figure 2D). Below, we explore the mechanisms affecting neuronal PCL in greater detail.

### Molecular mechanisms affecting neuronal PCL

#### G-protein coupled receptors

The activation of GPCRs and the subsequent changes in the levels of secondary messengers has been shown to affect PCL. For example, ciliary  $G\alpha_s$ -coupled GPCRs such as EP4 [3] and HTR6 [23] can stimulate the production of cAMP, thus promoting elongation of cilia [37], partly by redistributing actin stress fibres [33], whereas the reduction of cAMP levels by  $G\alpha_i$ -coupled GPCRs such as DRD2 [12] and MCHR1 [24] promotes shortening of cilia (Figure 2A). By contrast, increases in the somatic concentration of cAMP can decrease PCL [38]. Lithium chloride promotes the growth of PCL and inhibits **adenylate cyclase 3 (AC3)** [39] but is also likely to act on other signalling pathways [40] (Figure 2A). Similarly, increased ciliary  $Ca^{2+}$  levels (Figure 2B) activate calcium and/or calmodulin (CaM) kinase upstream from Aurora A kinase (AURKA) [41,42] and are associated with reduced PCL [37]. However, calcium-dependent protein kinase (PKC) can also increase PCL via MAPK [43], suggesting that the effects of calcium on PCL depend on its concentration, dynamics, and cellular location, and the cell's state [33].

#### Actin stress fibres and Rho-associated kinases

Actin stress fibres are likely to regulate PCL, since their destabilisation by pharmacological [44,45] or genetic [46] means promotes longer primary cilia (Figure 2C). Mechanistically, CDK10/CycM phosphorylates the RhoA-associated kinase PKN2, which activates **Rho-associated kinase (ROCK)** to promote the expansion of actin stress fibres [46], which, in turn, induces the nuclear translocation of Yes-associated protein/transcriptional coactivator with PDZ-binding motif (YAP/TAZ) [47] to promote the transcription of AURKA and other negative regulators of PCL (Figure 2B) [44,45]. Similarly, the loss of polycystin-1 (PC-1)-mediated repression of ROCK signalling leads to increased actin stress fibres and shortening of cilia through the LIMK-cofilin pathway [48], whereas pharmacological inhibition of ROCK by the small molecule Y-27362 promotes the growth of PCL [44,48] (Figure 2C). Finally, cofilin activation by MAPK promotes the inhibition of stress fibres formation [49] and increases PCL [43] (Figure 2C).

#### Tubulin and the IFT machinery

The microtubules of the axoneme are stable compared with cytoplasmic microtubules [50], and they are most dynamic at the ciliary tip, where they can be readily polymerized or depolymerized (Figure 2D). Increased levels of soluble tubulin, which forms the building blocks of the axoneme, lead to increased cilia length, while stabilisation of the microtubule with the drug taxol leads to shortened or absent cilia [51]. Microtubule stabilisation reduces the pool of soluble tubulin and, therefore, may erode the distal tip of the axoneme [51]. The posttranslational modification of

tubulin by polyglutamylation and detyrosination affects the motor function of Kinesin-1 and -2, which are required to transport tubulin inside the cilium and may therefore contribute to the regulation of PCL by promoting elongation [52–54], since increasing the activity of anterograde IFT separately increases PCL [37]. Indeed, it has been proposed that PCL might be regulated by the amount and frequency of cargos transported by the IFT into the cilium [55]. Finally, the histone deacetylase HDAC6 acts downstream from AURKA to deacetylate ciliary microtubules and to promote disassembly of the axoneme and reduction in PCL [36] (Figure 2D).

#### Cell cycle-associated proteins

Since the ciliary basal body must detach from the plasma membrane to organize the microtubule apparatus during mitosis, ciliogenesis and the cell cycle are tightly connected [1]. Postmitotic neurons continue to express cyclins and other proteins typically associated with cell cycle regulation, suggesting that they continue to play a role in regulating PCL. For example, the master cell cycle regulatory complex APC–Cdc20 is required for the normal morphogenesis of dendrites [56]. Cyclin E negatively regulates CDK5 in postmitotic neurons [57] and has been associated with reduced PCL in other cell types [58]. Furthermore, accumulation of the centriolar E3 ligase FBXO41 promotes disassembly of cilia in both mitotic and postmitotic cells in a manner that requires cytoskeletal rearrangements of actin [15].

#### Autophagy and mTOR

Autophagy may regulate PCL, since acute serum starvation is a common experimental tool to induce ciliation *in vitro* [59]. By contrast, sustained serum starvation has been shown to downregulate IFT20 and reduce PCL [59] (Figure 2E), whereas the sustained inhibition of autophagy with 3-methyladenine [60] or through disruption of the positive regulators of autophagy (ATG5 and RPGRIP1L) is associated with increased PCL [60]. While serum starvation arrests cell division, and ciliogenesis is coupled to the cell cycle [61], autophagy may be independent of the cell cycle since its chemical or genetic inhibition in hypothalamic neurons is sufficient to reduce PCL [31] (Figure 2E). Furthermore, diseases associated with defects in autophagy, such as focal cortical dyslamination and thyroid Hürthle cell tumours, are characterized by an abnormal PCL [62]. Activation of the autophagy-related serine/threonine protein kinase mTORC1 increases PCL by promoting protein synthesis related to primary cilia assembly [32,63], and its inhibition by the drug rapamycin [64] or by Tuberous Sclerosis Complex 1 (TSC1) leads to reduction in PCL [63] (Figure 2E). Indeed, mTORC1's activity is downregulated under nutrient-deprived conditions through a signalling cascade involving the activation of AMP-activated protein kinase (AMPK) and TSC1, which promote elongation of cilia when ablated [32] (Figure 2E). The mTOR pathway is also the target of Glycogen synthase kinase-3 beta (GSK3 $\beta$ ), which is inhibited by phosphoinositide 3 kinase (PI3K)/AKT signalling downstream from diverse ligands, including the metabolic hormones leptin and insulin [14], which are associated with regulating PCL [65]. Specifically, the inhibition of GSK3 $\beta$  increases  $\beta$ -catenin stability and leads to the growth of PCL [66,67], potentially explaining the effects of lithium chloride on elongation of cilia [40,68]. Inhibition of GSK3 $\beta$  also allows the transcription factor RFX1 to promote anterograde IFT gene expression [66] and relieves the repression of mTORC1 by TSC1 [64] (Figure 2E).

#### Ciliary membrane vesicles

The cilium's length may also be regulated by the incorporation or shedding of membrane-bound vesicles by exocytosis, endocytosis, or ectocytosis. Ciliary exosomes resulting from the exocytosis of multivesicular bodies (MVBs) are thought to be released from the base of cilia where endosomes can also be absorbed, whereas **ectosomes** are produced from the ciliary membrane at the axonemal tip [2,68–71] (Figure 2F). It is unclear whether these processes operate

in concert to homeostatically maintain the protein content of the cilium to optimize signalling, or whether they can be engaged independently to alter PCL. Ectosomes contain a fivefold higher concentration of GPCRs than the ciliary membrane [72] and may act as a ‘safety valve’ to rapidly eliminate activated ciliary GPCRs when their retrieval via retrograde IFT fails or is too slow [72]. Although ciliary shedding of ectosomes was first appreciated in cells with defective GPCR retrieval, this phenomenon appears to be active in normal primary cilia [72] and may therefore actively reduce PCL by removing portions of the ciliary membrane. For example, prominin-1-enriched cilia-derived membrane vesicles are associated with shorter primary cilia in neuroepithelial cells, suggesting that their shedding is a mechanism of regulating PCL [73], and cilia from *bbs8* mutant neurons in *Caenorhabditis elegans* display variable length in a manner potentially related to altered ectocytosis [74].

### Potential consequences of changes in PCL

While changes in PCL have been observed in mammalian neurons in response to physiological signals, the consequences of these changes have largely been unexplored. We propose that if the primary cilium acts as a sensory antenna [3,18], changes in the length of this antenna could affect its function as an organelle by modifying its ability to sense and transduce extracellular signals, which, in turn, could affect cellular function and tissue- or organism-level behaviour. Below, we propose four distinct but not mutually exclusive models (Figure 3) of the potential functional consequences of changes in PCL. We expect those to vary by ligand, developmental time point, and cell type.

#### (i) Passive indicator model

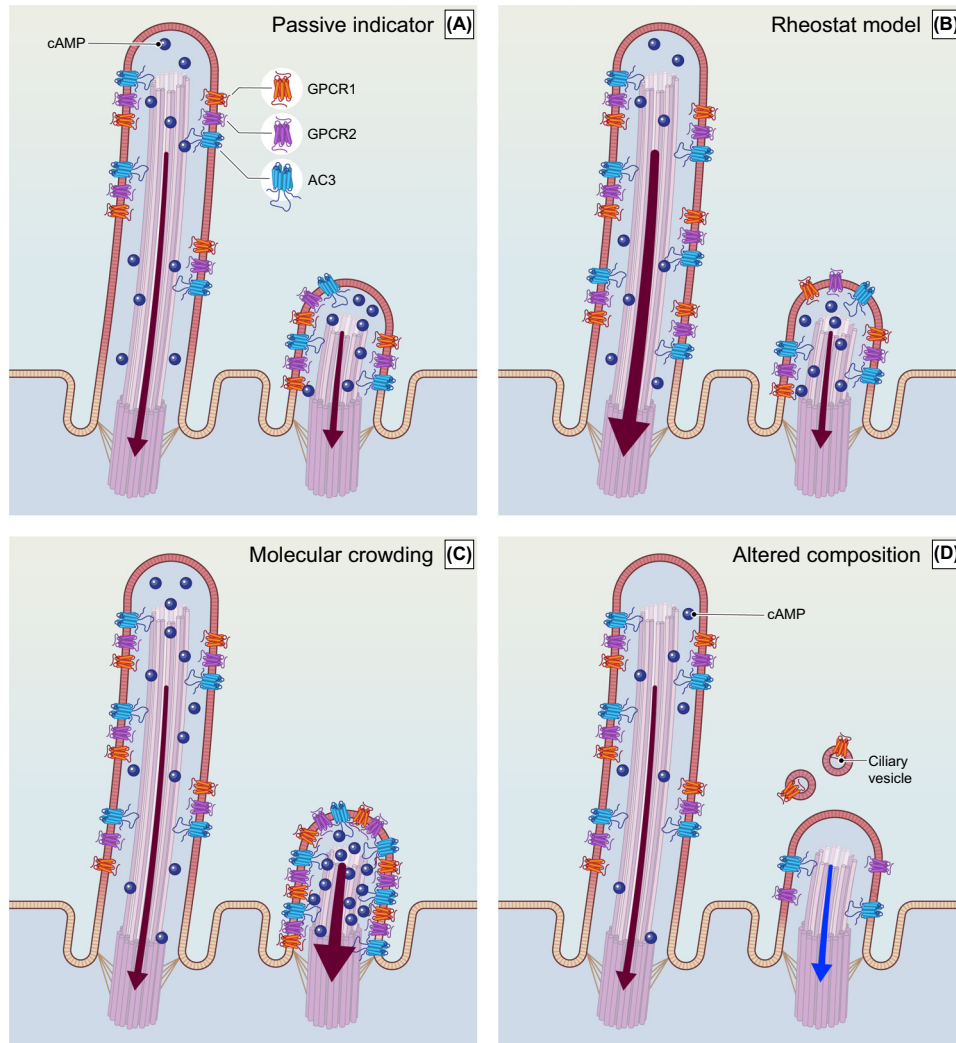
Although the primary cilium plays an important role in signalling [18], changes in its length may simply indicate a change in the cell’s state or an echo of past ciliary signalling that does not alter its current or future signalling capacity (Figure 3A). For example, ectosome shedding from the ciliary tip to fine-tune receptor signalling might lead to a reduction in PCL only as a secondary consequence [72]. Furthermore, changes in PCL induced by MAPK, PP-1, and cofilin were not accompanied by obvious changes in ciliary function, as determined by measuring the concentration of cytosolic calcium induced by fluid-shear stress in endothelial cells [43], although we cannot exclude the possibility that functional changes could have been detected with a suitably relevant and sensitive assay. Similarly, experimental manipulations that increased steady-state PCL *in vivo* did not interfere with normal development or behaviour in mice [68,75].

#### (ii) Rheostat model

A longer ‘sensory antenna’ with a correspondingly larger surface area could increase the number of receptors and downstream signalling mediators without altering their concentration (Figure 3B), so that changes in PCL alter the cilia’s sensitivity to signals, much like a rheostat (variable resistor) regulates the flow of current in an electrical circuit. The rheostat model assumes that the mechanisms regulating proteins’ entry and exit from the primary cilium operate at least as quickly as the observed changes in PCL. This concept is supported by computational models suggesting that the capture rate of molecules by the ciliary receptors scales linearly with PCL [76]. Indeed, the exogenous expression of ciliary proteins has been associated with elongation of cilia in the neocortical neurons of mice [17], suggesting that PCL may change in response to changes in the receptors’ entry or activity to normalize their concentration. We note that despite the constant concentration of receptors and downstream targets (e.g., AC3) imagined in the rheostat model, the concentration of secondary messengers (e.g., cAMP) may still build up to higher levels in the small volume of a longer cilium before they reach the ‘sink’ of the cytoplasm.



### Consequences of changes in cilia length



Trends in Cell Biology

**Figure 3. Hypothetical models of the functional consequences of changes in PCL.** While many different receptor types and signalling pathways are active in primary cilia, for clarity we schematise ciliary signalling as the interaction of GPCR with AC3 and the downstream production of cAMP as a secondary messenger, with the thickness of the arrow representing the efficacy of signal transduction. (A) The passive indicator model proposes that changes in PCL do not alter ciliary signalling capacity. (B) In the rheostat model, increased PCL potentiates ciliary signalling capacity due to the increased number of ciliary proteins. (C) The local concentration model proposes that if the total number of ciliary proteins is not substantially changed, a shorter cilium would increase their concentration to potentiate or alter ciliary signalling. (D) In the altered composition model, PCL influences the composition and variety of receptors to alter the signalling capacity of the primary cilium. Abbreviations: AC3, adenylate cyclase 3; cAMP, cyclic AMP; GPCR, G-protein coupled receptor; PCL, primary cilia length.

Experimentally, ciliary SHH signalling was potentiated in mouse embryonic fibroblasts with primary cilia that were lengthened either genetically by the ablation of *Cnks2a* [77], or pharmacologically by the inhibitor of actin polymerisation Cytochalasin B [78]. In the hypothalamus, alterations in PCL might modulate the intensity of signalling from metabolic factors [30] since leptin- and insulin-induced cilia elongation has been proposed to increase neuronal sensitivity

to other metabolic signals [14], leading to increased activation of anorexigenic proopiomelanocortin (POMC) neurons and decreased food intake [22].

#### (iii) Local concentration model

There are a number of ways in which the local concentration of ciliary proteins could be altered. For example, if the rate of protein trafficking into and out of the cilium is slow relative to the rate of change in PCL, then the total number of ciliary membrane proteins would remain relatively constant, increasing their concentration and potentially increasing the number of intermolecular interactions as cilia shrink. This increased local concentration of ciliary proteins could potentiate or alter their signalling in at least three ways: (i) increasing the likelihood that receptors such as GPCRs will interact with downstream targets, (ii) increasing the concentration of second messengers due to the decreased ciliary volume, and (iii) increasing the formation of receptor homomers, heteromers, and higher-order oligomers which could, in turn, shape their responses to ligands [79]. Indeed, longer primary cilia in senescent fibroblasts have been suggested to reduce their capacity to respond to SHH by diluting their receptors and downstream targets [80]. Genetic or pharmacological induction of multiple cilia in the same cell diluted the ciliary proteins and attenuated ciliary signalling, since the multiple primary cilia occupy the same ciliary pocket, and thus the trafficking machinery becomes rate-limiting [81]. Furthermore, the appetite-regulating ciliary GPCRs MCHR1 and SSTR3 heterodimerize [79], and the receptor for the appetite-promoting hormone ghrelin (GSHR) heterodimerizes with multiple other GPCRs including the melanocortin 3 receptor (MC3R) [82], suggesting that changes in PCL that affect local protein concentration could alter neuronal responses to metabolic signals. While it is unclear how ciliary receptors interact with each other within the cilium, they may do so in lipid rafts, which can accumulate at the base of the cilium and have been shown to propagate adipogenic insulin signalling after IR/IGF1R heterodimerization [83].

#### (iv) Altered composition model

Mechanisms that might alter PCL such as ectosome shedding could alter the composition of ciliary proteins and lipids to influence the subsequent signalling, as observed when GPR161 was removed upon SHH binding [72]. The ciliary composition could also be changed by other mechanisms, for example, if longer cilia more effectively accommodate lowly expressed or inefficiently targeted proteins than shorter cilia, in which case ciliary elongation would lead to the presentation of a greater diversity of ciliary proteins.

### Concluding remarks

We have summarized the mechanisms potentially involved in the control of PCL in mammalian neurons and proposed models of how these changes could affect ciliary signalling and, consequently, alter the cell's state. Here, we propose a set of experiments to determine which model (or combination of models) might explain how PCL could affect signalling pathways, keeping in mind that results from one experimental system may not apply universally (see [Outstanding questions](#)). For example, one could use cilia-targeted cAMP or calcium sensors [84] to test how experimentally shortening or lengthening primary cilia might alter ciliary signalling in response to a panel of ligands that act via ciliary GPCRs. One could also use techniques such as somatic calcium imaging or analyses of transcriptional changes to test how changes in PCL might alter the cells' state. One challenge to this approach is that manipulations that affect PCL may also affect the function of cilia or other compartments, so it would be necessary to validate the findings from one approach using an orthogonal method. Ciliary proximity proteomics could be used to experimentally test how physiologically induced changes in PCL alter ciliary protein composition [85,86]. These experiments would provide rich datasets, but are costly, relatively low-throughput, and have yet to be optimized

### Outstanding questions

How is the length of primary cilia regulated and do these mechanisms differ between mitotic and postmitotic cells?

How do changes in PCL affect cell signalling and behaviour, particularly in the neurons?

Is the complement of ciliary proteins affected by changes in length? If yes, how does this affect cell signalling?

Which signalling pathways are most sensitive to changes in PCL?

Are the mechanisms regulating PCL cell-type-specific or universal?

Can the mechanisms that regulate PCL be therapeutically harnessed to treat ciliopathies and diseases accompanied by changes in PCL?

for neurons. We also note that the molecular mechanisms and models described above have largely been informed by studies in postmitotic mammalian neurons, in which cilia may change in length but otherwise remain relatively stable. However, we propose that the mechanisms and consequences of changes in PCL may be more broadly relevant, since primary cilia in most cell types participate in signalling. In embryonic development, changes in PCL might tune cellular responsiveness, and shape the timing and outcomes of differentiation and tissue morphogenesis. For example, FGF signalling affects the length and function of cilia during development, and developmental defects related to FGF signalling could partly derive from ciliary dysfunction [87]. Changes in PCL are likely to be relevant to disease biology, since the rescue of cilia length attenuates disease phenotypes and mitigates SHH signalling defects [88,89]. Finally, alterations in PCL are a cellular phenotype that could provide a foundation for image-based clustered regularly interspaced short palindromic repeats (CRISPR) screens, in which the PCL in a given cell is correlated with the identity of the perturbed gene [90]. Such studies could catalogue the genes involved in ciliary formation, function, and length control, as well as revealing a ciliary role for genes associated with human disease. Understanding the causes and consequences of changes in PCL could significantly advance our understanding of primary cilia in development and in the function of adult cells and tissues, and perhaps may also reveal a new class of therapeutic targets for diseases that involve signalling pathways mediated by primary cilia.

### Acknowledgments

We are grateful to Dr David Mick at the Universität des Saarlandes and Dr Jiani Gou at the University of Calgary for their helpful comments in the preparation of this manuscript. We are deeply thankful to Andreina Esposito for her strength, support, and guidance that helped make this work possible. The figures were produced with the assistance of Maria Romanova at Molecular House. F. T.M. is a New York Stem Cell Foundation – Robertson Investigator (NYSRF-R-156) and is supported by the Wellcome Trust and the Royal Society (211221/Z/18/Z). V.M. is supported by the Sackler Trust, the Medical Research Council Doctoral Training Programme, and the School of Clinical Medicine Cambridge Trust Scholarship. For the purpose of open access, the authors have applied a CC-BY public copyright license to any author-accepted manuscript version arising from this submission.

### Declaration of interests

The authors declare no competing interests.

### References

- Wang, L. and Dynlacht, B. (2018) The regulation of cilium assembly and disassembly in development and disease. *Development* 145, dev151407
- Nachury, M.V. and Mick, D.U. (2019) Establishing and regulating the composition of cilia for signal transduction. *Nat. Rev. Mol. Cell Biol.* 20, 389–405
- Schou, K.B. *et al.* (2015) Ins and outs of GPCR signaling in primary cilia. *EMBO Rep.* 16, 1099–1113
- Ye, F. *et al.* (2018) BBSome trains remove activated GPCRs from cilia by enabling passage through the transition zone. *J. Cell Biol.* 217, 1847–1868
- Badgandi, H.B. *et al.* (2017) Tubby family proteins are adapters for ciliary trafficking of integral membrane proteins. *J. Cell Biol.* 216, 743–760
- Berbari, N.F. *et al.* (2008) Bardet–Biedl syndrome proteins are required for the localization of G protein-coupled receptors to primary cilia. *Proc. Natl. Acad. Sci. U. S. A.* 105, 4242–4246
- Wingfield, J.L. *et al.* (2018) Trafficking of ciliary membrane proteins by the intraflagellar transport/BBSome machinery. *Essays Biochem.* 62, 753–763
- Reiter, J.F. and Leroux, M.R. (2017) Genes and molecular pathways underpinning ciliopathies. *Nat. Rev. Mol. Cell Biol.* 18, 533–547
- Green, J.A. and Mykytyn, K. (2014) Neuronal primary cilia: an underappreciated signaling and sensory organelle in the brain. *Neuropsychopharmacology* 39, 244–245
- Guadiana, S.M. *et al.* (2016) Type 3 adenylyl cyclase and somatostatin receptor 3 expression persists in aged rat neocortical and hippocampal neuronal cilia. *Front. Aging Neurosci.* 8, 127
- Hu, L. *et al.* (2017) Serotonin 5-HT6 receptors affect cognition in a mouse model of Alzheimer’s disease by regulating cilia function. *Alzheimers Res. Ther.* 9, 76
- Miyoshi, K. *et al.* (2014) Lack of dopaminergic inputs elongates the primary cilia of striatal neurons. *PLoS One* 9, e97918
- Chakravarthy, B. *et al.* (2012) Reduction of the immunostainable length of the hippocampal dentate granule cells’ primary cilia in 3xAD-transgenic mice producing human Aβ1-42 and tau. *Biochem. Biophys. Res. Commun.* 427, 218–222
- Han, Y.M. *et al.* (2014) Leptin-promoted cilia assembly is critical for normal energy balance. *J. Clin. Invest.* 124, 2193–2197
- King, C.R. *et al.* (2019) Fbxo41 promotes disassembly of neuronal primary cilia. *Sci. Rep.* 9, 8179
- Kobayashi, Y. *et al.* (2021) Ciliary GPCR-based transcriptome as a key regulator of cilia length control. *FASEB BioAdv.* 3, 744–767
- Guadiana, S.M. *et al.* (2013) Arborization of dendrites by developing neocortical neurons is dependent on primary cilia and Type 3 adenylyl cyclase. *J. Neurosci.* 33, 2626–2638
- Anvarian, Z. *et al.* (2019) Cellular signalling by primary cilia in development, organ function and disease. *Nat. Rev. Nephrol.* 15, 199–219
- Insinna, C. and Besharse, J.C. (2008) Intraflagellar transport and the sensory outer segment of vertebrate photoreceptors. *Dev. Dyn. Off. Publ. Am. Assoc. Anat.* 237, 1982–1992
- Loktev, A.V. and Jackson, P.K. (2013) Neuropeptide Y family receptors traffic via the Bardet–Biedl syndrome pathway to signal in neuronal primary cilia. *Cell Rep.* 5, 1316–1329
- Gigante, E.D. and Caspary, T. (2020) Signaling in the primary cilium through the lens of the Hedgehog pathway. *WIREs Dev. Biol.* 9, e377


22. Sun, J.S. *et al.* (2021) Ventromedial hypothalamic primary cilia control energy and skeletal homeostasis. *J. Clin. Invest.* 131, 138107
23. Brodsky, M. *et al.* (2017) 5-HT6 receptor blockade regulates primary cilia morphology in striatal neurons. *Brain Res.* 1660, 10–19
24. Alhassen, W. *et al.* (2022) Regulation of brain primary cilia length by MCH signaling: evidence from pharmacological, genetic, optogenetic, and chemogenic manipulations. *Mol. Neurobiol.* 59, 245–265
25. Spasic, M. and Jacobs, C.R. (2017) Lengthening primary cilia enhances cellular mechanosensitivity. *Eur. Cell. Mater.* 33, 158–168
26. Bhatia, A. *et al.* (2022) *Biochemistry, Dopamine Receptors, StatPearls*
27. Marley, A. *et al.* (2013) GPR88 reveals a discrete function of primary cilia as selective insulators of GPCR cross-talk. *PLoS One* 8, e70857
28. Tu, H.-Q. *et al.* (2022) Rhythmic cilium in SCN neuron is a gatekeeper for the intrinsic circadian clock. *bioRxiv* Published online February 1, 2022. <https://doi.org/10.1101/2022.01.26.477948>
29. Schwartz, M.W. *et al.* (2000) Central nervous system control of food intake. *Nature* 404, 661–671
30. Oh, E.C. *et al.* (2015) Metabolic regulation and energy homeostasis through the primary cilium. *Cell Metab.* 21, 2121–2131
31. Ávalos, Y. *et al.* (2022) Palmitic acid control of ciliogenesis modulates insulin signaling in hypothalamic neurons through an autophagy-dependent mechanism. *Cell Death Dis.* 13, 659
32. Takahashi, K. *et al.* (2018) Glucose deprivation induces primary cilium formation through mTORC1 inactivation. *J. Cell Sci.* 131, jcs208769
33. Avasthi, P. and Marshall, W.F. (2012) Stages of ciliogenesis and regulation of ciliary length. *Differ. Res. Biol. Divers.* 83, S30–S42
34. Hawes, B.E. *et al.* (2000) The melanin-concentrating hormone receptor couples to multiple G proteins to activate diverse intracellular signaling pathways. *Endocrinology* 141, 4524–4532
35. Ruan, H.-B. *et al.* (2013) Cracking the O-GlcNAc code in metabolism. *Trends Endocrinol. Metab.* 24, 301–309
36. Tian, J.L. and Qin, H. (2019) O-GlcNAcylation regulates primary ciliary length by promoting microtubule disassembly. *iScience* 12, 379–391
37. Besschetnova, T.Y. *et al.* (2010) Identification of signaling pathways regulating primary cilium length and flow-mediated adaptation. *Curr. Biol. CB* 20, 182–187
38. Hansen, J.N. *et al.* (2020) Nanobody-directed targeting of optogenetic tools to study signaling in the primary cilium. *eLife* 9, e57907
39. Ou, Y. *et al.* (2009) Adenylate cyclase regulates elongation of mammalian primary cilia. *Exp. Cell Res.* 315, 2802–2817
40. Miyoshi, K. *et al.* (2009) Lithium treatment elongates primary cilia in the mouse brain and in cultured cells. *Biochem. Biophys. Res. Commun.* 388, 757–762
41. Hu, H.-B. *et al.* (2021) LPA signaling acts as a cell-extrinsic mechanism to initiate cilia disassembly and promote neurogenesis. *Nat. Commun.* 12, 662
42. Pugacheva, E.N. *et al.* (2007) HEF1-dependent Aurora A activation induces disassembly of the primary cilium. *Cell* 129, 1351–1363
43. Abdul-Majeed, S. *et al.* (2012) Mechanisms regulating cilia growth and cilia function in endothelial cells. *Cell. Mol. Life Sci.* 69, 165–173
44. Kim, J. *et al.* (2015) Actin remodelling factors control ciliogenesis by regulating YAP/TAZ activity and vesicle trafficking. *Nat. Commun.* 6, 6781
45. Smith, C.E.L. *et al.* (2020) Primary cilia, ciliogenesis and the actin cytoskeleton: a little less resorption, a little more actin please. *Front. Cell Dev. Biol.* 8, 622822
46. Guen, V.J. *et al.* (2018) A homozygous deleterious mutation in CDK10 is associated with agenesis of corpus callosum, retinopathy and deafness. *Am. J. Med. Genet. A* 176, 92–98
47. Seo, J. and Kim, J. (2018) Regulation of Hippo signaling by actin remodeling. *BMB Rep.* 51, 151–156
48. Streets, A.J. *et al.* (2020) Polycystin-1 regulates ARHGAP35-dependent centrosomal RhoA activation and ROCK signaling. *JCI Insight* 5, e135385
49. Oleinik, N.V. *et al.* (2010) ALDH1L1 inhibits cell motility via dephosphorylation of cofilin by PP1 and PP2A. *Oncogene* 29, 6233–6244
50. Orbach, R. and Howard, J. (2019) The dynamic and structural properties of axonemal tubulins support the high length stability of cilia. *Nat. Commun.* 10, 1838
51. Sharma, N. *et al.* (2011) Soluble levels of cytosolic tubulin regulate ciliary length control. *Mol. Biol. Cell* 22, 806–816
52. Gadadhar, S. *et al.* (2017) Tubulin glycylation controls primary cilia length. *J. Cell Biol.* 216, 2701–2713
53. Wloga, D. *et al.* (2017) Posttranslational modifications of tubulin and cilia. *Cold Spring Harb. Perspect. Biol.* 9, a028159
54. He, K. *et al.* (2020) The emerging role of tubulin posttranslational modifications in cilia and ciliopathies. *Biophys. Rep.* 6, 89–104
55. Wren, K.N. *et al.* (2013) A differential cargo-loading model of ciliary length regulation by IFT. *Curr. Biol.* 23, 2463–2471
56. Kim, A.H. *et al.* (2009) A centrosomal Cdc20-APC pathway controls dendrite morphogenesis in postmitotic neurons. *Cell* 136, 322–336
57. Odajima, J. *et al.* (2011) Cyclin E constrains Cdk5 activity to regulate synaptic plasticity and memory formation. *Dev. Cell* 18, 655–668
58. Maskey, D. *et al.* (2015) Cell cycle-dependent ubiquitylation and destruction of NDE 1 by CDK 5- FBW 7 regulates ciliary length. *EMBO J.* 34, 2424–2440
59. Pampliega, O. and Cuervo, A.M. (2016) Autophagy and primary cilia: dual interplay. *Curr. Opin. Cell Biol.* 39, 1–7
60. Struchtrup, A. *et al.* (2018) The ciliary protein RPGRI1L governs autophagy independently of its proteasome-regulating function at the ciliary base in mouse embryonic fibroblasts. *Autophagy* 14, 567–583
61. Goto, H. *et al.* (2017) Mechanisms of ciliogenesis suppression in dividing cells. *Cell. Mol. Life Sci.* 74, 881–890
62. Morleo, M. and Franco, B. (2019) The autophagy–cilia axis: an intricate relationship. *Cells* 8, 905
63. Rosengren, T. *et al.* (2018) TSC1 and TSC2 regulate cilia length and canonical Hedgehog signaling via different mechanisms. *Cell. Mol. Life Sci. CMLS* 75, 2663–2680
64. Yuan, S. *et al.* (2012) Target-of-rapamycin complex 1 (Torc1) signaling modulates cilia size and function through protein synthesis regulation. *Proc. Natl. Acad. Sci.* 109, 2021–2026
65. Ballou, L.M. and Lin, R.Z. (2008) Rapamycin and mTOR kinase inhibitors. *J. Chem. Biol.* 1, 27–36
66. Kang, G.M. *et al.* (2015) Leptin elongates hypothalamic neuronal cilia via transcriptional regulation and actin destabilization. *J. Biol. Chem.* 290, 18146–18155
67. Kim, J.-G. *et al.* (2017) Wnt3A induces GSK-3 $\beta$  phosphorylation and  $\beta$ -catenin accumulation through RhoA/ROCK. *J. Cell. Physiol.* 232, 1104–1113
68. Ford, M.J. *et al.* (2018) A cell/cilia cycle biosensor for single-cell kinetics reveals persistence of cilia after G1/S transition is a general property in cells and mice. *Dev. Cell* 47, 509–523.e5
69. Wood, C. *et al.* (2013) The cilium secretes bioactive ectosomes. *Curr. Biol. CB* 23, 906–911
70. Wang, J. *et al.* (2014) *C. elegans* ciliated sensory neurons release extracellular vesicles that function in animal communication. *Curr. Biol. CB* 24, 519–525
71. Cocucci, E. and Meldolesi, J. (2015) Ectosomes and exosomes: shedding the confusion between extracellular vesicles. *Trends Cell Biol.* 25, 364–372
72. Nager, A.R. *et al.* (2017) An actin network dispatches ciliary GPCRs into extracellular vesicles to modulate signaling. *Cell* 168, 252–263.e14
73. Dubreuil, V. *et al.* (2007) Midbody and primary cilium of neural progenitors release extracellular membrane particles enriched in the stem cell marker prominin-1. *J. Cell Biol.* 176, 483–495
74. Razzauti, A. and Laurent, P. (2021) Ectocytosis prevents accumulation of ciliary cargo in *C. elegans* sensory neurons. *eLife* 10, e67670
75. Bangs, F.K. *et al.* (2015) Lineage specificity of primary cilia in the mouse embryo. *Nat. Cell Biol.* 17, 113–122
76. Hickey, D. *et al.* (2021) Ciliary chemosensitivity is enhanced by cilium geometry and motility. *eLife* 10, e66322
77. Loukil, A. *et al.* (2021) A complex of distal appendage-associated kinases linked to human disease regulates ciliary trafficking and stability. *Proc. Natl. Acad. Sci.* 118, e2018740118
78. Drummond, M.L. *et al.* (2018) Actin polymerization controls cilia-mediated signaling. *J. Cell Biol.* 217, 3255–3266



79. Green, J.A. *et al.* (2012) Heteromerization of ciliary G protein-coupled receptors in the mouse brain. *PLoS One* 7, e46304
80. Breslin, L. *et al.* (2014) Ciliary abnormalities in senescent human fibroblasts impair proliferative capacity. *Cell Cycle* 13, 2773–2779
81. Mahjoub, M.R. and Stearns, T. (2012) Supernumerary centrosomes nucleate extra cilia and compromise primary cilium signaling. *Curr. Biol.* 22, 1628–1634
82. Rediger, A. *et al.* (2011) Mutually opposite signal modulation by hypothalamic heterodimerization of ghrelin and melanocortin-3 receptors. *J. Biol. Chem.* 286, 39623–39631
83. Yamakawa, D. *et al.* (2021) Primary cilia-dependent lipid raft/caveolin dynamics regulate adipogenesis. *Cell Rep.* 34, 108817
84. Jiang, J.Y. *et al.* (2019) Direct visualization of cAMP signaling in primary cilia reveals up-regulation of ciliary GPCR activity following Hedgehog activation. *Proc. Natl. Acad. Sci.* 116, 12066–12071
85. Cho, K.F. *et al.* (2020) Proximity labeling in mammalian cells with TurboID and split-TurboID. *Nat. Protoc.* 15, 3971–3999
86. May, E.A. *et al.* (2021) Time-resolved proteomics profiling of the ciliary Hedgehog response. *J. Cell Biol.* 220, e202007207
87. Neugebauer, J.M. *et al.* (2009) FGF signalling during embryo development regulates cilia length in diverse epithelia. *Nature* 458, 651–654
88. Huangfu, D. *et al.* (2003) Hedgehog signalling in the mouse requires intraflagellar transport proteins. *Nature* 426, 83–87
89. Husson, H. *et al.* (2020) Correction of cilia structure and function alleviates multi-organ pathology in Bardet-Biedl syndrome mice. *Hum. Mol. Genet.* 29, 2508–2522
90. Kanfer, G. *et al.* (2021) Image-based pooled whole-genome CRISPRi screening for subcellular phenotypes. *J. Cell Biol.* 220, e202006180

Review

# Viticultural Manipulation and New Technologies to Address Environmental Challenges Caused by Climate Change

Qun Sun <sup>1</sup>, Gabriel Granco <sup>2</sup> , Leah Groves <sup>1</sup>, Jully Voong <sup>2</sup> and Sonet Van Zyl <sup>1,\*</sup>

<sup>1</sup> Department of Viticulture and Enology, California State University, Fresno, 2360 E. Barstow, Avenue, Fresno, CA 93740, USA; qsun@mail.fresnostate.edu (Q.S.); groves\_leah@mail.fresnostate.edu (L.G.)

<sup>2</sup> Department of Geography and Anthropology, California State Polytechnic University, Pomona, CA 91768, USA; ggranco@cpp.edu (G.G.); jullyvoong@cpp.edu (J.V.)

\* Correspondence: svanzyl@mail.fresnostate.edu

**Abstract:** Climate change is a critical challenge for the global grape and wine industry, as it can disrupt grapevine growth, production, and wine quality. Climate change could influence the cost-effectiveness and growth of the wine industry in different wine regions since grapevine development is deeply dependent on weather (short-term) and climate (long-term) conditions. Innovation and new technologies are needed to meet the challenge. This review article addresses the impact of climate change on grapevines, such as vine phenology, pest and disease pressure, crop load, and grape and wine composition. It also reviews recent advances in the areas of viticultural manipulation and relevant technologies to potentially reduce the impact of climate change and help growers improve grape quality. Remote sensing is used for vineyard microclimate monitoring; thermal sensors combined with UAVs, aircraft, or satellites are used for water management; soil electrical conductivity sensors have been developed for soil mapping. Viticultural manipulations, such as regulated deficit irrigation for water use efficiency and berry-ripening delay for growing quality fruit, are also discussed. The review assesses future directions for further technological development, such as soil and vine water monitoring devices, precision viticulture, and artificial intelligence in vineyards.

**Keywords:** climate change; viticultural manipulation; grapevines; wines; new technology; adaption



**Citation:** Sun, Q.; Granco, G.; Groves, L.; Voong, J.; Van Zyl, S. Viticultural Manipulation and New Technologies to Address Environmental Challenges Caused by Climate Change. *Climate* **2023**, *11*, 83. <https://doi.org/10.3390/cli11040083>

Academic Editor: Nir Y. Krakauer

Received: 9 March 2023

Revised: 31 March 2023

Accepted: 1 April 2023

Published: 6 April 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

“Climate change refers to a change in the state of the climate that can be identified (e.g., by using statistical tests) by changes in the mean and/or the variability of its properties and that persists for an extended period, typically decades or longer” [1]. The global average temperature change started in 1880 [2]. According to the latest U.S. Climate Normals 2021 [3], the average temperature in 1981–2010 for the contiguous USA was 52.8 °F, and in 1991–2020, it was 53.3 °F. Carbon dioxide (CO<sub>2</sub>) emissions from human activities are the main drive toward global climate change [4]. As the CO<sub>2</sub> concentration is predicted to continue rising, Earth surface temperature will keep increasing [5].

Changing climate can lead to agricultural change. Severe warming due to climate change has led to increased drought in agricultural production areas and, thus, negatively affects plant water availability and crop production and quality. For example, based on a new released U.S. Drought Monitor, some areas of Kansas and Texas reached exceptional drought levels, which reduced surface water, dried out soils and vegetation, and altered the timing of water availability [6]. The grape and wine industry are no exception to the climate change impacts. For instance, grape growing relies heavily on the length of the growing season, and environmental temperatures and changes in the climate affect these factors [7]. Global warming is expected to change the temperature range in most viticultural areas and will directly influence wine quality. Thus, there is a need to investigate the multiple ways climate change has modified the grape and wine industry.

This review article focuses on the impact of climate change on grapevines, such as phenology, pest and disease pressure, yield, and grape and wine composition. The recent advances in the areas of viticultural manipulation and relevant technologies are also reviewed. This review article contributes to the literature on supporting grape growers and winemakers to better understand the challenges presented by climate change issues. It will also provide them with adaptation options to face the challenges and improve grape and wine quality.

## 2. Overview of Climate Change Models for Wine-Producing Regions

Climate change constitutes a double-edged sword for wine regions. The warming temperature in cool climate areas improves grape and wine quality. For example, with temperature increases of 2.3–5.3 °C in the northern European regions [8], the extended frost-free growing seasons will make it possible to grow warmer-climate varieties and improve current fruit production and quality [8,9]. An empirical study in Moldova also found that both land size and labor have positive significant influences on the productivity of wineries. This finding could be an inspiration for other countries that have similar potential in rural areas to increase wine production [10]. In addition, cool climate regions in North America, such as Oregon and Washington State (USA) and British Columbia (Canada), would also be able to grow new varieties [11]. Grape production in Michigan in the USA has increased due to its shift in climate; the warmer temperatures have been canceling out early season frost. Michigan's future climate is more likely to be suitable to growing all varieties of grapes, including grape varieties that are climatically sensitive to cold temperatures [12,13]. Climate change will also cause grape-growing regions to be more suitable for growing at higher elevations [7].

However, other warm-climate wine regions will be negatively affected by high temperatures and water stress due to the reduction in the amount of water available or irregular precipitation. It is predicted that temperatures will increase by 0.42 °C per decade, which will make it challenging for some wine regions to produce high quality grapes and wines [14]. Regions with Mediterranean climatic conditions are the largest wine regions in the world owing it to long growing seasons with dry/hot summers and mild/wet winters [15]. The regions are located on the western sides of continents between 30° and 40° latitude and will likely face the most severe impacts from climate change [16]. Other wine regions, such as southern Europe, are becoming too hot to produce premium wine since they have already reached, or even exceeded, the optimum range of growing conditions for the currently cultivated grape varieties. Australian wine regions have reported early ripening of Chardonnay, Shiraz, and Cabernet Sauvignon due to increased temperatures [17]. California's climate is one of the most challenged in North America. During the last decade, California has faced several of the most extreme climate and environmental conditions, such as severe drought and wildfires [18].

In our study, ten wine regions were selected as representative of the main producing regions and of the diversity of climatic conditions encountered in wine production. The average annual temperature was used for the climatic conditions and climate change conditions. The data analysis consisted of calculating the mean temperatures for each region under a baseline condition and two climate change periods (mid-century = 2040–2060 and end-century = 2080–2100). The difference between the temperatures in relation to the baseline was calculated to identify the direction of change in temperature. A positive sign indicated a warming condition, while a negative sign indicated a cooling condition. Given the global scope of our analysis, the use of average annual temperature was justified, as these data are commonly used to characterize producing regions, even though average annual temperature can mask extreme weather events that would be prejudicial to grapes or to the quality of the wine [19].

The data for the geographic areas were based on the maps produced by vineyards.com and manually digitized by the authors. The data for the baseline were average annual temperature data for 1970–2000 and were obtained from the WorldClim 2.0 dataset at

a spatial resolution of 30 s [20]. The data for average annual temperature under the climate change condition were from the National Aeronautics and Space Administration (NASA) Goddard Institute for Space Studies (GISS) climate model [21]. We used the downscaled version of GISS data under shared socioeconomic pathway (SSP) 585 at 30 s spatial resolution for 2040–2060 and 2080–2100 from the WorldClim 2.0 database [14]. The GISS data under SSP 585 represent a climate model calibrated under representative concentration pathway 8.5 and shared socioeconomic pathway 5, which represent intensive use of fossil fuels and energy and economic development. From now on, we will refer to the downscaled version as GISS.

For the baseline conditions (Figure 1), annual average temperatures in the ten regions varied from 9.11 °C for the Mosel Valley region, Germany, to 16.54 °C for the Stellenbosch region, South Africa (Figure 1). It is interesting to note the difference between Mendoza, Argentina, and Maipo Valley, Chile, as Mendoza recorded 10 °C and Maipo Valley was 4 °C warmer at 14.38 °C. In the baseline, five regions were in the two lower temperature ranges.



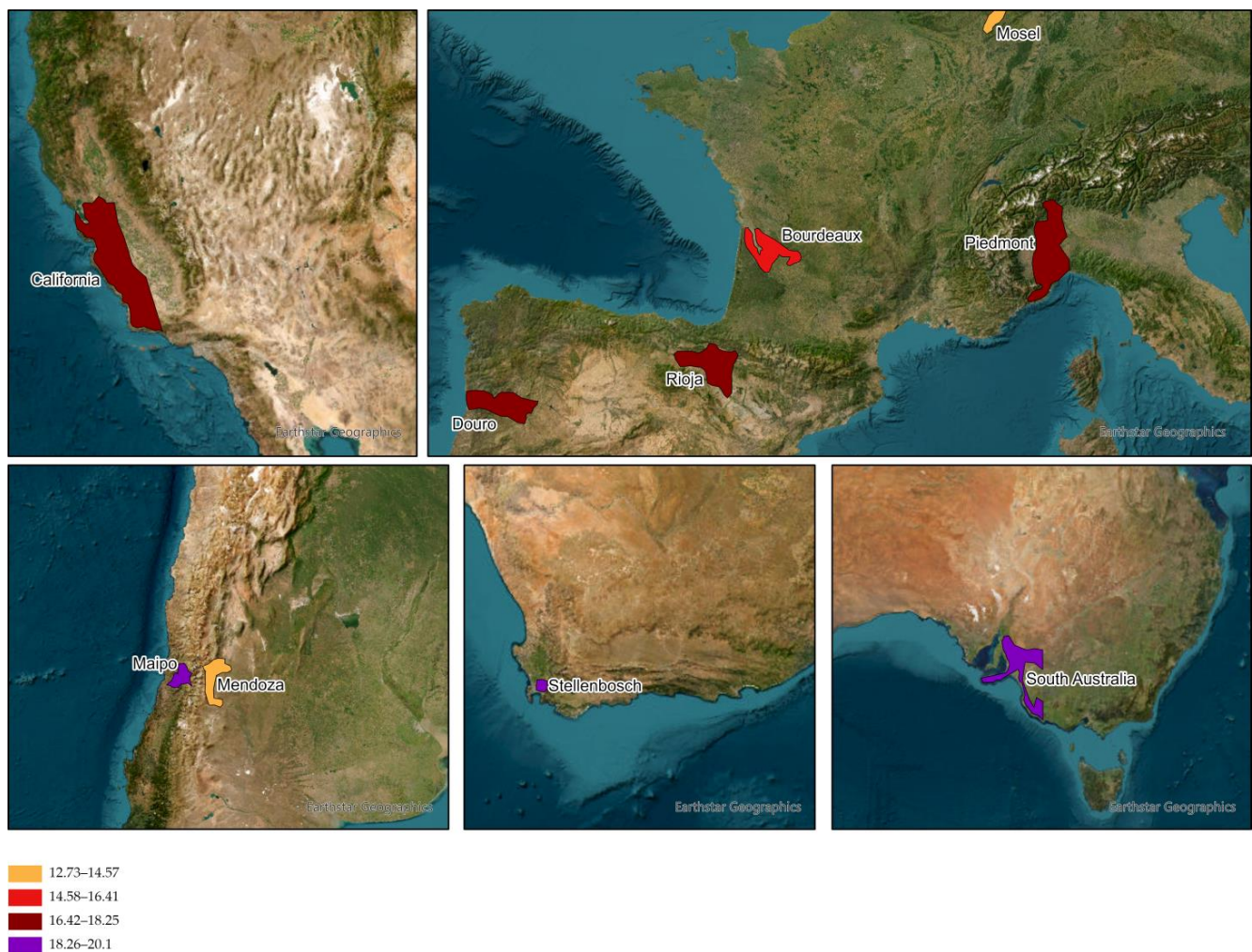
**Figure 1.** Average annual temperature for selected wine-producing regions for baseline period (1970–2000) in Celsius. Source: Average annual temperature [20], wine-producing regions manually created by the authors, World Imagery by Esri and Earthstar Geographics.



Under the future climate conditions estimated by the GISS, there will be remarkable warming in all regions. Figures 2 and 3 show an increase in temperature, which prompted us to add new temperature ranges to our scale to accommodate it. For the mid-century period, the temperature range was from 11.33 °C to 18.22 °C. There was no region under 10 °C. We identified the highest increase in temperature (2.5 °C) as occurring in Piedmont, Italy, and the lowest increase in temperature (1.6 °C) in Central Coast California, United States. For the end-century period, the temperature range was from 12.72 °C to 20.09 °C. The lowest temperature class for this period corresponded to a mid-to-high temperature in the baseline climatic condition. We identified the highest increase in temperature (4.4 °C) as occurring in Piedmont, Italy, and the lowest increase in temperature (3.3 °C) in Bordeaux, France. Furthermore, three regions (Maipo Valley, South Australia, Stellenbosch) had temperatures in the range from 18.2 °C to 20.1 °C.



**Figure 2.** Average annual temperature for selected wine-producing regions for middle of century period (2041–2060) in Celsius. Source: GISS—average annual temperature for 2041–2060 [20], wine-producing regions manually created by the authors, World Imagery by Esri and Earthstar Geographics.



**Figure 3.** Average annual temperature for selected wine-producing regions for end of century period (2081–2100) in Celsius. Source: GISS—average annual temperature for 2081–2100 [20], wine-producing regions manually created by the authors, World Imagery by Esri and Earthstar Geographics.

### 3. Impact of Climate Change on Grapevines, Berries, and Wine

Grapevine and wine production is vulnerable to the effect of climate change. Heat and water stress related to climate change can affect vine phenology, pest and disease pressure, crop yield, and berry and wine composition.

#### 3.1. Grapevine Phenology

Grapevine phenology has two developmental cycles known as the vegetative and reproductive cycles. These cycles are complex and include vine and fruit growth for the current season as well as the next. These cycles are responsible for berry formation and growth of the vine. There is a strong correlation between grapevine phenology and environmental factors, such as daylight length, heat, water, soil, and light. Climate change factors, such as elevated  $\text{CO}_2$ , elevated temperatures, and water stress, have played an important role in grapevine growth patterns [22].

Carbon dioxide is a major concern with climate change as the atmospheric  $\text{CO}_2$  concentration is expected to reach 600 ppm. High  $\text{CO}_2$  levels may promote physiological changes in vines. Increasing  $\text{CO}_2$  levels will limit the production of plant hormones, such as ethylene and jasmonic acid, which are both significant in the plants' defense response [23].

Temperatures can affect grape quality by leading to an earlier onset of the growing season, creating premature véraison from early heat, which can affect enzyme activation or



cause poor ripening. This would cause the growing season to be shortened and accelerate berry ripening [24]. When the sugar accumulation starts earlier and progresses more rapidly in the warmer periods of the growing season, the phenolic composition and berry anthocyanin concentration can be negatively affected, causing less than desirable harvest levels. For example, during flowering and the berry growth period, heat extremes can cause berry softening and changes in berry color. It has been shown that the accumulation of anthocyanins, which are responsible for berry coloration, is lower when maturation occurs at higher temperatures [25]. Premature véraison might lead to less aroma and flavor compound accumulation, which can affect berry flavor development [26]. Similar observations have been made in different wine regions [27–30]. In summary, increasingly high temperatures would lead to negative outcomes for grape and wine quality.

A major factor affecting the growth and productivity of the vine is water availability. Climate change is associated with unpredictable precipitation patterns and more severe drought conditions that can be expected to impact yield and the overall growth of vines. When water shortages occur early in the season, they can reduce yield by affecting bud fertility, as the reproductive structures inside the dormant buds are sensitive during this time [31]. Once the grapevine gets past budbreak, the reproductive growth is relatively unaffected. Another effect of climate change is the combination of higher temperatures and unpredictable precipitation, leading to water deficits. High rates of evapotranspiration and increased plant water requirements are caused by higher temperatures [32]. When water deficits are up to 50% of the evapotranspiration threshold, there is almost no effect on yield; however, when the threshold is surpassed, the yield decreases. This phenomenon is more prominent during the budbreak to bloom phenological phase [33,34].

### 3.2. Pest and Disease Pressure

Studies have found that climate change is inducing greater pest and disease pressure in vineyards. A healthy vine can create resistance or fight off potential attacks due to the plant's defense system. Pests or pathogens usually affect the vine during specific exposed periods of the vine's lifecycle. However, climate change can modify the period a plant will be exposed to a pathogen. For instance, high temperatures would promote pathogen development and increase survival rates, which can change the susceptibility of a host (plant) to pests and diseases [35].

Plant diseases can be used as an indicator of climate change. This can be complicated with all the biological interactions that result in disease [35]. Disease will occur more often when the vines are stressed in warmer climates. There is a concern regarding diseases that increased CO<sub>2</sub> levels will decrease plants' ability to decompose, so leaves or plant material on the ground can cause fungal spore development if not managed properly [36,37]. This, along with extremely hot temperatures that can cause berry sunburn and damage the berry skin, could increase the *Botrytis cinerea* infection rate in grapes [38]. When the phenology of the plant and the pathogen align, more plant diseases can occur.

Depending on the magnitude of global warming, it may influence the phenology of insects by impacting the timing of their emergence and feeding patterns. Since there may be a change in the timing of grapevine phenology—for example, budbreak or foliar growth—this could lead to a change in insect survival, since insects' timing is determined by the plant. If they are not emerging at the stages of growth needed for survival, it could cause the population to undergo food starvation or be unable to meet survival needs [39]. Elevated CO<sub>2</sub> concentrations can lead to the accumulation of non-structural carbohydrates in plant, which results in lower tissue nitrogen concentrations. This can lead to the need for insects to consume more foliage to meet their nitrogen needs [40–42]. Thus, the effects of climate change on insects are complex and involve several unknown factors, such as the introduction of new pests, competition among pests, and the presence of beneficial insects.

However, many of the highly mobile enemy insects can track climate change conditions, while this could take a while for less mobile species [39]. Vine mealybugs, for example, are less likely to leave their hideouts under the bark and on the roots to mi-

grate towards leaves and fruit during hot conditions. However, if the daily temperatures are to increase by 2 °C or 4 °C, most regions would expect to see an increase in overall mealybug density. In California, the mealybug densities are predicted to be higher in the cooler vineyards, such as Napa and Sonoma, while being lower in warmer areas, such as the Coachella Valley [43]. The extremely hot summer temperatures in the Coachella Valley can cause mealybug mortality. In comparison, the summer temperatures in the San Joaquin Valley are not hot enough to cause high mortality rates; thus, the mealybug populations can increase significantly during the growing season. In cooler areas, such as the coastal regions, mealybug abundance follows a similar pattern to the San Joaquin Valley but with fewer summer generation cycles [44]. As many of the distributions of the pest and pathogens will be changed, interactions between them and the plants will need to be monitored. Additionally, researchers have found that pest and weed management are the main hotspots for greenhouse gas emissions in the lifecycle of grape production [45,46].

### 3.3. Grapevine Yield

Grapevine yield depends on soil fertility and climatic conditions. High yields can be achieved with moderately high temperatures, sufficient light conditions, and enough nitrogen and water. Increased temperatures are beneficial for crop yield in some cool climate regions. Nemani et al. (2001) [29] found that yields and berry quality were improved in Napa and Sonoma Valleys due to lower occurrence of frost and a longer growing season. However, grapevines grown under excessive heat stress suffer photosynthesis limitation, thus contributing to significant yield reductions. Heat waves may result in a yield decrease of up to 35% in some viticultural regions [47]. Drought conditions impair grape yield, and the decrease can be variety-dependent [48]. During drought conditions, stomatal closure and the impairment of the photosynthetic machinery limit photosynthesis [48]. Increased water deficit due to reductions in precipitation in conjunction with increases in evapotranspiration influences yield. Studies have found that water deficit negatively affects yield [49,50]. Berry weight is one of the yield components most affected by water availability and is used for calculating the yield of a vineyard. A study showed that Shiraz vines with water deficit after flowering demonstrated significant reductions in berry weight compared to sufficiently watered vines, particularly during high-temperature seasons [51]. Another source of yield variability is temperature. A study on Sangiovese and Cabernet Sauvignon in Italy in relation to climate change reported that warmer weather resulted in higher yield variability [52].

### 3.4. Berry and Wine Composition

Grape berries are composed of several hundreds of chemical compounds, including water, fermentable sugars, organic acids, nitrogen compounds, minerals, pectins, phenolic compounds, and aromatic compounds [53]. Environmental conditions, such as soil, topography, and climate, influence yields, grape composition and sensory attributes, and the quality of the wines. Higher temperatures can accelerate grape metabolism, leading to changes in the biosynthesis of basic components [54]. A report found that metabolic pathways in grapes changed once the ambient temperature was 30 °C [25]. Many studies provide evidence of grape and wine composition changes, including dramatic changes in pH, total acidity, and alcohol [5,14,55].

#### 3.4.1. Sugar, Acid, and Alcohol

Elevated temperatures have been associated with sugar accumulation and organic degradation, resulting in an unbalanced sugar–acid ratio [56]. Wine made from these kinds of berries contains higher alcohol content and falls short in freshness and aromatic complexity. Warmer conditions in Slovenia led to a large reduction in total acidity in early-ripening wine varieties [57,58]. Grape malic acid levels are typically low in warm climate regions, since it tends to degrade at high temperatures. Lecourieux et al. found that imposing heat treatment (+8 °C, 14 days) on grape clusters at véraison and during berry ripening



significantly decreased the berry concentration of malic acid, while it increased some amino acids, such as phenylalanine,  $\gamma$ -aminobutyric acid, proline, and leucine. The alterations in acid concentrations could have been due to the deep remodeling of transcriptomes in heated berries [59].

### 3.4.2. Anthocyanins

Anthocyanins in berries are phenolic compounds that are responsible for berry coloration. Temperature has been found to be a critical factor affecting anthocyanin synthesis due to enzymes in the metabolic pathway being temperature-sensitive [60,61]. If sugar accumulation starts earlier and proceeds more rapidly in the growing season under high temperatures, the berry anthocyanin concentration cannot reach the desirable levels at the harvest. This is especially true in warm-climate wine regions. Ripened berries typically have an unbalanced composition, with higher total soluble solids, low acidity, and fewer anthocyanins. Moreover, anthocyanins are highly unstable and susceptible to thermal degradation. Research has indicated that anthocyanins tend to accumulate better at 20 °C than 30 °C [62]. The anthocyanin accumulation decreases once the temperature rises above 30 °C [63,64]. *Véraison*, heat treatment, and ripening heat treatment (+8 °C, 14 days) decreased the concentrations of anthocyanins at harvest, such as delphinidin-3-O-glucoside, cyanidin-3-O-glucoside, petunidin-3-O-glucoside, peonidin-3-O-glucoside, delphinidin-3-O-(6'-acetyl) glucoside, cyanidin-3-O-(6'-acetyl) glucoside, petunidin-3-O-(6'-acetyl) glucoside, and peonidin-3-O-(6'-acetyl) glucoside [59].

### 3.4.3. Aroma

Increased temperatures and solar radiation will alter the secondary metabolites in berries, thus impacting flavor development [5]. The temperature range of 20–22 °C appears to be optimal for aroma formation during the grape maturation stage for most varieties [65]. Increased volatilization of aroma compounds at high temperatures has been observed. Belancic et al. reported that the terpenol content of Moscatel de Alejandria (Muscat of Alexandria) and Moscatel Rosada (Muscat Rose) was lower as a result of the vines' overexposure to sunlight and higher berry temperature [66]. The concentrations of some aroma compounds may increase due to high temperatures, but this imparts negative effects on the wine since it breaks the balance of the aroma profile. For example, Marais et al. found that the concentrations of 1,1,6-trimethyl-1,2-dihydronaphthalene (TDN) varied with climatic regions. Warm temperatures increased the formation of TDN, which might have had a negative effect on Riesling wine, with overpowering petrol notes [67]. In a study by Lecourieux et al. (2017) [59], Cabernet Sauvignon berries exposed to high temperature showed decreased aromatic potential due to deregulation of numerous aroma and aroma precursor-related genes. The results suggested that heat treatment contributed to the decrease in volatile terpenoids caused by the repression of many key enzymes in the biosynthetic pathway. Carotenoid biosynthesis also decreased with heat treatment. High-temperature exposure also led to a drastic reduction in 2-methoxy-3-isobutylpyrazine (IBMP) content in ripe berries due to the repression of the key gene *VviOMT3*. *VviOMT3* was reported to be responsible for the synthesis of the IBMP [68].

The impact of climate change on grapevines, berries, and wine are summarized and presented in Table 1.

**Table 1.** Summary of impact of climate change on grapevines, berries, and wine.

Category	Effects of Climate Change	References
Grapevine phenology	Shortens the growing period and fastens berry ripening Causes high rates of evapotranspiration and increases vine water requirements	[31,33,34]

Table 1. Cont.

Category	Effects of Climate Change	References
Pest and disease pressure	Promotes pathogen development and pest increased survival rate Changes the susceptibility of vines Reduces beneficial insects	[35–37,40]
Grapevine yield	Improves yields in cool wine regions due to less frost occurring and longer growing season Limits vine photosynthesis and decreases significant yields due to excessive heat stress Causes higher yield variability	[29,47,48,52]
Berry and wine composition	Accelerates grape metabolism Increases sugar accumulation and reduces organic acid, resulting in an unbalanced sugar–acid ratio Increases alcohol content Decreases anthocyanin accumulation Has an impact on flavor development	[57–59,66,67]

#### 4. Strategies for Climate Change Adaptation

“Climate change adaptation is the process of adjusting to current or expected climate change and its effects” [1], and it is the key to the future of agriculture, which depends heavily on weather and climatic conditions. To continuously achieve high quality grapes and wine, different viticultural manipulations and technological solutions are being developed to mitigate the negative effects of climate change.

##### 4.1. Vineyard Location and Vine Row Orientation

The radiation intercepted by the grapevine canopy is dependent on several factors, such as topography, slope direction, steepness, and vineyard altitude. Vineyard location selection could be a cost-effective way to allow the winegrower to adapt locally to climate change. For example, cultivating wine grapes at a higher altitude is a useful strategy since vineyard altitudes play a significant role in combatting the impact of climate change [7]. High-altitude viticulture exposes the vine to more ultraviolet light but allows a low temperature to be maintained, which favors higher quality grapes and premium wine. One study conducted in Southern Brazil found that fully mature Cabernet Sauvignon berries could be produced in a high-altitude vineyard with high total soluble solids and phenolics contents [69].

Row orientation may also be a useful tool as a viticulture practice to adapt to climate change. There is a climate and temperature difference between the Northern and Southern Hemispheres. Due to sun exposure, north-facing slopes are typically cooler than south-facing slopes in the Northern Hemisphere. The Southern Hemisphere has the opposite situation. Studies have demonstrated that row orientation could influence vine physiology, the grapevine growth cycle, crop load [70–72], and berry and wine quality [73]. Research conducted in a Shiraz vineyard in Robertson, South Africa, found that (1) the highest skin anthocyanin and phenolic contents were present in grapes grown with a northwest–southeast (NW–SE) row orientation and (2) grapes from the NW–SE and NE–SW rows carried the most favorable sensory notes [73]. Therefore, for wine regions with elevated temperatures, it would be more appropriate to select a vineyard row orientation that leads to low canopy light interception.

##### 4.2. Plant Material

The selection of plant material is an affordable and useful solution for adapting to climate change. Drought-resistant varieties, rootstocks, and clones are expected to decrease the vulnerability of vineyards to water deficits and help avoid reductions in quality caused by elevated temperatures during berry ripening.

#### 4.2.1. Rootstocks

Drought-tolerant rootstocks can support scions' growth and increase grape production when water supply is limited. Grapevine rootstocks with different degrees of drought resistance are summarized in Table 2. Among these rootstocks, 110 Richter, 140 Ruggeri, 196.17 Castel, and 44–53 M show a high degree of drought tolerance, while O39-16, 420A, 1616C, Riparia Gloire, and 5C have low resistance. Other known rootstocks have low, low to moderate, and moderate to high drought tolerance [74]. Newer rootstocks, such as Matador, have not been evaluated for drought tolerance [75], and no information is available for RS-3, RS-9, Minotaur, or Kingfisher.

With the changes in climate conditions causing higher drought incidence, water and soil salinity is becoming an increasingly serious problem for grapevines. Certain rootstocks can protect scion varieties from accumulating significant amounts of saline ions. These mechanisms are not fully understood and research is underway. Existing rootstocks have been tested to determine their salt tolerance levels. Rootstocks such as Ramsey, Dogridge, 140 Ruggeri, and 1103 Paulsen have shown high tolerance to salt in the field [76].

**Table 2.** Grapevine rootstocks: degrees of drought tolerance.

Grapevine Rootstock Drought Tolerance				
High	Moderate to High	Medium	Low to Moderate	Low
110 Richter	99 Richter	5BB	101–14 Mgt	O39-16
140 Ruggeri	1103 Paulsen	Schwarzmann	Harmony	420A
44–53 M	Ramsey	Dogridge	SO4	5C
196.17 Castel		Freedom	St. George 3309C	1616C Riparia Gloire

Source: Adapted from [74,77].

#### 4.2.2. Varieties

Switching to drought-resistant and heat-tolerant grape varieties could also help wine-makers adapt to climate change [78]. Based on historical cultivation, grape growers have selected several varieties for drought tolerance, such as Grenache, Carignan, Cinsault, and most Mediterranean varieties. Several studies have been conducted to explore drought-resistant varieties [79,80]. In a study conducted in the Denominación de Origen Calificada Rioja in Spain, the red varieties Garnacha Roya, Alicante Bouschet, Trepát, Morate, and Agawan and the white varieties Maturana Blanca and Garnacha Blanca were the promising varieties for addressing global warming. These varieties have been shown to produce berries with lower sugar levels and higher acidity content during harvest in comparison to other traditional grapevine varieties [81]. Since berry maturity and ripening times vary significantly between the different varieties, grape growers can also select later-ripening varieties, such as Cabernet Sauvignon, Grenache, Sangiovese, Barbera, Nebbiolo, Mourvedre, Aglianico, Marsanne, Cortese, etc., to extend the harvest date and to ensure that the desired grape quality parameters are met.

#### 4.2.3. Clones

A certain level of genetic variability exists within a given grape variety. Clones have been historically selected for traits such as high productivity, disease resistance, cold hardiness, etc. In the context of a changing climate, new clones can be selected for drought resistance and improved water use efficiency. It would also be a useful strategy to change the grape growing cycle to delay maturity and reduce sugar accumulation at an early stage. Van Leeuwen conducted a study on different Cabernet Franc clone characteristics and found a difference in berry sugar accumulation between clones, providing valuable guidance for growers in choosing appropriate clones in warm climate regions [82].

#### 4.3. Training Systems and Canopy Management

Training system selection can be used to reduce high temperatures in the fruit zone of grapevines. The goblet system is an ancient untrellised vine-training method developed by wine growers in the Mediterranean region. This system made it possible to grow grapes in semi-dry and dry areas for a long time period because it was cost-effective and drought-resistant [83]. However, goblet-trained vines tend to produce a low crop load and cannot easily be adapted to mechanical cultural practices, such as mechanized shoot thinning, berry/cluster thinning, pruning, and harvesting. Hence, it is not practical for wine regions that focus on high-yield grape production. The San Joaquin Valley in California is a major wine production area characterized by a hot climate. Many of the wine grapes planted in this area adopted the California sprawl training system or the simple curtain system. Grapes are grown on two wires (a fruiting/cordon wire and a foliage wire), resulting in sprawling vines without rigorous shoot positioning [84]. Sprawl systems can provide the canopy shading required to prevent sunburn on the grapes, and they are also cheap to construct and suitable for grape varieties with moderate to high vigor.

Canopy management is increasingly recognized as an important tool for improving vineyard water management since it can affect the vine leaf area index and the transpiration rate. Various canopy management practices can be adopted throughout the year that enable the grower to mitigate the negative effects of climate change. Shoot trimming or topping is one useful canopy management strategy to delay berry ripening and slow down sugar accumulation that has been reported in several studies [85–87]. Leaf removal after véraison has been reported as an efficacious technique to adapt to climate change because it delays sugar accumulation and postpones the harvest date. A two-year study on mechanical post-véraison leaf removal from Sangiovese in a commercial vineyard in central Italy showed a significant reduction in total soluble solids (from 23.9°B to 22.7°B) and similar concentrations of other chemical components [88]. This study demonstrated that the timing and intervention intensity of canopy management are critical.

#### 4.4. Water Management

Grapevines are more drought-tolerant compared to other fruit crops, but a certain amount of water is needed to support vine growth and berry development. The water demand for a grapevine is dependent on the region, variety, rootstock, training system, planting density, and yield. A study found that grapevines needed annual rainfall of between 34 and 503 mm from April to October in a cool climate region of California (Carneros district of Napa Valley) [89], while in the other five warm climate regions in California (Madera, Livermore, Kearney, Paso Robles, and Temecula), grapevines needed between 450 and 800 mm of water from April to September [90]. Irrigation management is one of the strategies adopted in wine regions across the world to address climate change.

Irrigation systems can be selected to maximize water use efficiency, such as subsurface irrigation, drip irrigation, and high-pressure irrigation systems. “Subsurface irrigation system is below the ground surface so that water is supplied directly to the root zone of the plant” [91]. This system has benefits such as reducing evaporation losses and providing less hindrance to surface cultivation works. “Drip irrigation is a process of slow application of water on, above or beneath the soil by the surface, sub-surface, bubbler, and spray or pulse system” [91]. This is an efficient water application method since the applied rate is almost equal to the use rate of the crop. “Sprinkler irrigation is an advanced method of irrigation in which water is sprayed to air and allowed to fall on the ground similar to rainfall” [92]. In this method, the water is distributed through a nozzle connected to a network of pipes while under pressure. The main advantages of sprinkler irrigation are the savings in water/labor, and it is suitable for all soils with infiltration rates lower than 4 cm/h. Drip irrigation systems combined with regulated deficit irrigation (RDI) at an early or late stage provide the most control over water use. In other words, the grape grower can apply the precise amount of water that each grapevine needs. Precision scheduling and timing of the irrigation water supply are needed, and this can be accomplished with vine water



status monitoring through stem water potential measurements. The RDI technique can be applied for different vines and berry growing stages to accomplish different objectives. McCarthy et al. (1997) reported that application of RDI after flowering successfully reduced the berry weight of Shiraz vines [51]. Dry et al. (2001) reported that RDI induced an accumulation of anthocyanins in red wine grape varieties in Australian vineyards [93]. In Australia, RDI is current a vineyard management tool that is widely used to improve grape and wine quality. RDI is also a valuable management strategy to enhance vineyard profitability. Bellvert et al. (2020) used cost–benefit analysis to calculate the total savings in a 247-acre vineyard with the precision irrigation strategy for Tempranillo, Cabernet Sauvignon, and Syrah varieties. They reported a net benefit of USD 9.6 per acre in 2016 and USD 19.7 per acre in 2017 [94].

Remote sensing is a new technology used in vineyards to optimize precision irrigation and improve water use efficiency. Manned aircraft and satellites are currently the two main platforms for collecting data remotely. However, these systems are not widely used in vineyards due to several limitations, such as low spatial resolution, the high initial investment cost, and susceptibility to climatic conditions [95]. Bellvert (2014) successfully used airborne imaging from unmanned aerial vehicles (UAVs) to map the spatial variability in water requirements across a 27-acre block of Pinot Noir vineyards located in Lleida in Spain [96]. UAVs interlinked with wireless sensor networks (WSNs) represent a promising monitoring approach in vineyards. Using a multi-sensor approach (high-resolution UAV remote sensing and WSN proximal sensing), Di Gennaro et al. (2017) obtained a combined dataset that could be used to predict individual vine stress levels and berry quality [97]. Sun et al. (2017) used Landsat satellite imagery to predict grape yields and the effects of drought on the yield for Pinot Noir vineyards in California [98].

#### 4.5. Soil Management

Soil management is an important tool to prevent water loss in vineyards. Soil affects available water capacity, water infiltration, and evapotranspiration. Agronomic practices and soil amendments can be used to improve the soil surface state and soil structure in terms of porosity, stoniness, and deepness. With increasing drought conditions in semi-dry and dry areas, drought-tolerant soils could be reserved for vines sensitive to water stress. Chrysargyris et al. (2018) found that combining no-irrigation and no-tillage practices was a useful vineyard strategy for Maratheftiko, a heat-resistant variety native to Cyprus [79]. Pinamonti (1998) reported that the use of compost mulched soils in a Merlot vineyard improved soil structure, water-holding capacity, and the nutrient profile [99]. Soil water evaporation and temperature variations were also reduced [99]. Monteiro and Lopes (2007) conducted a 3-year study on the effects of permanent resident vegetation and permanent sown cover crops on a 15-year-old Cabernet Sauvignon vineyard located in the Estremadura region of Portugal [100]. Cover crop treatments showed significantly low daily water use during berry ripening due to low soil evaporation. Vineyard cover crop usage also reduced vine vegetative growth and increased total phenols and anthocyanins in grape berries. Cover cropping is a useful vineyard floor management tool as it can also help maintain living microorganisms in the soil. Clover and native grass planted in soils had greater microbial biomass than tilled soil in a Merlot vineyard located in Sacramento, California [101].

It is important to manage cover crops within a reasonable time frame. Water competition between vines and cover crops needs to be avoided; otherwise, it can lead to severe water stress in vines and have negative effects on vine growth, yield, and berry quality. The accurate selection of species and varieties is also critical. A study in an organic Manto Negro vineyard located in central Majorca, Spain, reported that self-seeding or perennial species, such as *Trifolium* sp., *Medicago* sp., *Brachycalycinum* sp., and *Dactylis* sp., improved soil properties. Vine vigor and yield were reduced due to water resource competition between cover crops and vines [102].

Soil amendments are traditionally used to improve soil structure. Biochar application is the most studied amendment strategy. Biochar is a co-product of the thermochemical conversion of biomass, and it has a microscopically porous structure that makes it possible to hold water. The effects of biochar depend on its physical and structural characteristics, the application rate, and the soil properties. Baronti et al. reported that biochar application in a central Italian vineyard increased soil water content and improved photosynthetic activity but had no effect on soil hydrophobicity [103]. A study carried out in a non-irrigated vineyard in Tuscany, Italy, investigated the effects of biochar applications on crop load and berry quality. The results demonstrated that biochar improved soil properties and increased soil water content. The crop load increased from 16% to 66% during the four years of research conducted without an effect on berry quality [104]. All these findings support the feasibility of a biochar-based strategy for vineyards in drought area adaptations to climate change and water shortages. For California growers, the Sonoma Biochar Initiative implemented by the Sonoma Ecology Center (SEC) provides farm manager training and education on how to produce and use biochar (<https://sonomabiocharinitiative.org/>, accessed on 10 July 2022)).

New technologies are playing an important role in soil management aimed at adaptation to climate change. The use of geographic information systems (GISs) and remote sensing can improve the management of vineyards by providing information at the field level more quickly than manual data collection [105,106]. Soil moisture sensors have emerged as a promising approach in vineyards, and they can support growers in making correct and timely decisions regarding irrigation, prevent vines from suffering drought stress, and save considerable water. Soil moisture content can be determined using gravimetric methods, nuclear gauges, or other methods based on electromagnetic, tensiometric, hygrometric, or remote sensing processes. The most common sensors are based on electromagnetic technology. Dilrukshi et al. (2019) evaluated the accuracy and repeatability of a commercial soil moisture sensor (SKU: SEN0193) [107]. They also integrated the sensor with a data acquisition system and developed a cost-effective automated soil moisture monitoring system to reduce soil moisture stress and avoid excessive water application. Stevanato et al. (2019) assessed the capabilities of a novel cosmic-ray neutron sensor (CRNS) in two different fields located in Potsdam in Germany and Lagosanto in Italy [108]. The study demonstrated that the CRNS could be a powerful tool for soil moisture measurements in large agricultural areas. Soil electrical conductivity sensors have been developed for soil mapping purposes. Rodriguez-Perez et al. (2021) evaluated the feasibility of using apparent electrical conductivity (eCa) in a Napa Valley vineyard in California [109]. The investigation indicated that eCa could be used to predict soil physical and chemical properties. The study clarified the feasibility of using soil electrical conductivity sensors to estimate soil properties to save time and cost.

#### 4.6. Other Adaptation Strategies

Further promising strategies to adapt to climate change include the use of shade nets; antitranspirants, such as kaolin particle film and chitosan; and late pruning. These strategies could be used to mitigate the impact of high temperature and have been examined for some cultivars and climatic conditions.

Sabir et al. (2020) investigated the effects of different colored shade nets on the berry skin color and functional properties of “Alphonse Lavallée” table grapes under controlled glasshouse conditions [110]. The results demonstrated that application of yellow netting led to the highest anthocyanin content (15.1 mg/100 mL juice), followed by green (14.2 mg/100 mL juice) and black (13.2 mg/100 mL juice) nets. Caravia et al. (2016) found that overhead shade treatment imposed on Shiraz vines from véraison to harvest gave the most consistent effect in mitigating heat stress [111]. The treatment also delayed berry mass loss and decreased wine alcohol content.

Antitranspirants are used in vineyards to counteract climate change since they can reduce excessive water loss and improve berry water moisture. Two types of plant antitrans-

spirants are currently available with different functions: film polymers (e.g., kaolin, Vapor Gard, etc.) and stomatal closing compounds, such as chitosan. Kaolin ( $\text{Al}_2\text{Si}_2\text{O}_5(\text{OH})_4$ ) is a white, chemically inert, and non-toxic clay material. It can reflect photosynthetically active radiation (PAR), ultraviolet (UV), and infrared radiation (IR) and can be used to mitigate the effects of excessive heat and radiation [112]. It is a promising and cost-effective mineral for the improvement of grape composition and wine quality in hot-climate regions with severe drought conditions. Garrido et al. (2019) found that kaolin film applied to leaves increased the photosynthetic activity of both exocarps and seed integuments of “Alvarinho” berries growing under low-light canopy conditions in the Demarcated Region of Vinho Verde, Braga, Portugal [113]. Dinis et al. (2016) showed that foliar kaolin application was beneficial to Touriga Nacional grapevines in the Douro Region of northern Portugal, which has a Mediterranean climate with a warm temperate and dry summer [114]. Palliotti et al. (2013) investigated the post-véraison application of the film-forming antitranspirant Vapor Gard (VG; a.i. di-1-p-menthene) on Sangiovese grapes in central Italy [73]. Their research showed that antitranspirant application delayed grape ripening and reduced berry sugar accumulation. The effectiveness of the Vapor Gard antitranspirant spray was also confirmed in [115]. The authors applied antitranspirant sprays to Barbera vines pre-flowering (PF), pre-véraison (PV), and at both dates (PFPV) in 2013 and 2014 in Piacenza, Italy. They found that all antitranspirant spray treatments reduced gas exchange by up to 46%, while the PV and PFPV treatments delayed sugar accumulation in both seasons without affecting color development. Brillante et al. (2016) observed that the fruit quality and consumer preference for wines made with the traditional antitranspirant pinolen were not comparable to those for wine made with kaolin-treated berries [116]. Chitosan is a naturally occurring polysaccharide produced from seafood shells. It can be used to simulate biotic and abiotic stresses since chitosan can induce abscisic acid activity. When vines are in a stress stage, abscisic acid plays a key role in minimizing stomata opening and maintaining water. Górnik et al. (2008) studied the effect of chitosan (Biochikol 020 PC) on the response of grapevines to drought and heat stress [117]. Studies have shown that chitosan can be widely used to improve rooting and heat and drought resistance. Singh et al. (2020) found that chitosan (0.01% in 0.01% acetic acid) application also increased several monomeric anthocyanins in Tinto Cão berry skins [118]. Silva et al. (2020) undertook a comparison study of chitosan solutions and chitosan nanoparticles applied to the grape variety Sousão [119]. They found that, compared to chitosan nanoparticle treatments, chitosan solutions were more effective in increasing the total phenol, anthocyanin, and tannin contents in berries, as well as antibacterial activity.

Late spur-pruning was originally performed in cool climate regions to delay bud burst and avoid frost damage. It has emerged as a potentially viable adaptation strategy in vineyards that can alleviate the impact of global warming by delaying vine development and berry ripening. Gatti et al. (2018) applied late-pruning strategies to a cane-pruned system of Pinot Noir vines in central Italy when the distal part of the cane reached an average of three leaves [120]. Late pruning significantly delayed ripening and extended berry maturity into a cooler season over two trial seasons. It was approved as a simple technique suitable for preserving the balanced must composition required for a range of sparkling wine styles, whereas yield per vine was reduced by 35% due to lower shoot fruitfulness. Buesa et al. (2021) applied late-pruning techniques to Bobal and Tempranillo varieties in eastern Spain when budbreak initiated [121]. Late pruning delayed grape ripening in both cultivars and resulted in higher berry anthocyanins at harvest, which contributed to the wine color intensity. However, late pruning had small detrimental effects on yield, showing a 10% reduction. Moran et al. (2017) limited pruning to two to three leaves and found that it effectively spread the harvest, increased the berry anthocyanin concentration, improved the wine chemical composition, and altered sensory profiles without negative effects on the yield for Barossa Valley Shiraz [122].

In summary, there are many different practices that can be used in the vineyard to help growers combat and mitigate the effects of climate change. The opportunities and weak-

nesses of adaptation strategies are crucial factors for farmers to consider when planning to adopt these practices. It has been found that wineries in Sonoma Valley with more than 15 years of experience had a positive correlation with adopting sustainable practices [123]. Growers' participation in partnership arrangements also had a positive correlation with adoption of new practices [124]. Some technologies are innovative, but most are not new and have been employed for a very long time. Each adaptation practice has advantages and shortcomings. Table 3 summarizes each practice and compares the advantages and limitations of the applications. These insights into adaptations can be informative for the future development of systematic methods to optimize their effectiveness.

**Table 3.** Summary of the strategies for climate change adaptation.

Adaptation Strategies	Opportunities	Weaknesses	Application Suggestions
Vineyard location and vine row orientation	Cost-effective ways to allow the winegrower to adapt locally to climate change	Initial cost of conducting research on a new location Require collection of information of altitude, climate, soil, etc.	Workshop and field demonstrations are necessary to help growers understand research and persuade them to explore a new location
Rootstock selection	Enables growers to select for grapevines that are more resistant to drought or disease, thus making it possible to maintain or increase vine productivity Helps to overcome soil problems, such as texture, pH, and density	Selection of a rootstock can be quite complex Selection becomes even more challenging with the changing climate	Consider how closely existing rootstock choices interact with other management strategies (e.g., irrigation and cover cropping) Consider rootstock's ability to survive in specific conditions, such as with different soil, pests, and viruses
New variety selection	A natural solution to delay berry ripening	Might be difficult to perform in some wine regions with traditional appellations	Continue monitoring and evaluating the new varieties over time Workshop and field demonstrations are necessary to help growers choose the right variety for their vineyards
Clone selection	A long-term solution involving changing the berry growing cycle to delay maturity and reduce sugar accumulation	Clonal selection is time-consuming it is necessary to follow a long, rigorous procedure Requires a variety of skills and techniques and specific equipment	Continue monitoring and evaluating the new clones over time to guarantee the reliability and quality of the selections
Training systems	An efficient solution to delay berry ripening	Initial investment required to investigate a new training system	Workshop and field demonstrations are necessary to help growers choose the right training system for their vineyards
Canopy management	An efficient solution to delay berry ripening	Need to train winegrowers	Timing and intervention intensity of canopy management are critical
Water management	Optimizes water use efficiency Reduces fertilizer waste Sustainable vineyard management	Initial and maintenance costs for precision irrigation and new technology are high High labor skills are required	More seasons are needed to achieve the desirable effect
Soil management	Prevents water loss Improves soil properties, vine performance, and berry quality over the long term	A long time is required to attain healthy soil Initial costs for soil sensors and soil mapping equipment	Soil amendment should be used for several seasons to achieve the desirable effect
Colored shade nets	Protect vines from excessive solar radiation Delay berry ripening	High labor cost	Need to consider different factors, such as fabric material, density percentage, size/shape of holes, and color



Table 3. Cont.

Adaptation Strategies	Opportunities	Weaknesses	Application Suggestions
Kaolin	<p>Cost-efficient</p> <p>Forms a physical barrier against various pests</p> <p>Reduces moisture to prevent diseases</p> <p>Acts against solar radiation to prevent berry sunburn</p> <p>Natural and safe agent for the winegrowers and can be used for organic vineyards</p>	<p>It is easily washed off by rain and kaolin spray is not suitable for overhead irrigation</p> <p>Requires constant application, especially for new leaves</p> <p>Requires strict attention to detail</p> <p>Missed leaves open window for pests and disease</p>	<p>It must be mixed thoroughly and applied via a sprayer with continuous agitation</p> <p>Be careful not to overspray</p> <p>Although it is generally regarded as safe for humans, it is still important to protect workers while spraying (long sleeves, long pants, closed-toed shoes, a mask or respirator)</p>
Other antitranspirants	<p>Reduce water loss through leaves by reducing transpiration</p> <p>Simple and viable technique to control berry sugar accumulation and obtain less alcoholic wines</p>	<p>Might increase fruit surface temperature and sunburn due to lack of transpiration</p> <p>Might decrease fruit quality and consumer preference for the wines</p>	<p>Application timing is critical</p> <p>Make sure lower leaf epidermis is fully wetted by the antitranspirant chemical</p> <p>Spray cautions same as for kaolin application</p>
Late pruning	<p>Simple and inexpensive</p> <p>Suitable for light-bodied wines, such as white, rosé, and sparkling</p>	<p>Might reduce yield for some varieties</p>	<p>More seasons are needed to achieve the desirable effect</p>

Source: Author's summary of the literature cited in Section 4.

## 5. Conclusions and Future Work

In view of the current and continuing changes in the global climate, the grape and wine industry faces big challenges. Climate change compounds issues such as heat, drought, and water stress, which are known to affect vine phenology, pest and disease incidences, crop yield, berry quality, and wine tasting. There are potential adaptation strategies that can be used to address these issues. On the one hand, there are effective and inexpensive practices, such as kaolin application and late pruning. On the other hand, there are practices that will have a long-term effect on vineyard sustainability, such as water and soil management. Rootstock, clone, and new variety selection are time-consuming and it is necessary to follow a long, rigorous process, which also requires a variety of skills and techniques and specific equipment. New technologies, such as soil moisture sensors and soil electrical conductivity sensors, can help in monitoring the available water and soil properties. Precision viticulture can provide more control with regard to water usage and save production costs. The Natural Resources Conservation Service implemented by the USDA has developed programs to help farmers and growers with regional hub information, assessments of soil carbon, systems to help account for carbon and greenhouse gas systems, and grants. There are programs that help with incentives for growers and ranchers who will sequester carbon, reduce atmospheric greenhouse gases, and improve soil health. The present work cannot present an in-depth review of all areas being studied in terms of viticulture and winemaking. Future work can build on this review and zoom-in on specific topics or adaptations. Another limitation in the literature is the lack of an established protocol through which researchers could publish their experimental findings, which hinders direct quantitative analysis of the diverse adaptation strategies. Future researchers can expand the present review and perform a bibliometric analysis of adaptation practices and management strategies.

In order to achieve long-term benefits, modifications to plant material should be considered a priority. They offer a natural way to obtain grapes with low sugar content and high acid profiles and to maximize their quality through delayed ripening. Late-ripening varieties can be found among the traditional varieties in some wine-growing regions. These are environmentally friendly and do not increase production costs. However, this strategy is obviously difficult to perform in some wine regions with traditional appellations. For example, winegrowers can only use local varieties in European wine-growing regions. More research on the use of non-local varieties in certain regions is necessary. With more drought-resistant rootstocks and clones available, regulations for these new plant materials

need to be modified. Better collaboration between the government, research institutes, industry partners, and growers will facilitate new plant material applications.

The use of new technology is a promising strategy to address global warming issues. However, some new tools are still expensive, and their uses and applications are limited for grape growers. Therefore, further development of affordable precision tools is needed. Innovation-based agriculture will need technological standardization to ensure the compatibility and safety of equipment. Meanwhile, standards should be kept updated to adapt to technological changes. Lastly, more educational workshops and technology demonstrations for growers are necessary.

**Author Contributions:** Conceptualization, Q.S., G.G. and S.V.Z.; methodology, Q.S. and G.G.; software, G.G.; validation, Q.S., S.V.Z. and G.G.; formal analysis, Q.S., G.G., L.G., J.V. and S.V.Z.; investigation, Q.S., G.G., L.G., J.V. and S.V.Z.; resources, S.V.Z. and G.G.; data curation, Q.S., G.G. and S.V.Z.; writing—original draft preparation, Q.S., G.G., L.G., J.V. and S.V.Z.; writing—review and editing, Q.S., S.V.Z. and G.G.; visualization, G.G.; supervision, Q.S. and G.G.; project administration, Q.S., G.G. and S.V.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** G.G.'s research was partially funded by the National Institute of Food and Agriculture (NIFA/USDA/AFRI), grant number 2022-67023-36150.

**Data Availability Statement:** No new data were created for this review.

**Acknowledgments:** The authors acknowledge the anonymous reviewers for their insightful comments. G.G. acknowledges the CPP CAFÉ Creates for productive writing sessions and the students in the Granco Research Lab.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## References

1. Pachauri, R.K.; Meyer, L.A. Climate Change 2014: Synthesis Report. Available online: [https://www.ipcc.ch/site/assets/uploads/2018/05/SYR\\_AR5\\_FINAL\\_full\\_wcover.pdf](https://www.ipcc.ch/site/assets/uploads/2018/05/SYR_AR5_FINAL_full_wcover.pdf) (accessed on 15 June 2022).
2. NASA. GISS Surface Temperature Analysis (v4). Available online: [https://data.giss.nasa.gov/gistemp/graphs\\_v4/](https://data.giss.nasa.gov/gistemp/graphs_v4/) (accessed on 15 June 2022).
3. National Centers for Environmental Information. Decadal Update from NCEI Gives Forecasters and Public Latest Averages for 1991–2020. Available online: <https://www.ncei.noaa.gov/news/noaa-delivers-new-us-climate-normals> (accessed on 15 June 2022).
4. Granier, C.; Bessagnet, B.; Bond, T.; D'Angiola, A.; van der Gon, H.D.; Frost, G.J.; Heil, A.; Kaiser, J.; Kinne, S.; Klimont, Z.; et al. Evolution of anthropogenic and biomass burning emissions of air pollutants at global and regional scales during the 1980–2010 period. *Clim. Chang.* **2011**, *109*, 163–190. [CrossRef]
5. Schultz, H. Climate change and viticulture: A European perspective on climatology, carbon dioxide and UV-B effects. *Aust. J. Grape Wine Res.* **2000**, *6*, 2–12. [CrossRef]
6. Riganti, C.; Sanchez-Lugo, A.U.S. Drought Monitor. Available online: <https://droughtmonitor.unl.edu/> (accessed on 20 January 2020).
7. Hannah, L.; Roehrdanz, P.R.; Ikegami, M.; Shepard, A.V.; Shaw, M.R.; Tabor, G.; Zhi, L.; Marquet, P.A.; Hijmans, R.J. Climate change, wine, and conservation. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 6907–6912. [CrossRef]
8. Christensen, J.H.; Hewitson, B.; Busuioc, A.; Chen, A.; Gao, X.; Held, I.; Jones, R.; Kolli, R.K.; Kwon, W.T.; Laprise, R.; et al. Regional climate projections. In *Climate Change 2007: The Physical Science Basis*; Solomon, S.D., Qin, M., Manning, Z., Chen, M., Eds.; Cambridge University Press: New York, NY, USA, 2007; pp. 847–940.
9. Olesen, J.E.; Bindi, M. Consequences of climate change for European agricultural productivity, land use and policy. *Eur. J. Agron.* **2002**, *16*, 239–262. [CrossRef]
10. Darma, S.; Lestari, D.; Darma, D.C. The Productivity of Wineries—An Empirical in Moldova. *J. Agric. Crop.* **2021**, *8*, 50–58. [CrossRef]
11. White, M.A.; Diffenbaugh, N.S.; Jones, G.V.; Pal, J.S.; Giorgi, F. Extreme heat reduces and shifts United States premium wine production in the 21st century. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 11217–11222. [CrossRef] [PubMed]
12. Schultze, S.R.; Sabbatini, P.; Luo, L. Effects of a warming trend on cool climate viticulture in Michigan, USA. *Springerplus* **2016**, *5*, 1119. [CrossRef]
13. Wanyama, D.; Bunting, E.L.; Goodwin, R.; Weil, N.; Sabbatini, P.; Andresen, J.A. Modeling Land Suitability for *Vitis vinifera* in Michigan Using Advanced Geospatial Data and Methods. *Atmosphere* **2020**, *11*, 339. [CrossRef]

14. Jones, G.V.; White, M.A.; Cooper, O.R.; Storchmann, K. Climate Change and Global Wine Quality. *Clim. Chang.* **2005**, *73*, 319–343. [[CrossRef](#)]
15. Ponti, L.; Gutierrez, A.P.; Boggia, A.; Neteler, M. Analysis of Grape Production in the Face of Climate Change. *Climate* **2018**, *6*, 20. [[CrossRef](#)]
16. Lionello, P.; Malanotte-Rizzoli, P.; Boscolo, R.; Alpert, P.; Artale, V.; Li, L.; Luterbacher, J.; May, W.; Trigo, R.; Tsimplis, M. The Mediterranean climate: An overview of the main characteristics and issues. *Dev. Earth Environ. Sci.* **2006**, *4*, 1–26. [[CrossRef](#)]
17. Sadras, V.; Petrie, P. Climate shifts in south-eastern Australia: Early maturity of Chardonnay, Shiraz and Cabernet Sauvignon is associated with early onset rather than faster ripening. *Aust. J. Grape Wine Res.* **2011**, *17*, 199–205. [[CrossRef](#)]
18. Bedsworth, L.; Cayan, D.; Franco, G. California's Forth Climate Change Assessment: Statewide Summary Report. Available online: [https://www.energy.ca.gov/sites/default/files/2019-11/Statewide\\_Reports-SUM-CCCA4-2018-013\\_Statewide\\_Summary\\_Report\\_ADA.pdf](https://www.energy.ca.gov/sites/default/files/2019-11/Statewide_Reports-SUM-CCCA4-2018-013_Statewide_Summary_Report_ADA.pdf) (accessed on 10 July 2022).
19. Gambetta, G.A.; Kurtural, S.K. Global warming and wine quality: Are we close to the tipping point? *OENO One* **2021**, *55*, 353–361. [[CrossRef](#)]
20. Fick, S.E.; Hijmans, R.J. WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **2017**, *37*, 4302–4315. [[CrossRef](#)]
21. Kelley, M.; Schmidt, G.A.; Nazarenko, L.S.; Bauer, S.E.; Ruedy, R.; Russell, G.L.; Ackerman, A.S.; Aleinov, I.; Bauer, M.; Bleck, R.; et al. GISS-E2.1: Configurations and Climatology. *J. Adv. Model. Earth Syst.* **2020**, *12*, e2019MS002025. [[CrossRef](#)] [[PubMed](#)]
22. Webb, L.B.; Whetton, P.H.; Barlow, E.W.R. Modelled impact of future climate change on the phenology of winegrapes in Australia. *Aust. J. Grape Wine Res.* **2007**, *13*, 165–175. [[CrossRef](#)]
23. Reineke, A.; Selim, M. Elevated atmospheric CO<sub>2</sub> concentrations alter grapevine (*Vitis vinifera*) systemic transcriptional response to European grapevine moth (*Lobesia botrana*) herbivory. *Sci. Rep.* **2019**, *9*, 2995. [[CrossRef](#)]
24. Yau, I.-H.; Davenport, J.R.; Moyer, M.M. Developing a Wine Grape Site Evaluation Decision Support System for the Inland Pacific Northwestern United States. *Horttechnology* **2014**, *24*, 88–98. [[CrossRef](#)]
25. Mori, K.; Sugaya, S.; Gemma, H. Decreased anthocyanin biosynthesis in grape berries grown under elevated night temperature condition. *Sci. Hortic.* **2005**, *105*, 319–330. [[CrossRef](#)]
26. Mullins, M.G.; Bouquet, A.; Williams, L.E. *Biology of the Grapevine*; Cambridge University Press: Cambridge, UK, 1992; pp. 120–135.
27. Jones, G.V.; Davis, R.E. Climate Influences on Grapevine Phenology, Grape Composition, and Wine Production and Quality for Bordeaux, France. *Am. J. Enol. Vitic.* **2000**, *51*, 249–261. [[CrossRef](#)]
28. Caprio, J.M.; Quamme, H.A. Weather conditions associated with grape production in the Okanagan Valley of British Columbia and potential impact of climate change. *Can. J. Plant Sci.* **2002**, *82*, 755–763. [[CrossRef](#)]
29. Nemani, R.R.; White, M.; Cayan, D.R.; Jones, G.V.; Running, S.W.; Coughlan, J.C.; Peterson, D.L. Asymmetric warming over coastal California and its impact on the premium wine industry. *Clim. Res.* **2001**, *19*, 25–34. [[CrossRef](#)]
30. Petrie, P.; Sadras, V. Advancement of grapevine maturity in Australia between 1993 and 2006: Putative causes, magnitude of trends and viticultural consequences. *Aust. J. Grape Wine Res.* **2008**, *14*, 33–45. [[CrossRef](#)]
31. Baeza, P.; Junquera, P.; Peiro, E.; Lissarrague, J.R.; Uriarte, D.; Vilanova, M. Effects of Vine Water Status on Yield Components, Vegetative Response and Must and Wine Composition. In *Advances in Grape and Wine Biotechnology*; IntechOpen: London, UK, 2019. [[CrossRef](#)]
32. Duchêne, E.; Schneider, C. Grapevine and climatic changes: A glance at the situation in Alsace. *Agron. Sustain. Dev.* **2005**, *25*, 93–99. [[CrossRef](#)]
33. Ramos, M.C.; Martínez-Casasnovas, J.A. Soil water balance in rainfed vineyards of the Penedès region (Northeastern Spain) affected by rainfall characteristics and land levelling: Influence on grape yield. *Plant Soil* **2010**, *333*, 375–389. [[CrossRef](#)]
34. Ramos, M.C.; Pérez-Álvarez, E.P.; Peregrina, F.; de Toda, F.M. Relationships between grape composition of Tempranillo variety and available soil water and water stress under different weather conditions. *Sci. Hortic.* **2019**, *262*, 109063. [[CrossRef](#)]
35. Pathak, T.B.; Maskey, M.L.; Dahlberg, J.A.; Kearns, F.; Bali, K.M.; Zaccaria, D. Climate Change Trends and Impacts on California Agriculture: A Detailed Review. *Agronomy* **2018**, *8*, 25. [[CrossRef](#)]
36. Nazir, N.; Bilal, S.; Bhat, K.A.; Shah, T.; Badri, Z.; Bhat, F.; Wani, T.; Mugal, M.; Parveen, S.; Dorjey, S. Effect of Climate Change on Plant Diseases. *Int. J. Curr. Microbiol. Appl. Sci.* **2018**, *7*, 250–256. [[CrossRef](#)]
37. Zayan, S.A. Impact of climate change on plant diseases and IPM strategies. In *Plant Diseases—Current Threats and Management Trends*; IntechOpen: London, UK, 2020. [[CrossRef](#)]
38. Steel, C.; Greer, D. Effect of climate on vine and bunch characteristics: Bunch rot disease susceptibility. *Acta Hortic.* **2008**, *785*, 253–262. [[CrossRef](#)]
39. Thomson, L.J.; Macfadyen, S.; Hoffmann, A.A. Predicting the effects of climate change on natural enemies of agricultural pests. *Biol. Control* **2010**, *52*, 296–306. [[CrossRef](#)]
40. Taub, D.R.; Wang, X. Why are Nitrogen Concentrations in Plant Tissues Lower under Elevated CO<sub>2</sub>? A Critical Examination of the Hypotheses. *J. Integr. Plant Biol.* **2008**, *50*, 1365–1374. [[CrossRef](#)]
41. Zavala, J.; Gog, L.; Giacometti, R. Anthropogenic increase in carbon dioxide modifies plant-insect interactions. *Ann. Appl. Biol.* **2016**, *170*, 68–77. [[CrossRef](#)]

42. Fajer, E.D.; Bowers, M.D.; Bazzaz, F.A. The Effects of Enriched Carbon Dioxide Atmospheres on Plant—Insect Herbivore Interactions. *Science* **1989**, *243*, 1198–1200. [[CrossRef](#)] [[PubMed](#)]
43. Gutierrez, A.P.; Daane, K.M.; Ponti, L.; Walton, V.M.; Ellis, C.K. Prospective evaluation of the biological control of vine mealybug: Refuge effects and climate. *J. Appl. Ecol.* **2007**, *45*, 524–536. [[CrossRef](#)]
44. Daane, K.M.; Bentley, W.J.; Smith, R.J.; Haviland, D.R.; Weber, E.A.; Battany, M.; Gospert, C.; Millar, J.G. Planococcus Mealybugs (Vine Mealybugs). In *Grape Pest Management*, 3rd ed.; Bettiga, L.J., Ed.; University of California Agricultural and Natural Resources: Davis, CA, USA, 2013; pp. 246–260.
45. Tian, D.; Zhang, M.; Wei, X.; Wang, J.; Mu, W.; Feng, J. GIS-Based Energy Consumption and Spatial Variation of Protected Grape Cultivation in China. *Sustainability* **2018**, *10*, 3248. [[CrossRef](#)]
46. Steenwerth, K.L.; Strong, E.B.; Greenhut, R.F.; Williams, L.; Kendall, A. Life cycle greenhouse gas, energy, and water assessment of wine grape production in California. *Int. J. Life Cycle Assess.* **2015**, *20*, 1243–1253. [[CrossRef](#)]
47. Fraga, H.; Molitor, D.; Leolini, L.; Santos, J.A. What Is the Impact of Heatwaves on European Viticulture? A Modelling Assessment. *Appl. Sci.* **2020**, *10*, 3030. [[CrossRef](#)]
48. Gambetta, G.A.; Herrera, J.C.; Dayer, S.; Feng, Q.; Hochberg, U.; Castellarin, S.D. The physiology of drought stress in grapevine: Towards an integrative definition of drought tolerance. *J. Exp. Bot.* **2020**, *71*, 4658–4676. [[CrossRef](#)]
49. Junquera, P.; Lissarrague, J.R.; Jiménez, L.; Linares, R.; Baeza, P. Long-term effects of different irrigation strategies on yield components, vine vigour, and grape composition in cv. Cabernet-Sauvignon (*Vitis vinifera* L.). *Irrig. Sci.* **2012**, *30*, 351–361. [[CrossRef](#)]
50. Myburgh, P.A. Response of *Vitis vinifera* L. cv. Merlot to Low Frequency Irrigation and Partial Root Zone Drying in the Western Cape Coastal Region—Part II. Vegetative Growth, Yield and Quality. *S. Afr. J. Enol. Vitic.* **2016**, *32*, 104–116. [[CrossRef](#)]
51. Mccarthy, M.G. The effect of transient water deficit on berry development of cv. Shiraz (*Vitis vinifera* L.). *Aust. J. Grape Wine Res.* **1997**, *3*, 2–8. [[CrossRef](#)]
52. Bindi, M.; Fibbi, L.; Gozzini, B.; Orlandini, S.; Miglietta, F. Modelling the impact of future climate scenarios on yield and yield variability of grapevine. *Clim. Res.* **1996**, *7*, 213–224. [[CrossRef](#)]
53. Waterhouse, A.L.; Sacks, G.L.; Jeffery, D.W. *Understanding Wine Chemistry*, 1st ed.; John Wiley & Sons, Ltd.: New York, NY, USA, 2016. [[CrossRef](#)]
54. Dokoozlian, N.K.; Kliewer, W.M. Influence of Light on Grape Berry Growth and Composition Varies during Fruit Development. *J. Am. Soc. Hortic. Sci.* **1996**, *121*, 869–874. [[CrossRef](#)]
55. de Orduña, R.M. Climate change associated effects on grape and wine quality and production. *Food Res. Int.* **2010**, *43*, 1844–1855. [[CrossRef](#)]
56. Schultz, H.R. Global Climate Change, Sustainability, and Some Challenges for Grape and Wine Production. *J. Wine Econ.* **2016**, *11*, 181–200. [[CrossRef](#)]
57. Spayd, S.E.; Tarara, J.M.; Mee, D.L.; Ferguson, J.C. Separation of Sunlight and Temperature Effects on the Composition of *Vitis vinifera* cv. Merlot Berries. *Am. J. Enol. Vitic.* **2002**, *53*, 171–182. [[CrossRef](#)]
58. Sweetman, C.; Sadras, V.O.; Hancock, R.D.; Soole, K.L.; Ford, C.M. Metabolic effects of elevated temperature on organic acid degradation in ripening *Vitis vinifera* fruit. *J. Exp. Bot.* **2014**, *65*, 5975–5988. [[CrossRef](#)]
59. Lecourieux, F.; Kappel, C.; Pieri, P.; Charon, J.; Pillet, J.; Hilbert, G.; Renaud, C.; Gomès, E.; Delrot, S.; Lecourieux, D. Dissecting the biochemical and transcriptomic effects of a locally applied heat treatment on developing cabernet sauvignon grape berries. *Front. Plant Sci.* **2017**, *8*, 53. [[CrossRef](#)]
60. Cohen, S.D.; Tarara, J.M.; Kennedy, J.A. Assessing the impact of temperature on grape phenolic metabolism. *Anal. Chim. Acta* **2008**, *621*, 57–67. [[CrossRef](#)]
61. Yamane, T.; Shibayama, K. Effects of Trunk Girdling and Crop Load Levels on Fruit Quality and Root Elongation in ‘Aki Queen’ Grapevines. *J. Jpn. Soc. Hortic. Sci.* **2006**, *75*, 439–444. [[CrossRef](#)]
62. Yamane, T.; Jeong, S.T.; Goto-Yamamoto, N.; Koshita, Y.; Kobayashi, S. Effects of temperature on anthocyanin biosynthesis in grape berry skins. *Am. J. Enol. Vitic.* **2006**, *57*, 54–59. [[CrossRef](#)]
63. Ori, K.; Goto-Yamamoto, N.; Kitayama, M.; Hashizume, K. Loss of anthocyanins in red-wine grape under high temperature. *J. Exp. Bot.* **2007**, *58*, 1935–1945. [[CrossRef](#)]
64. Kliewer, W.M.; Antcliff, A.J. Influence of defoliation, leaf darkening, and cluster shading on the growth and composition of ‘Sultana’ grapes. *Am. J. Enol. Vitic.* **1970**, *21*, 26–36.
65. Blancquaert, E.H.; Oberholster, A.; Da-Silva, J.M.R.; Deloire, A.J. Effects of abiotic factors on phenolic compounds in the grape berry—A review. *S. Afr. J. Enol. Vitic.* **2018**, *40*, 1–14. [[CrossRef](#)]
66. Belancic, A.; Agosin, E.; Ibacache, A.; Bordeu, E.; Baumes, R.; Razungles, A.; Bayonove, C. Influence of Sun Exposure on the Aromatic Composition of Chilean Muscat Grape Cultivars Moscatel de Alejandria and *Moscatel rosada*. *Am. J. Enol. Vitic.* **1997**, *48*, 181–186. [[CrossRef](#)]
67. Marais, J.; Versini, G.; Van Wyk, C.; Rapp, A. Effect of Region on Free and Bound Monoterpene and C13-N orisoprenoid Concentrations in Weisser Riesling Wines. *S. Afr. J. Enol. Vitic.* **1992**, *13*, 71–77. [[CrossRef](#)]
68. Guillaumie, S.; Ilg, A.; Rety, S.; Brette, M.; Trossat-Magnin, C.; Decroocq, S.; Léon, C.; Keime, C.; Ye, T.; Baltenweck-Guyot, R.; et al. Genetic Analysis of the Biosynthesis of 2-Methoxy-3-Isobutylpyrazine, a Major Grape-Derived Aroma Compound Impacting Wine Quality. *Plant Physiol.* **2013**, *162*, 604–615. [[CrossRef](#)]



69. Falcão, L.D.; Brighenti, E.; Rosier, J.-P.; Bordignon-Luiz, M.T.; Burin, V.M.; Chaves, E.S.; Vieira, H.J. Vineyard altitude and mesoclimate influences on the phenology and maturation of Cabernet-Sauvignon grapes from Santa Catarina State. *OENO One* **2010**, *44*, 135. [[CrossRef](#)]
70. Intrieri, C.; Silvestroni, O.; Rebucci, B.; Poni, S.; Filippetti, I. The effects of row orientation on growth, yield, quality and dry matter partitioning in chardonnay vines trained to simple curtain and spur-pruned cordon. In Proceedings of the 11th Meeting of the Study Group for Vine Training Systems, Sicily, Italy, 6–12 June 1999; pp. 254–262.
71. Hunter, J.J.; Volschenk, C.G.; Zorer, R. Vineyard row orientation of *Vitis vinifera* L. cv. Shiraz/101-14 Mgt: Climatic profiles and vine physiological status. *Agric. For. Meteorol.* **2016**, *228–229*, 104–119. [[CrossRef](#)]
72. Hunter, J.J.; Volschenk, C.G.; Booyse, M. Vineyard row orientation and grape ripeness level effects on vegetative and reproductive growth characteristics of *Vitis vinifera* L. cv. Shiraz/101-14 Mgt. *Eur. J. Agron.* **2017**, *84*, 47–57. [[CrossRef](#)]
73. Hunter, J.J.; Volschenk, C.G. Chemical composition and sensory properties of non-wooded and wooded Shiraz (*Vitis vinifera* L.) wine as affected by vineyard row orientation and grape ripeness level. *J. Sci. Food Agric.* **2017**, *98*, 2689–2704. [[CrossRef](#)] [[PubMed](#)]
74. Christensen, L.P. Rootstock selection. In *Wine Grape Varieties in California*; Bettiga, L.J., Golino, D.A., McGourty, G., Smith, R.J., Verdegaal, P.S., Weber, E., Eds.; University of California Agriculture and Natural Resources: Davis, CA, USA, 2003; pp. 12–15.
75. Zasada, I.A.; Howland, A.D.; Peetz, A.B.; East, K.; Moyer, M. *Vitis* spp. Rootstocks Are Poor Hosts for Meloidogyne hapla, a Nematode Commonly Found in Washington Winegrape Vineyards. *Am. J. Enol. Vitic.* **2018**, *70*, 1–8. [[CrossRef](#)]
76. Zhou-Tsang, A.; Wu, Y.; Henderson, S.W.; Walker, A.R.; Borneman, A.R.; Walker, R.R.; Gilliam, M. Grapevine Salt Tolerance. *Aust. J. Grape Wine Res.* **2021**, *27*, 149–168. [[CrossRef](#)]
77. Corso, M.; Bonghi, C. Grapevine rootstock effects on abiotic stress tolerance. *Plant Sci. Today* **2014**, *1*, 108–113. [[CrossRef](#)]
78. Monteverde, C.; De Sales, F. Impacts of global warming on southern California’s winegrape climate suitability. *Adv. Clim. Chang. Res.* **2020**, *11*, 279–293. [[CrossRef](#)]
79. Chrysargyris, A.; Xylia, P.; Litskas, V.; Mandoulaki, A.; Antoniou, D.; Boyias, T.; Stavrinides, M.; Tzortzakis, N. Drought stress and soil management practices in grapevines in Cyprus under the threat of climate change. *J. Water Clim. Chang.* **2018**, *9*, 703–714. [[CrossRef](#)]
80. Wang, Z.-L.; Xue, T.-T.; Gao, F.-F.; Zhang, L.; Han, X.; Wang, Y.; Hui, M.; Wu, D.; Li, H.; Wang, H. Intraspecific recurrent selection in *V. vinifera*: An effective method for breeding of high quality, disease-, cold-, and drought-resistant grapes. *Euphytica* **2021**, *217*, 1–15. [[CrossRef](#)]
81. Martínez de Toda, F.; García, J.; Balda, P. Adaptación al calentamiento climático de veinte variedades de vid, minoritarias de la DOCa Rioja, por su potencial de acidez. *Zubia* **2017**, *29*, 83–94.
82. Van Leeuwen, C.; Roby, J.P.; Alonso-Villaverde, V.; Gindro, K. Impact of clonal variability in *Vitis vinifera* Cabernet Franc on grape composition, wine quality, leaf blade stilbene content and downy mildew resistance. *J. Agric. Food Chem.* **2013**, *61*, 19–24. [[CrossRef](#)] [[PubMed](#)]
83. Van Leeuwen, C.; Destrac, A. Modified grape composition under climate change conditions requires adaptations in the vineyard. *OENO One* **2017**, *51*, 147–154. [[CrossRef](#)]
84. Gladstone, E.A.; Dokoozlian, N.K. Influence of leaf area density and trellis/training systems on the microclimate within grapevine canopies. *Vitis* **2003**, *32*, 123–131. [[CrossRef](#)]
85. Cartechini, A.; Palliotti, A.; Lungarotti, C. Influence of timing of summer hedging on yield and grape quality in some red and white grapevine cultivars. *Acta Hort.* **1998**, *512*, 101–110. [[CrossRef](#)]
86. Stoll, M.; Scheidweiler, M.; Lafontaine, M.; Schultz, H.R. Possibilities to reduce the velocity of berry maturation through various leaf area to fruit ratio modifications in *Vitis vinifera* L. *Riesling*. *Progrès Agric. Vitic.* **2009**, *127*, 68–71.
87. Filippetti, I.; Allegro, G.; Mohaved, N.; Pastore, C.; Valentini, G.; Intrieri, C. Effects of late-season source limitations induced by trimming and antitranspirants canopy spray on grape composition during ripening in *Vitis vinifera* cv. Sangiovese. In Proceedings of the 17th International GIESCO Symposium, Asti-Alba, Italy, 29 August–2 September 2011; pp. 259–262.
88. Palliotti, A.; Panara, F.; Silvestroni, O.; Lanari, V.; Sabbatini, P.; Howell, G.; Gatti, M.; Poni, S. Influence of mechanical postveraison leaf removal apical to the cluster zone on delay of fruit ripening in Sangiovese (*Vitis vinifera* L.) grapevines. *Aust. J. Grape Wine Res.* **2013**, *19*, 369–377. [[CrossRef](#)]
89. Williams, L.E. Determination of evapotranspiration and crop coefficients for a chardonnay vineyard located in a cool climate. *Am. J. Enol. Vitic.* **2014**, *65*, 159–169. [[CrossRef](#)]
90. Williams, L.E.; Baeza, P. Relationships among ambient temperature and vapor pressure deficit and leaf and stem water potentials of fully irrigated, field-grown grapevines. *Am. J. Enol. Vitic.* **2007**, *58*, 173–181. [[CrossRef](#)]
91. Reddy, R.N. *Irrigation Engineering*; Gene-Tech: San Francisco, CA, USA, 2010; p. 69.
92. Biswas, R.K. *Drip and Sprinkler Irrigation*; NIPA: Pitam Pura, New Delhi, India, 2015; pp. 8, 142.
93. Dry, P.R.; Loveys, B.R.; McCarthy, M.G.; Stoll, M. Strategic irrigation management in Australian vineyards. *J. Int. Sci. Vigne Vin* **2001**, *35*, 129–139. [[CrossRef](#)]
94. Bellvert, J.; Mata, M.; Vallverdú, X.C.; Paris, C.; Marsal, J. Optimizing precision irrigation of a vineyard to improve water use efficiency and profitability by using a decision-oriented vine water consumption model. *Precis. Agric.* **2021**, *22*, 319–341. [[CrossRef](#)]
95. Xiang, H.; Tian, L. Development of a low-cost agricultural remote sensing system based on an autonomous unmanned aerial vehicle (UAV). *Biosyst. Eng.* **2011**, *108*, 174–190. [[CrossRef](#)]

96. Bellvert, J.; Zarco-Tejada, P.J.; Girona, J.; Fereres, E. Mapping crop water stress index in a 'Pinot-noir' vineyard: Comparing ground measurements with thermal remote sensing imagery from an unmanned aerial vehicle. *Precis. Agric.* **2014**, *15*, 361–376. [[CrossRef](#)]
97. Di Gennaro, S.F.; Matese, A.; Gioli, B.; Toscano, P.; Zaldei, A.; Palliotti, A.; Genesio, L. Multisensor approach to assess vineyard thermal dynamics combining high-resolution unmanned aerial vehicle (UAV) remote sensing and wireless sensor network (WSN) proximal sensing. *Sci. Hortic.* **2017**, *221*, 83–87. [[CrossRef](#)]
98. Sun, L.; Gao, F.; Anderson, M.C.; Kustas, W.P.; Alsina, M.M.; Sanchez, L.; Sams, B.; McKee, L.; Dulaney, W.; White, W.A.; et al. Daily mapping of 30 m LAI and NDVI for grape yield prediction in California vineyards. *Remote Sens.* **2017**, *9*, 317. [[CrossRef](#)]
99. Pinamonti, F. Compost mulch effects on soil fertility, nutritional status and performance of grapevine. *Nutr. Cycl. Agroecosyst.* **1998**, *51*, 239–248. [[CrossRef](#)]
100. Monteiro, A.; Lopes, C.M. Influence of cover crop on water use and performance of vineyard in Mediterranean Portugal. *Agric. Ecosyst. Environ.* **2007**, *121*, 336–342. [[CrossRef](#)]
101. Ingels, C.A.; Scow, K.M.; Whisson, D.A.; Drenovsky, R.E. Effects of cover crops on grapevines, yield, juice. Composition, soil microbial ecology, and gopher activity. *Am. J. Enol. Vitic.* **2005**, *56*, 19–29. [[CrossRef](#)]
102. Pou, A.; Gulías, J.; Moreno, M.M.; Tomás, M.; Medrano, H.; Cifre, J. Cover cropping in "Vitis vinifera" L. cv. Manto negro vineyards under Mediterranean conditions: Effects on plant vigour, yield and grape quality. *J. Int. Sci. Vigne Du Vin* **2011**, *45*, 223–234. [[CrossRef](#)]
103. Baronti, S.; Vaccari, F.; Miglietta, F.; Calzolari, C.; Lugato, E.; Orlandin, S.; Pini, R.; Zulian, C.; Genesio, L. Impact of biochar application on plant water relations in *Vitis vinifera* (L.). *Eur. J. Agron.* **2014**, *53*, 38–44. [[CrossRef](#)]
104. Genesio, L.; Miglietta, F.; Baronti, S.; Vaccari, F.P. Biochar increases vineyard productivity without affecting grape quality: Results from a four years field experiment in Tuscany. *Agric. Ecosyst. Environ.* **2015**, *201*, 20–25. [[CrossRef](#)]
105. Moral, F.J.; Rebollo, F.J.; Paniagua, L.L.; Garcia-Martin, A. A GIS-based multivariate clustering for characterization and ecoregion mapping from a viticultural perspective. *Span. J. Agric. Res.* **2016**, *14*, e0206. [[CrossRef](#)]
106. Hall, A.; Jones, G.V. Spatial analysis of climate in winegrape-growing regions in Australia. *Aust. J. Grape Wine Res.* **2010**, *16*, 389–404. [[CrossRef](#)]
107. Dilrukshi, E.A.A.; Nagahage, I.S.P.; Fujino, T. Calibration and validation of a low-cost capacitive moisture sensor to integrate the automated soil moisture monitoring system. *Agriculture* **2019**, *9*, 141. [[CrossRef](#)]
108. Stevanato, L.; Baroni, G.; Cohen, Y.; Fontana, C.L.; Gatto, S.; Lunardon, M.; Marinello, F.; Moretto, S.; Morselli, L. A novel cosmic-ray neutron sensor for soil moisture estimation over large areas. *Agriculture* **2019**, *9*, 202. [[CrossRef](#)]
109. Rodriguez-Perez, J.R.; Plant, R.E.; Lambert, J.; Smart, D.R. Using apparent soil electrical conductivity (ECa) to characterize vineyard soils of high clay content. *Precis. Agric.* **2011**, *12*, 775–794. [[CrossRef](#)]
110. Sabir, A.; Sabir, F.; Jawshle, A.I.M. Quality changes in grape berry as affected by the use of different colored shade nets proposed to alleviate the adverse effects of climate change. *Asian J. Agric. Food Sci.* **2020**, *8*. [[CrossRef](#)]
111. Caravia, L.; Collins, C.; Petrie, P.R.; Tyerman, S.D. Application of shade treatments during Shiraz berry ripening to reduce the impact of high temperature. *Aust. J. Grape Wine Res.* **2016**, *22*, 422–437. [[CrossRef](#)]
112. Brito, C.; Dinis, L.T.; Moutinho-Pereira, J.; Correia, C. Kaolin, an emerging tool to alleviate the effects of abiotic stresses on crop performance. *Sci. Hortic.* **2019**, *250*, 310–316. [[CrossRef](#)]
113. Garrido, A.; Seródio, J.; Vos, R.D.; Conde, A.; Cunha, A. Influence of foliar kaolin application and irrigation on photosynthetic activity of grape berries. *Agronomy* **2019**, *9*, 685. [[CrossRef](#)]
114. Dinis, L.T.; Ferreira, H.; Pinto, G.; Bernardo, S.; Correia, C.M.; Moutinho-Pereira, J. Kaolin-based, foliar reflective film protects photosystem II structure and function in grapevine leaves exposed to heat and high solar radiation. *Photosynthetica* **2016**, *54*, 47–55. [[CrossRef](#)]
115. Gatti, M.; Galbignani, G.; Garavani, A.; Bernizzoni, F.; Tombesi, S.; Palliotti, A.; Poni, S. Manipulation of ripening via antitranspirants in cv. Barbera (*Vitis vinifera* L.). *Aust. J. Grape Wine Res.* **2016**, *22*, 245–255. [[CrossRef](#)]
116. Brillante, L.; Belfiore, N.; Gaiotti, F.; Lovat, L.; Sansone, L.; Poni, S.; Tomasi, D. Comparing kaolin and pinolene to improve sustainable grapevine production during drought. *PLoS ONE* **2016**, *11*, e0156631. [[CrossRef](#)]
117. Górník, K.; Grzesik, M.; Romanowska-Duda, B. The effect of chitosan on rooting of grapevine cuttings and on subsequent plant growth under drought and temperature stress. *J. Fruit Ornamental. Plant Res.* **2008**, *16*, 333–343.
118. Singh, R.K.; Martins, V.; Soares, B.; Castro, I.; Falco, V. Chitosan application in vineyards (*Vitis vinifera* L. cv. Tinto Cão) induces accumulation of anthocyanins and other phenolics in berries, mediated by modifications in the transcription of secondary metabolism genes. *Int. J. Mol. Sci.* **2020**, *21*, 306. [[CrossRef](#)] [[PubMed](#)]
119. Silva, V.; Singh, R.K.; Gomes, N.; Soares, B.G.; Silva, A.; Falco, V.; Capita, R.; Alonso-Calleja, C.; Pereira, J.E.; Amaral, J.S.; et al. Comparative insight upon chitosan solution and chitosan nanoparticles application on the phenolic content, antioxidant and antimicrobial activities of individual grape components of Sousão variety. *Antioxidants* **2020**, *9*, 178. [[CrossRef](#)]
120. Gatti, M.; Pirez, F.J.; Frioni, T.; Squeri, C.; Poni, S. Calibrated, delayed-cane winter pruning controls yield and significantly postpones berry ripening parameters in *Vitis vinifera* L. cv. Pinot Noir. *Aust. J. Grape Wine Res.* **2018**, *24*, 305–316. [[CrossRef](#)]
121. Buesa, I.; Yeves, A.; Sanz, F.; Chirivella, C.; Intrigliolo, D.S. Effect of delaying winter pruning of Bobal and Tempranillo grapevines on vine performance, grape and wine composition. *Aust. J. Grape Wine Res.* **2021**, *27*, 94–105. [[CrossRef](#)]

122. Moran, M.A.; Bastian, S.E.; Petrie, P.R.; Sadras, V.O. Late pruning impacts on chemical and sensory attributes of Shiraz wine. *Aust. J. Grape Wine Res.* **2018**, *24*, 469–477. [[CrossRef](#)]
123. Pomarici, E.; Vecchio, R.; Mariani, A. Wineries' perception of sustainability costs and benefits: An exploratory study in California. *Sustainability* **2015**, *7*, 16164–16174. [[CrossRef](#)]
124. Hillis, V.; Lubell, M.; Hoffman, M. Sustainability partnerships and viticulture management in California. *J. Environ. Manag.* **2018**, *217*, 214–225. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.



## OPEN ACCESS

EDITED BY  
Susana M.P. Carvalho,  
University of Porto, Portugal

REVIEWED BY  
Georgios Koubouris,  
Hellenic Agricultural Organization  
DEMETER, Greece  
Franco Meggio,  
University of Padua, Italy

\*CORRESPONDENCE  
Suzy Y. Rogiers  
✉ suzy.rogiers@dpi.nsw.gov.au

SPECIALTY SECTION  
This article was submitted to  
Crop and Product Physiology,  
a section of the journal  
Frontiers in Plant Science

RECEIVED 10 November 2022

ACCEPTED 05 December 2022

PUBLISHED 21 December 2022

CITATION  
Rogiers SY, Greer DH, Liu Y, Baby T  
and Xiao Z (2022) Impact of climate  
change on grape berry ripening: An  
assessment of adaptation strategies for  
the Australian vineyard.  
*Front. Plant Sci.* 13:1094633.  
doi: 10.3389/fpls.2022.1094633

COPYRIGHT  
© 2022 Rogiers, Greer, Liu, Baby and  
Xiao. This is an open-access article  
distributed under the terms of the  
[Creative Commons Attribution License  
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or  
reproduction in other forums is  
permitted, provided the original  
author(s) and the copyright owner(s)  
are credited and that the original  
publication in this journal is cited, in  
accordance with accepted academic  
practice. No use, distribution or  
reproduction is permitted which does  
not comply with these terms.

# Impact of climate change on grape berry ripening: An assessment of adaptation strategies for the Australian vineyard

Suzy Y. Rogiers<sup>1,2,3\*</sup>, Dennis H. Greer<sup>3</sup>, Yin Liu<sup>2,3,4</sup>, Tintu Baby<sup>3</sup>  
and Zeyu Xiao<sup>2,3</sup>

<sup>1</sup>New South Wales Department of Primary Industries, Wollongbar, NSW, Australia, <sup>2</sup>Australian Research Council Training Centre for Innovative Wine Production, Urrbrae, SA, Australia, <sup>3</sup>Gulbali Institute, Charles Sturt University, Wagga Wagga, NSW, Australia, <sup>4</sup>School of Agriculture Environmental and Veterinary Science, Charles Sturt University, Wagga Wagga, NSW, Australia

Compressed vintages, high alcohol and low wine acidity are but a few repercussions of climate change effects on Australian viticulture. While warm and cool growing regions may have different practical concerns related to climate change, they both experience altered berry and must composition and potentially reduced desirable wine characteristics and market value. Storms, drought and uncertain water supplies combined with excessive heat not only depress vine productivity through altered physiology but can have direct consequences on the fruit. Sunburn, shrivelling and altered sugar-flavour-aroma balance are becoming more prevalent while bushfires can result in smoke taint. Moreover, distorted pest and disease cycles and changes in pathogen geographical distribution have altered biotic stress dynamics that require novel management strategies. A multipronged approach to address these challenges may include alternative cultivars and rootstocks or changing geographic location. In addition, modifying and incorporating novel irrigation regimes, vine architecture and canopy manipulation, vineyard floor management, soil amendments and foliar products such as antitranspirants and other film-forming barriers are potential levers that can be used to manage the effects of climate change. The adoption of technology into the vineyard including weather, plant and soil sensors are giving viticulturists extra tools to make quick decisions, while satellite and airborne remote sensing allow the adoption of precision farming. A coherent and comprehensive approach to climate risk management, with consideration of the environment, ensures that optimum production and exceptional fruit quality is maintained. We review the preliminary findings and feasibility of these new strategies in the Australian context.

## KEYWORDS

viticulture, management systems, adaptation, drought, heat stress, irrigation, soil health, *Vitis vinifera*



## 1 Emerging challenges

Terroir is the result of an interaction between climate, soil, landscape characteristics, topography and biodiversity for a particular cultivar within the vineyard and aside from the inherent natural environment, it also encompasses the cultural management of a site. It refers to “the interactions between the identifiable physical and biological environment and applied viticultural practices, providing distinctive characteristics for the products originating from this area” (OIV, 2010). Of these factors, temperature is undoubtedly a strong driving force for vine and fruit development (Jones and Davis, 2000; Jones and Alves, 2012). The Mediterranean climate is considered ideal for viticulture. Hence, warm, dry summers are accompanied by cool, wet winters and these combinations of temperature, light and water drive the desirable evolution of berry aroma, colour and flavour in hundreds of grape cultivars (Keller, 2010). That said, grapevines are grown with economic success across a range of climatic zones, resulting in highly diverse wine styles (Van Leeuwen et al., 2004). However, heat, drought, wildfires, excessive rain events and increased pest and disease pressure are posing new challenges for viticulture. Additionally, many viticultural regions are consistently experiencing a general phenological advancement in flowering, veraison and maturity. These trends and emerging challenges have been, at least partially, attributed to a changing climate (Caffarra and Eccel, 2010; Bonnefoy et al., 2013; Malheiro et al., 2013; Cola et al., 2017; Jarvis et al., 2017; Alikadic et al., 2019; Cameron et al., 2022).

Alongside the ambient temperature rise, Australia has been subjected to more extreme climatic events like heatwaves, wildfires and shifts in the timing and volume of rain (Abram et al., 2021). In February 2009, Eastern Australia witnessed extreme temperatures and one of the most devastating fires on record. Another extreme heat event occurred in December 2012–January 2013 across 70% of Australia with temperature records in every state and territory. The maximum temperature averaged across Australia was the hottest ever recorded at 40.3°C. Furthermore, a frost in November 2017 across South Australia wiped out 30,000 tonnes of grapes, while hail during flowering in 2017 in the Riverina region of NSW resulted in complete crop loss. These are genuine current examples of extreme climatic effects on viticultural productivity. Additionally, the unprecedented 2019–20 Black Summer bushfires occurred during record breaking temperatures and very low rainfall. These fires were classified as ‘megafires’ with nearly 19 million hectares of land destroyed and with an extreme impact on biodiversity and at least one billion vertebrate animals lost (Filkov et al., 2020). This disaster also resulted in extensive social and economic impacts, including smoke-related effects on public health and on vineyards in many of the affected wine-growing regions.

Australian viticulture stretches from the southernmost latitudes of Victoria, South Australia and Western Australia northward into New South Wales and Queensland and it has been projected that by 2050 a warming climate will reduce the suitable area available for cultivation by 22–73% (Hannah et al., 2013); this is despite factoring in estimates of emerging novel areas for cultivation. These projections, however, are based on existing cultivars and current management strategies. The adoption of new cultivars and integration of adaptive measures to tackle these mounting climate pressures will certainly lessen the severity of these predictions (Van Leeuwen et al., 2013; Mosedale et al., 2016). However, research and knowledge transfer are required now allowing appropriate practical strategies to be implemented (Santos et al., 2020).

### 1.1 Heat stress

The growing season mean temperature is an important driver for root, canopy and reproductive development. Grapevines can be cultivated in average growing season temperatures with a lower and upper threshold of 12–13°C and 22–24°C (Schultz and Jones, 2010). Once endo- and eco-dormancy are broken, temperatures above 7–10°C will drive budburst (Amerine and Winkler, 1944), followed by new vegetative growth and the emergence of the inflorescences. Warm temperatures will encourage development of the canopy, but later in the season, if temperatures reach above 35°C, heat stress will impact on the physiology of the vine. Temperatures can exceed 40°C for prolonged periods in Australia and this will have an impact on carbon assimilation and thus sugar accumulation by grapes (Greer and Weston, 2010). The process of photosynthesis is vulnerable to temperatures due reductions in carboxylation by ribulose 1,5-bisphosphate carboxylase/oxygenase (Rubisco) and the regeneration of ribulose 1,5-bisphosphate (RuBP) (Greer and Weedon, 2012). Heat is also often combined with low humidity and therefore, to prevent water stress, stomata will close.

The berries themselves are vulnerable to heat stress with repercussions on berry composition and wine quality. Aside from impact on primary metabolites such as sugars (Greer and Weston, 2010; Pillet et al., 2012), organic acids (Sweetman et al., 2014) and amino acids (Lecourieux et al., 2017), secondary metabolites responsible for the sensory attributes are also altered. Flavonoids are affected by high temperatures, but outcomes depend on the heat intensity, duration, phenological stage and genotype (Gouot et al., 2019). Anthocyanin changes were attributed to a combination of changes in gene expression, enzyme activity, degradation and relocation (Gouot et al., 2019). Additionally, molecular data pointed toward cell wall changes (Lecourieux et al., 2017) with potential ramifications on berry texture. Cell walls of berry skin cells became more extensible

perhaps to enable berry contraction and expansion to occur between the day and night extremes. Elevated canopy temperature can accelerate late ripening mesocarp cell death and berry desiccation (i.e. shrivelling) in susceptible cultivars (Bonada et al., 2013a). High temperature may exacerbate late ripening cell death by increasing respiration within the berries. Due to reduced oxygen diffusivity in late ripening berries, the mesocarp can become hypoxic, resulting in localised anaerobic fermentation which leads to ethanol accumulation (Xiao et al., 2018b). Excessive radiation, in combination with heat, can result in sunburn, especially on the western side of the canopy (Greer and Weedon, 2013; Greer et al., 2006). Sun exposure and the microclimate of the bunch will dictate berry temperature and, therefore, careful manipulation of the canopy may play an important role in berry attributes, and indeed berry survival during an extreme temperature event.

## 1.2 Water deficits

Water deficits can impact on vegetative growth, inflorescence development, berry set and berry development, though dependent on the phenological stage, severity and duration of the water deficit (Hardie and Considine, 1976). Photosynthesis occurs at the expense of transpiration, with the efficiency of carbon gain dependent on cultivar. Mild deficits may result in stomatal limitations, but more severe deficits will result in non-stomatal limitations, affecting photosynthesis (Lovisolo et al., 2010). Drought tolerant cultivars with adequate stomatal control and water-use efficiency are a priority; as leaf water potential declines, hydraulic conductivity can be maintained through stomatal closure. Shiraz and Semillon, for instance, are situated on the anisohydric end of the isohydric to anisohydric spectrum and can suffer from water stress (Schultz, 2003; Rogiers et al., 2009). However, cultivar behaviour is not always clearly aligned to one or the other end of the spectrum, and may be inconsistent (Charrier et al., 2018). Stomatal regulation and transpiration are strongly influenced by the environment, even during the night (Rogiers and Clarke, 2013). However, unfavourable conditions may not necessarily increase the risk of plant water stress (Dayer et al., 2021). For instance, the conditions preceding the water stress and the rate at which the water stress is imposed will influence the outcomes (Morabito et al., 2022). The age of the plant and the phenological stage also play a role. Water stress will hamper cell division and elongation, and thus overall growth and reproductive development. To cope with the increase in osmotic stress as a result of dehydration, cells can accumulate osmoprotectants such as sugars and amino acids. Leaf petiole ABA (abscisic acid) concentrations are positively correlated with root sucrose concentrations in water stressed Grenache and Semillon,

indicative of integration between ABA signalling and carbohydrate metabolism (Rogiers et al., 2011a) during water stress conditions. Water stress can also accelerate late ripening mesocarp cell death and exacerbate the berry dehydration effect in prone cultivars (Fuentes et al., 2010; Bonada et al., 2013b), resulting in increased sugar concentration (Sadras and McCarthy, 2007; Caravia et al., 2016), altered chemical composition (Šuklje et al., 2016) and sensory characteristics (Bonada et al., 2013b) of the berry.

## 1.3 Bushfires

The risk of fire is predicted to intensify in Australia as a result of rising temperature (Williams et al., 2001; Clarke et al., 2013; Abram et al., 2021). The fire season will likely increase in both duration and intensity resulting in bushfires that damage vines directly or compromise the crop through smoke exposure (Summerson et al., 2021). Smoke exposed fruit resulted in undesirable wine aromas with their intensity dependent on the phenological stage, characteristics of the fire and environmental conditions (Kennison et al., 2007; Krstic et al., 2015). The volatile phenols guaiacol, 4-methylguaiacol, *o*-, *m*-, and *p*-cresol; and syringol specifically contributed to the smoky aroma (Jones et al., 2022). These volatile phenols enter the fruit through their cuticle, with little transport from leaves. Once inside the berry, the phenols are glycosylated as a detoxifying mechanism (Dungey et al., 2011; Noestheden et al., 2018; Jones et al., 2022). Both the free and glycosylated forms have repercussions on aroma and flavour (Mayr et al., 2014).

## 1.4 Waterlogging

Heavy rainfall events and flooding are becoming more frequent and more intense globally, including Australia (Hague, 2021). Flooding as a result of excessive rain may lead to plant oxidative stress as a result of hypoxia and/or anoxia. The transport of O<sub>2</sub> from the leaves to the roots becomes insufficient because O<sub>2</sub> is consumed enroute, and there is a large resistance to gas movement in water saturated root conditions. Moreover, soil microorganisms compete with the roots for any remaining oxygen (Sauter, 2013). Once hypoxic conditions are perceived, a cascade of events led by hypoxic genes are switched on in all the plant's organs (Ruperti et al., 2019). Reduced oxygen levels in the roots results in lowered ability for aerobic respiration, however, alcoholic fermentation is able to generate limited energy. Toxic metabolites and reactive oxygen species (ROS) can accumulate under hypoxia or anoxia (Ponnampereuma, 1972). Moreover, the translocation of carbohydrates from the reserve sites in the roots to the rest of the plant may be hampered or rapidly utilised *in*

*situ* (Sauter, 2013). Reduced hydraulic conductance as a result of insufficient aquaporin activity (Colmer and Greenway, 2011) may further result in stomatal closure and consequently wilting. Canopy senescence, decreased root growth and root decay may also ensue. However, under adequate carbon supplies, new adventitious roots can be produced to maintain oxygen delivery (Striegler et al., 1993). In general, roots under hypoxic conditions are less able to take up water and macronutrients (Bailey-Serres and Voesenek, 2010). Soil microorganisms also require oxygen and nitrate availability declines as a result of less microbiological nitrification (Nguyen et al., 2018). The combined carbohydrate and nutrient starvation may eventually result in the death of the vine following flooding-induced hypoxia and anoxia.

## 1.5 Oxidative stress

Biotic and abiotic stress such as heat, drought, waterlogging, UV-B radiation, nutrient imbalances and salinity can result in the overproduction of ROS, leading to oxidative stress. ROS include hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>), superoxide, singlet oxygen, the hydroxyl radical and organic and inorganic peroxides. ROS propagate chain reactions to target nucleic acids, lipids, proteins and other biomolecules causing oxidative damage (Sharma et al., 2012). However, depending on the concentration, ROS are also secondary messengers for cellular processes, including stress responses, and it is the delicate equilibrium between scavenging and production that determine their role (Mittler et al., 2022). Baseline levels of ROS are produced in most cell compartments, including the chloroplast and the mitochondrion in the processes of photosynthesis and respiration, and under normal conditions cellular homeostasis is maintained. Elevated cellular oxidation plays multiple roles in grapevine growth and development. Bud burst was suggested to be associated with a localised modulation of oxidative signalling within the developing cambium and vascular tissues of the enclosed meristem (Meitha et al., 2015). The onset of berry ripening is linked to the accumulation of H<sub>2</sub>O<sub>2</sub> in the skin, in Pinot Noir (Pilati et al., 2014). Under stress conditions, ROS production and accumulation is further enhanced. For instance, under excessive heat, ROS accrue in the cytosol and the nucleus (Babbar et al., 2021). As a result of stomatal closure in response to drought, excess radiative energy can cause oxidation and hence impair the chloroplast, apoplast and cytosol. Because plants are often exposed to several stresses simultaneously (e.g. drought, high light and heat), overproduction of ROS may result in cell death and tissue necrosis. Cellular oxygen sensing (Xiao et al., 2018a; Xiao et al., 2018b), ROS detection technology (Chen and Fluhr, 2018) and molecular tools, including DNA, RNA and proteins, can be implemented to better understand vine response to heat and drought and conditions resulting in oxidative stress (Gomès et al., 2021).

## 1.6 Earlier maturity and decoupling of phenolic from sugar ripeness

Earlier grape maturity occurring than in the recent past is manifesting in many viticulture regions (Duchêne and Schneider, 2005; Wolfe et al., 2005; Etien et al., 2009; Duchêne et al., 2010; Urhausen et al., 2011; Webb et al., 2011; Morales-Castilla et al., 2020). For example, studies found that fruit maturity has advanced by 8 days per decade between 1985 and 2009 in southern Australia (Webb et al., 2011; Webb et al., 2012), while another estimated advancement by 0.5 to 3.1 days per year over 1993 to 2006 (Petrie and Sadras, 2008). In Chardonnay, Shiraz and Cabernet Sauvignon of south-eastern Australia, the early maturity is driven by the early onset of ripening as opposed to faster ripening (Sadras and Petrie, 2011). Earlier ripening may be driven by temperature, but also by vine water stress as dry soils can stimulate the production of the ripening hormone ABA. Optimum sugar levels in these warmer seasons are, however, not always concomitant with similar maturity in colour, flavour or aroma (Kliewer and Torres, 1972; Sadras and Moran, 2012; Sadras et al., 2013), culminating in the suggestion that 'sugar ripeness' is no longer co-ordinated with 'phenolic' ripeness (Van Leeuwen and Seguin, 2006; Goode, 2012). In other words, the anthocyanins and tannins have not matured to the same extent as they would in cooler years characterised by slower sugar accumulation. Berry acidity is another important quality parameter likely to decline in response to warming (Leolini et al., 2019). Warm nights can result in the respiratory loss of malic acid with the effect that the sugar-acid and aroma-acid balance is no longer optimal (Gatti et al., 2015). However, it has been suggested the trends in earlier ripening may not solely be the consequence of climate change, considering that better disease management, fertilizer application and deliberate yield reductions (to achieve an appropriate leaf area to yield ratio) have been implemented over the last decades (Webb et al., 2012). Regardless, the shift in phenology as a result of climate change demands viticultural practices to counteract these negative effects on vine resilience and berry attributes. A study of Australian premium wine found that quality ratings captured the impact of weather on wine prices (Oczkowski, 2016). It was also noted that production occurs at seasonal temperatures that are warmer than optimal. This is likely to be exacerbated with climate warming and will have economic repercussions.

## 1.7 Compressed vintages

A compressed vintage refers to a shorter harvest window for one particular cultivar. It also refers to a narrower harvest window for several cultivars in one particular region (Sadras et al., 2013; Petrie and Sadras, 2016), bringing about competition



for labour, harvesting machinery, cooling capacity, winery processing equipment and tank space (Sadras et al., 2013). The consequence is that harvest compression may result in extended 'hang time' due to the inability of growers and wineries to process large amounts of fruit at once. This extended ripening period can lead to overripe fruit with berry shrivelling (desiccation), the concentration of existing sugars so that alcohol levels rise, loss of acidity and even the degradation of anthocyanins as well as altered flavour and aroma profiles such as a loss of fresh fruit characters (Bonada et al., 2013a; Bonada et al., 2013b; Šuklje et al., 2016). Moreover, alcohol levels have been rising steadily across many wine regions (Alston et al., 2011). In Languedoc, France alcohol increased from 11 to 14%, pH increased from 3.5 to 3.75 and total acidity dropped from 6.0 to 4.5 g/L (Van Leeuwen et al., 2019). Bordeaux wines typically were 12.5% in the 1980s but today have risen to 16%. In Australia, alcohol concentrations for red wines have increased from 12.4% in 1984 to a peak of 14.5% in 2005 (Godden and Muhlack, 2010). These high alcohol levels are considered less 'food-friendly' (Jones, 2010; Jones, 2012), and considering the societal issues associated with high alcohol and growing trend for healthier lifestyles, consumers are expressing an interest in lower alcohol levels (Saliba et al., 2013).

## 1.8 Pest and disease pressure

Shifts in phenology and the geographic distribution of grapevine pest insects are taking place as a result of direct and indirect effects of climate change (Salinari et al., 2006; Reineke and Thiéry, 2016). The spatial and temporal distribution of insects and pathogens is largely determined by temperature, light and water considering these factors control their growth and development (Rosenzweig et al., 2001). These factors also modify the physiology and resilience of the grapevine as the host. Drought exacerbates underlying issues such as trunk diseases (*Eutypa* and *Botryosphaeria*), nematodes and borers while wet conditions are conducive to increased disease activity such as powdery and downy mildew, *Botrytis* and other bunch rots (Salinari et al., 2006; Steel et al., 2011; Galarneau et al., 2019). Drier conditions and reduced leaf wetness may lead to decreased infections; however, the accompanying warmer springtime temperatures will negate this due to earlier infections (Salinari et al., 2006). Because grapevines may become vulnerable to new pathogens as they spread geographically, regular monitoring and an adaptive environmentally conscious preventative program will be essential.

## 1.9 Changing soil properties and dynamics

The physicochemical and biological properties of the soil have a strong influence on soil quality and functionality, and

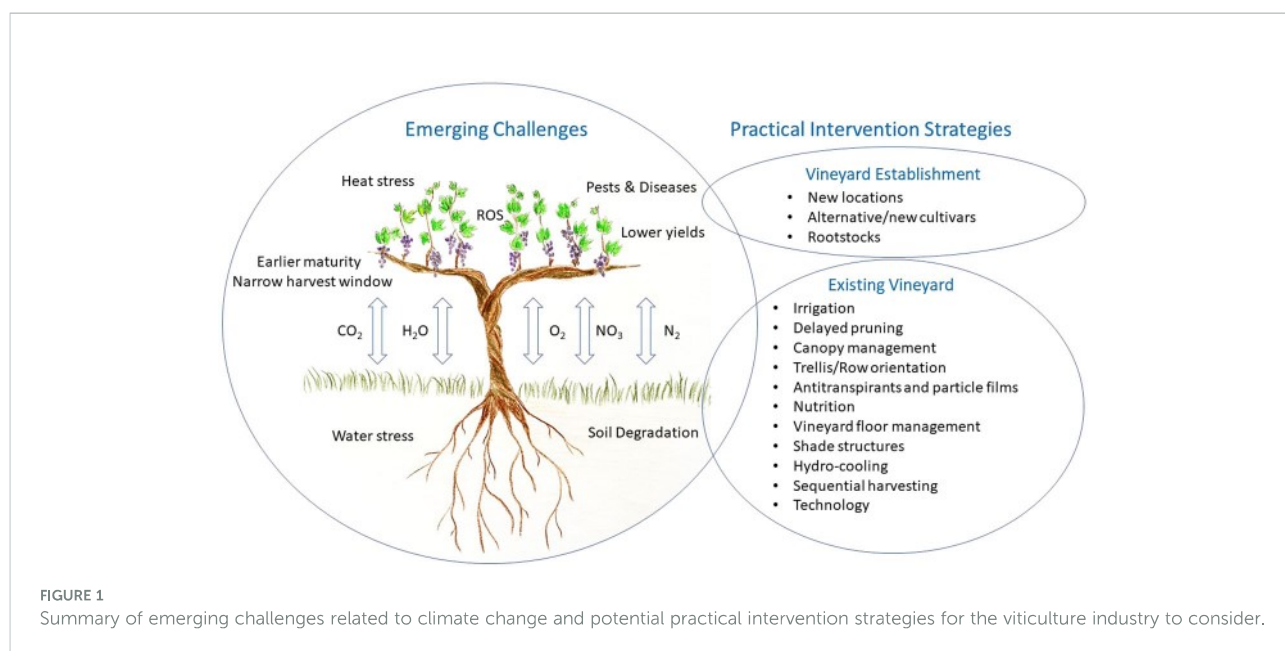
these impact on vine physiology and therefore yield and grape quality (Cataldo et al., 2021). Soil is a living biological entity with multiple functions beyond crop productivity (Doran, 2002; Lal, 2016). The soil contributes to water, carbon and mineral cycles, maintains essential ecosystem functionality by maintaining air and water quality, and providing a habitat for biodiversity (Yang et al., 2020). Analogous to rising air temperatures, the temperature of the soil has increased over the past decades in certain regions (Zhang et al., 2005), and high air temperatures have been associated with record soil temperatures (Ooi et al., 2009). To what extent cover crops ameliorate the relationship between air and soil temperature need investigation. However, soil temperature is an important driver for microbial activity (Bradford et al., 2019), carbon sequestration/release (Huang et al., 2018), greenhouse gas emissions (Smith et al., 2018), and below- and above-ground growth (Rogiers et al., 2014). Warmer soils can initiate earlier budbreak (Rogiers et al., 2014) and this may increase susceptibility of vine shoot growth to late-spring frosts.

The increasing variability in the quantity, intensity of rain and other extreme weather events also influence soil health (Allen et al., 2011). The physical loss of soil through mechanical cultivation and displacement through erosion is likely to be exacerbated by heavy rains. Prolonged drought will impact on water infiltration, soil organic matter and carbon sequestration. Additionally, low soil moisture decreases N availability in the upper soil horizons (Curtin et al., 2012). N is important to overall growth and berry composition such as aromatic secondary metabolites as well as yeast assimilable nitrogen (YAN) for fermentation. Drought will also influence the species and quantity of functional micro-organisms present because water acts as a resource, solvent and transport medium (Schimel, 2018). Aside from the soil itself, the composition and quantity of root exudates are affected by environmental parameters such as temperature and water, and these interact with the root microbiome to influence nutrient cycling, organic matter decomposition and plant growth regulation (Yang et al., 2020).

## 2 Intervention strategies

Without intervention, the combination of these stresses may lead to greater yield variability, sub-optimal berry composition and inconsistency in wine style and typicity. Aside from accessing cooler regions, these approaches could include: alternative cultivars and rootstocks, efficient irrigation strategies, delayed pruning, and consideration of row orientation, training systems along with canopy manipulation, plant tissue films, sustainable vineyard floor management and sequential harvesting (Figure 1). These strategies will need to be vineyard specific to match the topography, environmental conditions, cultivar, and style of wine desired, and it is likely





that a changing combination will be required depending on the urgency of immediate threats versus addressing longer term sustainability. Additionally, grower, community and consumer values/pressures will need to be considered along with labour availability and other socio-economic factors. Most importantly, any adaptation measure should consider potential short- and long- term impacts on the environment. Sustainable productivity will require a balanced approach with environmental stewardship and the preservation of our resources for future generations (Doran, 2002).

## 2.1 Accessing cooler sites

Suitability maps based on bioclimatic indices have been developed for future climate scenarios. Of the 61 wine regions in Australia, it is expected that quality wine grape production will be affected in 21 regions by 2070 if appropriate adaptations are not implemented (Hall and Jones, 2009). The investment into cooler regions is already apparent, particularly in Tasmania where wine production is increasing by 10% each year.

Moving vineyards to cooler elevations is another likely scenario. Higher altitudes have lower night temperatures, potentially resulting in higher acidity and aroma in grape berries with lower alcohol in the wine (Van Leeuwen and Senguin, 2006). Delayed budbreak and bloom are also characteristic of the lower temperatures at increased elevations (Falcão et al., 2010). Vineyards that are currently located on sun-exposed slopes, may spread into sites that traditionally avoided because these sites were low in temperature. The particular fine-scale conditions including aspect, slope, wind exposure, soil drainage and proximity to water bodies will impact on

suitability of a site, but often are not captured with macroclimatic modelling (Mosedale et al., 2016; Pipan et al., 2021). Regardless, shifts in location will entail high economic, environmental and social cost. Extension of vineyards into more marginal sites maybe at the expense of loss of natural habitats, impacting on biodiversity and natural water supplies.

## 2.2 Alternative cultivars, clones and rootstocks

Viticulture has the availability of hundreds of cultivars well suited to a wide range of climates, extending from cool, intermediate, warm to hot conditions (Jones, 2006). The conservation of current viticultural regions is possible by exploiting this diverse plant material. Finding the right cultivar or clone for a specific location is complex, however, and is dependent on the cultivar/clone's climate niche and the distinctive wine style that is desired. Any one cultivar can produce quite different wine styles, with cooler climates resulting in lighter, fresher and crisper styles of higher acidity while warmer climates result in wines with more body, colour, higher alcohol and darker fruit characters. Growing a cultivar outside its climatic niche may decrease certainty in consistent productivity, quality as well as wine style. High quality Pinot Noir, for example, has a particularly narrow range of optimal average growing season temperature, at 14–16°C (Jones et al., 2005) and its future suitability to a particular location could be in doubt (Jones and Webb, 2010). A greater effort to match the climate to many of the lesser-known cultivars will greatly enhance the adoption of these cultivars in a changing climate. The allocation of specific rootstocks that are able to shift

phenology and ripening rates (Van Leeuwen and Destrac-Irvine, 2017) will also be valuable but will require site specific research in line with cultivar credentials.

Drought tolerant cultivars may be considered for the future when water is scarce or too expensive. Likewise, late ripening cultivars may be required to address the expected advance of phenology (Duchêne et al., 2010). Choosing a cultivar to plant in a new vineyard today, knowing that it will experience a different climate in 20-30 years, is challenging and is also confounded with changing consumer preferences. Shiraz might be the iconic Australian wine for some time yet, but there is scope to bring in lesser-known cultivars or even newly bred cultivars that are better able to cope with climate change. To spread the risk, some winegrowers are adopting to plant several cultivars and/or rootstocks within a specific location.

Cultivars from countries such as Italy, are currently being planted in regions across Australia and their growth, productivity and wine production potential are being evaluated (Chalmers, 2019). In particular, the cultivar's distinctive character is assessed along with consumer preference studies to better gauge its potential for premium wine production. The breeding of cultivars suited to hot climates from additional ongoing research has yielded interesting cultivars, producing wines with, for example, increased colour intensity that may be used on their own or for blending purposes (Dry et al., 2022). Specifically, disease resistant cultivars bred by CSIRO have been evaluated across both cool and warm growing regions with good performance. Second generation mildew-resistant cultivars is the current focus of the breeding programs and the microvine, a dwarf grapevine mutant that flowers continuously (Dry and Thomas, 2015), will speed up the process considerably.

Aside from climate, soil is an important component of the vine's natural environment and the roots interact directly with the abiotic and biotic characteristics of the soil profile along a range of depths. Rootstocks offer resistance to phylloxera, nematodes and fungal pathogens but they also have varying tolerance to abiotic stresses (Marín et al., 2021). Vines with deeper and more extensive root systems are better able to access available underground water, while others have a greater intrinsic capacity to absorb water as a result of their higher fine root hydraulic conductivity and aquaporin activity (Gambetta et al., 2012). Rootstocks that are able to control vigour, or influence stomatal conductance through chemical or hydraulic signalling, can reduce vine transpiration and maintain plant water status. Drought tolerant rootstocks such as Ramsey, 110 Richter, the new generation M-series rootstocks, and salt-exclusion rootstocks such as 140 Ruggeri and 1103 Paulsen have been adopted to help vines cope with drought and soil salinity (Stevens and Walker, 2002; Pitt et al., 2018; Bianchi et al., 2020). Soil salinity has increased due to irrigation, of concern in the Murray River Valley of Australia, as high salinity impacts severely on growth and productivity. To improve tolerance to flooding, grafting onto Couderc 3309 resulted in less adverse

impacts (Striegler et al., 1993). Rootstocks can also influence phenology and ripening (May, 1994; Walker et al., 2000). Breeding programs can target rootstocks that shift berry ripening into a lower temperature period enabling berry anthocyanin and acidity to be maintained. Trials have been implemented to assess rootstocks for their ability to reduce potassium uptake by vines and berries (Ollat et al., 2015; Walker and Clingeleffer, 2016; Xiao et al., 2020) considering that grape juice pH is correlated with juice potassium and high pH has negative consequences for flavour, wine stability and colour (Rogiers et al., 2017).

## 2.3 Irrigation

Increases in evaporative demand and declines in precipitation as a consequence of climate change result in greater irrigation requirements for the vineyard. However, in Australia, it is expected that the quantity and quality of water available for irrigation will be reduced (Murray-Darling Basin Authority 2010). Drawing water from surface and groundwater sources has impacts on the surrounding natural environment as well as other agricultural industries and thus more judicious use of this precious resource is at the forefront of vineyard managers. The conversion from flood to drip irrigation in Australian grape growing regions has resulted in enormous water savings, and despite the additional energy required to run water pumps, the precision offered by drip irrigation has allowed the implementation of various mild deficit strategies to influence yield, berry composition and wine properties (Edwards and Clingeleffer, 2013). Regulated deficit irrigation (RDI) is another strategy used to limit competition between vegetative and reproductive growth in red cultivars and is usually applied between fruit set and veraison. While RDI and prolonged deficit reduced yield in Cabernet Sauvignon grown in a warm climate, there was no progressive reduction in yield over multiple seasons (Edwards and Clingeleffer, 2013). However, deficit irrigation in combination with elevated temperatures were detrimental to Shiraz wine phenolic substances and sensory traits (Bonada et al., 2015). Applying a water deficit prior to véraison may offset the delay in anthocyanin accumulation that can occur during high temperatures later in the season (Sadras and Moran, 2012). Investment in soil moisture sensors and methods to monitor vine water status will ensure that critical thresholds are not surpassed in managing the vineyard.

From the perspective of balancing canopy management and water-use efficiency, even though water savings are generally beneficial, if canopy development is impaired and bunch over-exposure occurs then it is suggested that sufficient water should be applied early in the season to allow full canopy development. The cooler microclimate because of the shading will lessen the severity of berry sunburn and shrivelling during late ripening.

Hypoxia and cell death in the berries can also occur during the late ripening stage in some cultivars (Xiao et al., 2018b), and this is exacerbated by water stress (Bonada et al., 2013b; Bonada et al., 2015; Xiao et al., 2018a).

The loss of berry cell vitality was characterised using the cell vitality dye (Krasnow et al., 2008; Tilbrook and Tyerman, 2008; Fuentes et al., 2010), which was associated with the reduced electrical impedance of the berries (Caravia et al., 2015). The cell death inhibitor gene *VvBAP1* was associated with increased drought tolerance (Cao et al., 2019). Shiraz is a variety associated with severe hypoxia and cell death in berries (Xiao et al., 2018b), increased hydraulic resistance into the berry after véraison (Tilbrook and Tyerman, 2008), as well as greater propensity for backflow of water from the berry to the vine in the later stage of ripening (Tilbrook and Tyerman, 2009). Due to greater severity of hypoxia and a decline in respiration, the energy status of berries may be curtailed in the late ripening stages. These changes in berry energy status might be correlated with altered metabolism and the greater force to transport solutes into and within the berries. The hydraulic and energy status of the berries, and the vitality of berry cells, may provide more insight into berry ripening, potentially informing future irrigation and fertilisation strategies.

Practical safeguard measures to deal with drought include the use of polymers or physical covers to reduce evaporation from storage dams, prioritising high value blocks when water is limited. Monitoring water demand through soil or plant-based sensors in combination with weather data, will ensure sustainable water use and potentially water savings for heatwave events. Ensuring dripper spacing is frequent enough to prevent alternating wet and dry zones will allow consistent root growth to spread across a wide area.

### 2.3.1 Irrigation during heatwaves

Irrigation is recommended prior to an impending heatwave to ensure that the optimal field capacity is maintained (Webb et al., 2010). Relative to the local climate and historical observations, a three-day period of unusual high maximum and minimum temperatures can be classified as a heatwave event (Nairn and Fawcett, 2015). During the heatwave, the vines should continue to be watered at regular intervals ensuring that transpiration is maintained to cool the canopy and bunches (Hayman et al., 2012). Application of water directly to the canopy is discussed below in 2.6 under hydro-cooling. The timing of the heatwave during the growing season will result in different consequences, with earlier heatwaves lowering fruit set and later heatwaves disrupting ripening. Heatwaves can cause vines to 'shut down'; that is, photosynthesis stops, due to stomatal closure, so that water can be conserved, and thus the rate of ripening slows (Greer and Weedon, 2013). Depending on the timing, severity and duration of the heatwave, it may take several weeks for the vine to resume its normal photosynthetic

activity resulting in overall low-quality fruit. In contrast, it has been suggested that dry soils brought about by higher temperatures and drought may advance grape maturity (Webb et al., 2012) associated with ABA production by the roots (Davies et al., 2000), a hormone correlated with ripening (Wheeler et al., 2009). Moreover, dry soils fluctuate in temperature to a greater extent during the day hence vine roots may experience warmer temperatures (Rogiers et al., 2014) when the soils are dry.

### 2.3.2 Post-harvest irrigation

Post-harvest irrigation has the benefit of prolonging leaf photosynthesis if conditions are warm enough so that vine carbohydrates can be replenished (Loescher et al., 1990). Grapevines rely on the stored reserves during early season root, shoot and inflorescence growth and a long post-harvest period with adequate soil moisture allows carbon capture through photosynthesis and may even encourage a new flush of fine root growth to aid in the uptake of nutrients (Mahmud, 2016). The timing of leaf senescence to some extent determines the length of carbohydrate reserve accumulation during the post-harvest period and this can be considerable in warm-climate viticulture.

### 2.3.3 Dry winters

Rainfall has been declining in late autumn and early winter in south-eastern Australia (Cai and Cowan, 2013) and may have negative impacts during budburst and spring canopy development. A recent study investigated the effects of low winter rainfall on vine growth and wine quality using rainout shelters (Bonada et al., 2017). Vines were irrigated with micro-sprinklers or drippers at different timepoints during the winter and it was evident that waiting to refill the profile until budburst resulted in excessive vegetative growth with negative consequences on yield and wine composition. It is likely that root growth dynamics and root longevity is altered by the changes in precipitation and irrigation patterns and this will have consequences on overall vine performance. Bonada et al. (2017) recommended that the soil profile be maintained throughout winter rather than delaying until spring to irrigate and refill the soil profile.

## 2.4 Delayed pruning

Delayed maturation of the berries into the cooler period of the growing season and decompressing the harvest window can be encouraged by delayed pruning. Applying pruning after budburst can delay véraison, allowing ripening to occur during cooler months. Pruning carried out at the 10 cm shoot length stage, following mechanical winter spur pruning, achieved a reduction in sugar accumulation in Sangiovese berries (Palliotti



et al., 2017). Shiraz berries from vines pruned at the 2-3 leaf stage reached sugar ripeness up to two weeks later than those with winter pruning applied and attained an improved anthocyanin to sugar ratio (Moran et al., 2017). This strategy was favourable for balancing tannin accumulation, colour intensity, fruit aroma and flavours (Moran et al., 2018).

In addition, delayed pruning can minimize the risk of spring frost damage of the young spring shoots (Friend and Trought, 2007) and improve pollination and fertilisation, by inducing flowering to occur in warmer months. Delayed spur-pruning of the previous season's apical nodes inhibits the development of basal nodes through apical dominance. Pruning carried out three months later than the usual winter pruning time increased yield by 60 to 90%, in Merlot, over a three-season trial (Petrie et al., 2017). However, the impact of delayed pruning on yield was considered to be dependent on interrelated factors such as cultivar, seasonal conditions as well as timing and severity of the pruning applied. Hence further development and assessment of this strategy is needed.

## 2.5 Row orientation, training systems and canopy manipulation

### 2.5.1 Vine balance

Balancing yield with new biomass growth is critical to maintaining long-term sustainable productivity. Nutrient and carbohydrate reserves are essential for supporting early spring growth prior to flowering and are subsequently refilled within the vine prior to ripening (Rogiers et al., 2011b). Heat and drought can influence the quantity of the photoassimilate that is fixed and then distributed to the three growing sinks (roots, shoots and bunches), as well as the reserve pool, predominantly sequestered as starch in the vine structural root system (Rogiers et al., 2011a).

Vine balance can be achieved by pruning of the dormant vines and thinning of the growing leaves/shoots. The number of nodes to be retained during winter pruning can affect vigour, yield, fruit composition and long-term productivity of the vines. Factors such as cultivar, climate and other management influences should be considered when determining the number of nodes retained. Keeping long-term records on yield and pruning weight will aid in monitoring vine performance for reliable vine balance assessment and crop load calculation. Leaf removal and shoot thinning can influence ripening of the fruit to varying degrees, depending on the cultivar, growing conditions, timing and severity of the thinning applied (De Bei et al., 2019). Excessive leaf removal around the bunch zone may result in sunburn of berries. Likewise, shoot thinning of Semillon at the 8-9 leaf stage in a hot Australian climate did not alter berry composition, probably because vine balance was not altered (De Bei et al., 2020). Further research on thinning techniques is required to assess the sustainable effects of this strategy.

### 2.5.2 Natural shading

The geometry of the canopy can be manipulated to ensure that the microclimate of the bunch zone is well-suited to the development of desirable berry attributes. Adequate solar radiation is required for optimal photosynthesis, bud fertility, and berry development, including polyphenol accumulation (Berli et al., 2008). However, over-exposure of leaves may result in photo-inhibition of photosynthesis (Greer et al., 1986) and leaf necrosis and mortality. Sunburnt fruit leads to increased bitterness with browning of the wine, and damaged skin can lead to pathogen invasion. Dark cultivars can experience berry temperature 15°C higher than ambient temperature and thus alterations in the primary and secondary metabolites important to wine quality are likely (Bergqvist et al., 2001). The excessive exposure of leaves and fruit to direct incoming and reflected radiation can be addressed by row orientation, vine architecture, planting density and canopy management. The hot afternoon sun can be avoided by orientating rows in the east-west direction as opposed to the north-south direction (Buesa et al., 2020). Allowing canopies to sprawl so that fruit are shaded will minimise sunburn and shrivelling in hot, sunny climates. While the natural shading may decrease air circulation and increase disease pressure, this may not be an issue in dry regions. Because light is important to anthocyanin production, shading may decrease colour in the red cultivars hence the balance between temperature and light will need to be considered. Removal of the leaves above the bunchzone of cvs. Bobal and Tempranillo vines prior to veraison, and grown under mild water stress, limited the accumulation of anthocyanins more than total soluble solids and was thus not recommended as a method to delay harvest (Buesa et al., 2018). Similarly, avoiding shoot trimming or leaf removal will decrease sun damage and also save on labour costs. Vertical Shoot Positioning (VSP) is not recommended in hot, dry conditions and when sun damage is an issue. Foliage wires can, however, be used to minimise canopy displacement caused by wind.

Pergola type systems that allow the spread of leaves above the hanging fruit are ideal for providing shade, and ventilation simultaneously. In contrast, the Gobelet, also known as the bush vine, is an ancient architectural system still used in the dry regions of Spain with reduced leaf area (Van Leeuwen et al., 2019). It, however, has the drawback of low productivity and is not conducive to mechanical harvesting. The bush vine tends to have a short trunk, however, trunk height can be manipulated to raise bunches away from the soil surface where temperatures may be warmer (Van Leeuwen et al., 2019)

## 2.6 Shade structures and hydro-cooling

The interactive effect of temperature and photon flux densities on reproductive development in Shiraz vines was



assessed in response to several densities of shade cloth over three seasons (Abeyasinghe et al., 2018). During severe heat events the 50% shade treatments were able to reduce canopy temperatures significantly, by approximately 4–5°C. While berry growth was delayed by a 50% shade cover, maximum berry size and the rate of sugar accumulation was not affected. It was interesting to note that shade cover up to 50% had no impact on the ripening process when there were no heat events. The berries had lower sugar concentrations due to greater water content and cell vitality was maintained, but most notably the wines from the fruit resulted in lower alcohol without reduced anthocyanin concentrations. Alongside density, the net's colour has an important impact on the spectral wavelengths that the canopy and fruit are exposed to, thus impacting on vegetative growth and reproductive development, including fruit composition (González et al., 2015).

Overhead sprinklers and misting may reduce heat and water stress as they are able to cool both canopies and berries. Semillon is a cultivar that is particularly sensitive to heat stress both during flowering and ripening, with reduced photosynthesis due to both stomatal and nonstomatal limitations and requiring up to 2 weeks to recover (Greer and Weston, 2010; Greer and Weedon, 2013). Hydrocooling, activated at a threshold temperature of 35°C, extended the period for leaf and berry expansion so that berries were larger. Canopy temperatures were lower, net CO<sub>2</sub> assimilation was higher and berry TSS was also slightly elevated (Greer and Weedon, 2014).

The cost and availability of water for hydrocooling may be an issue and adequate planning so that enough water is available for the entire season is required. The increase in disease pressure is likely to be low in heatwave conditions, but if conditions are warm and humid for a protracted period this may become an issue. The water quality should be adequate as saline water will result in leaf burn and defoliation.

## 2.7 Antitranspirants, sunscreens and film-forming barriers

Plant tissue films have been available to viticulture for several decades and have been trialled for their capacity to maintain tissue hydration (e.g. di-1-*p*-menthene: Palliotti et al., 2013; Gatti et al., 2016; Rogiers and Fahey, 2019), prevent sunburn (e.g. kaolin: Dinis et al., 2016; Conde et al., 2016; Conde et al., 2018), decrease disease incidence (e.g. chitosan: Meng et al., 2008) or even to ameliorate smoke taint (e.g. kaolin: Van der Hulst et al., 2019). These films are derived from natural or synthetic origins and, depending on their composition, time of application and concentration, may or may not be beneficial. When compared to protective netting, these particle film-forming and antitranspirant products may reduce fruit quality and consumer acceptance (Brillante et al., 2016). In a smoke

taint study, several materials were tested for their ability to prevent absorption of gaseous phenols but it was noted that they did not provide much protection (Culbert et al., 2021).

## 2.8 Sustainable vineyard floor management

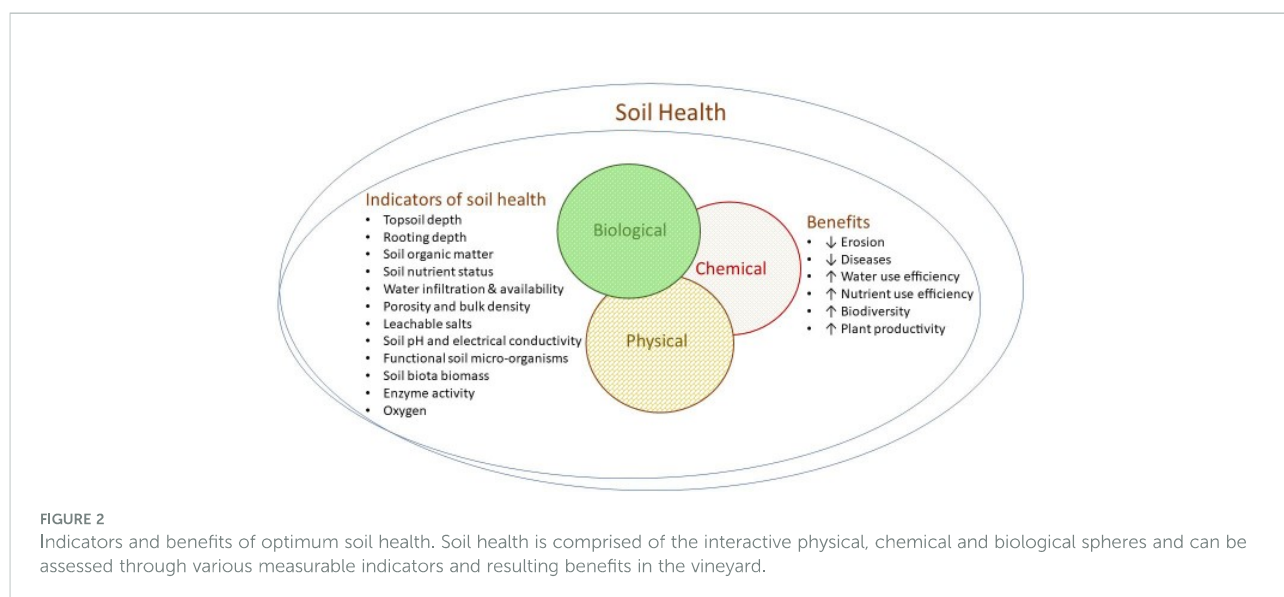
### 2.8.1 Inter-row and under-row vegetation

Bare soil as a result of cultivation contributes significantly to vineyard heating, water and carbon loss, soil erosion and greenhouse gas emissions (Santos et al., 2020). Allowing the resident vegetation to grow or planting cover crops can improve water infiltration, soil organic matter and microbial function (Celette et al., 2008; Steenwerth and Belina, 2008). Inter-row and under-row vegetation can encourage soil biodiversity and lessen disease pressure (Garcia et al., 2018). To minimise the competition for water between the vines and the inter-row plants they can be slashed or allowed to dry (Webb et al., 2010). Moreover, planting native species alongside a vineyard has the benefits of reducing wind, salinity, erosion and providing ecosystem services (Viers et al., 2013).

In regions with adequate rainfall, legume cover crops can supply soil organic matter and biologically fixed N to the grapevine (Ball et al., 2020). Nitrogen application to soil in the form of synthetic fertilisers, manures, composts or mulches can lead to nitrous oxide emissions, a particularly potent greenhouse gas (Longbottom and Petrie, 2015). Targeting the application of N during the phase of active nutrient uptake (before flowering and after harvest when there is a long post-harvest period) will reduce excessive N in the soil and curb nitrous oxide emissions. Employing drip irrigation instead of furrow or flood irrigation will also curb emissions as microbial activity leading to denitrification is reduced (Suddick et al., 2011).

### 2.8.2 Mulches, composts and soil conditioners

Undervine mulch, such as straw or vineyard prunings, has many benefits, aside from weed control, including the amelioration of soil water evaporation (López-Urrea et al., 2020), the regulation of soil temperature and improving biodiversity. Improving the carbon content of the soil by adding mulch and composts not only offsets carbon emissions but also improves soil structure and the water holding capacity of the soil. Likewise, composts increase nutrient reserves, the cation exchange capacity, and reduce the requirement for synthetic fertilisers. Organic matter supports soil biodiversity, which in turn suppresses pathogens and promotes beneficial as opposed to parasitic organisms. Figure 2 provides an overview of factors driving soil health and indicators for soil health assessment. Microbes are especially beneficial at the recycling of nutrients from organic matter and minerals. Mulches applied on top of the



manure or compost further reduces the release of nutrients such as nitrogen to the atmosphere and also helps control root-zone temperature. The characteristics of the soil amendment including its pH, EC, nutrient and heavy metal content should be considered prior to application. Composted grape marc, for instance, contains potassium and this has the potential to increase the pH of the grape must so that acid additions are required in the winery (Rogiers et al., 2017). That said, the addition of grape marc to soil in New Zealand had minor effects on juice K and pH (Mundy and Agnew, 2002).

Organic amendments such as biochar and brown coal waste (BCW) are a source of soil organic matter, which can influence the soil microclimate, microbial community structure, biomass turnover and mineralisation of nutrients (Amoah-Antwi et al., 2020). Coal derived humate is potentially effective as a soil conditioner in improving aggregate stability of acidic and sodic soils against adverse effects of cyclic seasonal wetting and drying conditions (Imbufe et al., 2005). Biochar is a carbon-rich material produced by heating organic material through the pyrolysis process in an oxygen limited environment. Feedstock and pyrolysis conditions (temperature, heating rate, oxygen supply, pressure, residence time, cooling down procedure) can result in biochars with differing chemical and physical properties (Fryda and Visser, 2015). Therefore, they are heterogeneous materials with a diverse range of properties that change over time. Biochar may improve soil function, increase nutrient availability to plants, promote plant productivity, remediate organic/inorganic contaminants, reduce N<sub>2</sub>O emissions, increase soil pH and lead to the net removal of carbon from the atmosphere (Laird, 2008; Kolton et al., 2017). Biochar's macropores filled with adsorbed nutrients create a habitat for beneficial

microorganisms (Atkinson et al., 2010). Biochar can provide protection against both root and foliar plant diseases through the direct manipulation of bacterial communities or by induced plant resilience (Elad et al., 2011). In vineyards, biochar may increase soil fertility, productivity or both (Baronti et al., 2014; Schmidt et al., 2014; Genesio et al., 2015; Mackie et al., 2015; Giagnoni et al., 2019; García-Jaramillo et al., 2021). Notably, a long-term study over ten years showed evidence of increases in soil water content and plant water status (Baronti et al., 2022). However further research is required using different parent materials and pyrolysis conditions to gain a better understanding of the potential for improving plant resilience and maintaining fruit quality under climate change.

## 2.9 Biostimulants

Application of biostimulants to grapevines have a high potential for promoting nutrient uptake, tolerance to biotic and abiotic stress and improving fruit yield and quality (Monteiro et al., 2022). Biostimulants are defined as “any substance or microorganism applied to plants with the aim to enhance nutrition efficiency, abiotic stress tolerance and/or crop quality traits, regardless of its nutrients content” (Du Jardin, 2015). Biostimulants commonly used in vineyards include seaweed extracts, humic substances, chitosan, exudates, and other plant extracts (Cataldo et al., 2022). Seaweed extracts contain macro- and micro-nutrients, amino acids, vitamins, auxins, abscisic acid and cytokinins. Likewise, brown seaweeds (*Ascophyllum nodosum* L.) have been shown to improve tolerance against biotic and abiotic stresses (Salvi et al., 2019; Taskos et al., 2019).

## 2.10 Sequential harvesting

Wine alcohol can be reduced through the simple method of adding water to the juice and amendments to the Australia New Zealand Food Standards Code now allows limited addition of water to high sugar must above 24.3°Brix. Alternatively, alcohol can be removed from wine using vacuum distillation or membrane systems based on reverse osmosis and evaporative perstraction, however the methods are not eco-sustainable and these wines tend to lack aroma through the loss of desirable volatile compounds responsible for fruity and floral characters (Longo et al., 2018). Other strategies include the use of yeast strains tailored for lower ethanol yields (Ristic et al., 2016; Hranilovic et al., 2020). Grapes harvested at a lower sugar level may have insufficient phenolic and aroma attributes that are characteristic of more mature fruit, and they may have too much acid and an undesirable herbaceous character. This can potentially be overcome by blending a low and high alcoholic wine so that acid levels are balanced and ripe flavours mask the unripe attributes. Following two sequential harvests, blending wines made from less ripe grapes with wine vinified from riper fruit resulted in a wine with similar sensory profiles of the later harvested fruit but with a lower alcohol content (Longo et al., 2018). This particular investigation was conducted on Petit Verdot and Verdelho, but the positive outcome certainly warrants further studies on other cultivars.

## 2.11 Technology, sensors, AI and online tools

Information on vineyard status can be obtained through sensors placed in contact with the soil or plant. For instance, soil nutrients can be assessed using ion-selective electrodes and ion-sensitive field-effect transistors (Adamchuck and Rossel, 2010). Direct measurement of plant water status for irrigation scheduling is more accurate than soil moisture monitoring alone as sensor placement may not be representative of full field conditions; moreover, plant measurement integrates evaporative demand. The most accurate method for assessing plant water potential is the pressure chamber, however this method is not conducive to automation. Dendrometers (Clonch et al., 2021), sap flow (Mancha et al., 2021), and acoustic sensors (Oletic et al., 2020) provide information on plant water status, and if temperature sensitivity, set-up inaccuracies and data interpretation difficulties can be overcome, these tools may be useful when correctly calibrated. Additionally, climate sensors alongside leaf temperature and wetness sensors provide continuous data on pest and disease pressure. These data can be captured using wireless sensor

network (WSN) technologies that allow the real-time remote monitoring of several areas within the vineyard.

While contact sensors have their place in vineyard management, contactless sensors are the new focus of the viticulture industry. Satellite imagery allows monitoring at a large scale however is limited by satellite orbit coverage patterns, visiting times, clouds and spectral resolution. Remote sensing using unmanned aerial vehicles (UAVs) or ground vehicles equipped with GPS and an array of sensors is capable of the collection of high-resolution data (mm) on phenology, water status, vigour, pest and disease incidence, weeds, yield and maturity (Seng et al., 2018; Maes and Steppe, 2019; Sassu et al., 2021). Thermal (Far-IR), RGB (red, green, blue), multispectral, hyperspectral sensors and chlorophyll fluorescence systems have all been applied to assess plant status and reproductive parameters (Matese et al., 2015; Tardaguila et al., 2021). Light detection and ranging (LiDAR) is appropriate for the determination of leaf area index and can be used to assess dieback as a result of trunk diseases (Ouyang et al., 2021). Hyperspectral and near-infrared (NIR) spectroscopy have been useful for the assessment of real-time berry composition during ripening (González-Caballero et al., 2012; Power et al., 2019). Additionally, both thermal imaging and NIR spectroscopy have been developed for real-time plant water status assessment (De Bei et al., 2011; Diago et al., 2018), with automated variable drip irrigation to address low-performing areas within a vineyard on the radar. Thus, high-resolution monitoring allows for management that takes into consideration the natural variability in the block (Fuentes and Gago, 2022).

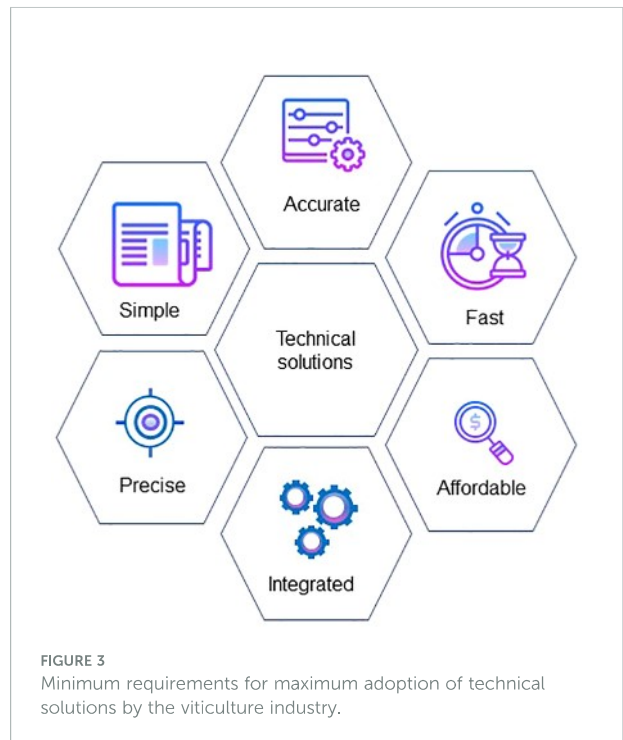
One of the greatest challenges that remains is the processing and integration of the large data sets acquired by these monitoring technologies (Manfreda et al., 2018), allowing a targeted management plan to be effected. AI and machine learning (ML) models are able to overcome these limitations and have been rapidly advancing in viticulture (Seng et al., 2018; Fuentes and Gago, 2022). These approaches can result in intelligent recommendations for efficient water, pesticide and nutrient use not only to target yield and composition but also to deal with temporal and spatial variability. Moreover, robotic innovations are currently under development for the monitoring, pruning, spraying and weeding of vineyards (Matese et al., 2015). Thus, these technologies allow for targeted treatment with improvements in efficiency and reductions in environmental repercussions.

The widespread use of smartphones and apps allow growers rapid and cost-effective assessments of individual vines within the vineyard. A smartphone with an attached thermal camera showed promise for assessing vine water status and to aid in irrigation scheduling (Petrie et al., 2019). Canopy vigour assessment through the VitiCanopy<sup>®</sup> app has GPS capabilities

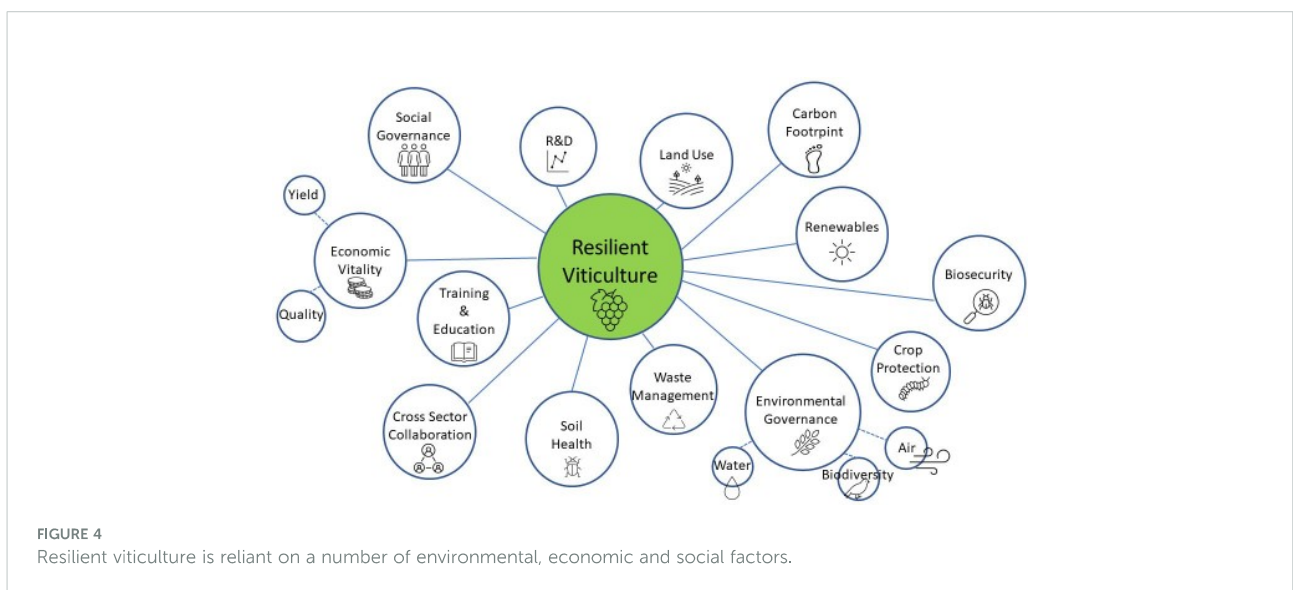
making it appropriate to map spatial canopy architecture (De Bei et al., 2016), and early prediction of yield is possible at flowering (Aquino et al., 2015) and berry (Liu et al., 2020) quantification. ML has also been applied for the detection of smoke contamination in berries using NIR spectroscopy (Fuentes and Tongson, 2018; Summerson et al., 2021). Furthermore, computer vision and deep learning were applied for the detection of nutrient disorders in grapevines (Debnath et al., 2021). RGB images of nutrient deficiencies were taken as symptoms progressed and these were used to develop, train and test models for the incorporation onto a smartphone. Flexible and rapid adjustment to the challenges posed by variable weather patterns can now be addressed through timely online notification systems, including SMS messages to a phone. There has been significant investment in the forecasting of heatwaves and communication to growers so that they are able to prepare for such events. For maximum adoption, technical solutions need to be simple, rapid, affordable, accurate, precise and integrated with other data gathering platforms (Figure 3).

### 3 Summary

The Australian wine industry is already implementing site-specific adaptive measures to deal with unpredictable weather, droughts and rising temperatures. Growers are conscious of their environmental footprint, their long-term sustainability, and are balancing these visions with economic viability (Figure 4). The increasing awareness of sustainable management practises will result in overall improvements in environmental equilibrium and soil health for future generations. As they become more economical and user-



friendly, technical innovations will be readily adopted and integrated with more traditional approaches to vineyard management. However, long-term field trials are required to fine-tune these adaptive strategies to particular situations. Additionally, the inclusion of stakeholders in a co-design framework for future R&D will accelerate the adoption of these mitigation strategies.





## Author contributions

Author contributions are as follows: Conceptualisation, investigation, visualisation, writing- original draft preparation: SR; Review and editing DG, ZX, YL, and TB. All authors have read and agreed to the published version of the manuscript.

## Funding

This work was supported by the Australian Research Council Training Centre for Innovative Wine Production ([www.ARCwinecentre.org.au](http://www.ARCwinecentre.org.au); project number IC170100008), funded by the Australian Government with additional support from Wine Australia, NSW Department of Primary Industries, Gulbali Institute and industry partners.

## References

- Abeyasinghe, S. K., Greer, D. H., and Rogiers, S. Y. (2018). The effect of photon flux densities imposed by selected shade treatments on berry growth and sugar accumulation in *Vitis vinifera* cv. Shiraz under vineyard conditions. *Vitis* 58, 7–16. doi: 10.5073/vitis.2019.58.7-16
- Abram, N. J., Henley, B. J., Sen Gupta, A., Lippmann, T. J., Clarke, H., Dowdy, A. J., et al. (2021). Connections of climate change and variability to large and extreme forest fires in southeast Australia. *Commun. Earth Environ.* 2, 1–17. doi: 10.1038/s43247-020-00065-8
- Adamchuk, V. I., and Rossel, R. A. (2010). "Development of on-the-go proximal soil sensor systems," in *Proximal soil sensing*. Eds. R. Viscarra Rossel, A. McBratney and B. Minasny (Dordrecht: Springer). doi: 10.1007/978-90-481-8859-8\_2
- Alikadic, A., Pertot, I., Eccel, E., Dolci, C., Zarbo, C., Caffarra, A., et al. (2019). The impact of climate change on grapevine phenology and the influence of altitude: A regional study. *Agric. For. Meteorol.* 271, 73–82. doi: 10.1016/j.agrformet.2019.02.030
- Allen, D. E., Singh, B. P., and Dalal, R. C. (2011). "Soil health indicators under climate change: a review of current knowledge," in *Soil health and climate change*. Eds. B. P. Singh, A. L. Cowie and K. Y. Chan (Berlin: Springer-Verlag), 25–45. doi: 10.1007/978-3-642-20256-8\_2
- Alston, J. M., Fuller, K. B., Lapsley, J. T., and Soleas, G. (2011). Too much of a good thing? causes and consequences of increases in sugar content of California wine grapes. *J. Wine Econ.* 6, 135–159. doi: 10.1017/S1931436100001565
- Amerine, M., and Winkler, A. (1944). Composition and quality of musts and wines of California grapes. *Hilgardia* 15, 493–675.
- Amoah-Antwi, C., Kwiatkowska-Malina, J., Thornton, S.F., Fenton, O., Malina, G., and Szara, E. (2020). Restoration of soil quality using biochar and brown coal waste. A review. *Sci. Total Environ.* 722:137852. doi: 10.1016/j.scitotenv.2020.137852
- Aquino, A., Millan, B., Gaston, D., Diago, M. P., and Tardaguila, J. (2015). *vitisFlower*<sup>®</sup>: development and testing of a novel android-smartphone application for assessing the number of grapevine flowers per inflorescence using artificial vision techniques. *Sensors* 15, 21204–21218. doi: 10.3390/s150921204
- Atkinson, C. J., Fitzgerald, J. D., and Hipps, N. A. (2010). Potential mechanisms for achieving agricultural benefits from biochar application to temperate soils: a review. *Plant Soil* 337, 1–18. doi: 10.1007/s11104-010-0464-5
- Babbar, R., Karpinska, B., Grover, A., and Foyer, C. H. (2021). Heat-induced oxidation of the nuclei and cytosol. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.617779
- Bailey-Serres, J., and Voesenek, L. A. (2010). Life in the balance: a signaling network controlling survival of flooding. *Curr. Opin. Plant Biol.* 13, 489–494. doi: 10.1016/j.pbi.2010.08.002
- Ball, K. R., Baldock, J. A., Penfold, C., Power, S. A., Woodin, S. J., Smith, P., et al. (2020). Soil organic carbon and nitrogen pools are increased by mixed grass and legume cover crops in vineyard agroecosystems: Detecting short-term management

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

effects using infrared spectroscopy. *Geoderma* 379, 114619. doi: 10.1016/j.geoderma.2020.114619

Baronti, S., Magno, R., Maienza, A., Montagnoli, A., Ungaro, F., and Vaccari, F. P. (2022). Long term effect of biochar on soil plant water relation and fine roots: Results after 10 years of vineyard experiment. *Sci. Total Environ.* 851, 158225. doi: 10.1016/j.scitotenv.2022.158225

Baronti, S., Vaccari, F. P., Miglietta, F., Calzolari, C., Lugato, E., Orlandini, S., et al. (2014). ). impact of biochar application on plant water relations in *Vitis vinifera* (L.). *Eur. J. Agron.* 53, 38–44. doi: 10.1016/j.eja.2013.11.003

Bergqvist, J., Dokoozlian, N., and Ebisuda, N. (2001). Sunlight exposure and temperature effects on berry growth and composition of Cabernet sauvignon and grenache in the central San Joaquin valley of California. *Am. J. Enol. Vitic.* 52, 1–7. doi: 10.5344/ajev.2001.52.1.1

Berli, F., D'Angelo, J., Cavagnaro, B., Bottini, R., Wuilloud, R., and Silva, M. F. (2008). Phenolic composition in grape (*Vitis vinifera* L. cv. Malbec) ripened with different solar UV-B radiation levels by capillary zone electrophoresis. *J. Agric. Food Chem.* 56, 2892–2898. doi: 10.1021/jf073421+

Bianchi, D., Caramanico, L., Grossi, D., Brancadoro, L., and Lorenzis, G. D. (2020). How do novel m-rootstock (*Vitis* spp.) genotypes cope with drought? *Plants* 9, 1385. doi: 10.3390/plants9101385

Bonada, M., Jeffery, D. W., Petrie, P. R., Moran, M. A., and Sadras, V. O. (2015). Impact of elevated temperature and water deficit on the chemical and sensory profiles of barossa Shiraz grapes and wines. *Aust. J. Grape Wine Res.* 21 (2), 240–253. doi: 10.1111/ajgw.12142

Bonada, M., Petrie, P. R., Edwards, E. J., and McCarthy, M. G. (2017). The impact of winter drought on vine growth and wine quality. *Aust. N.Z. Grapegrow. Winemak.* 644, 41.

Bonada, M., Sadras, V. O., and Fuentes, S. (2013a). Effect of elevated temperature on the onset and rate of mesocarp cell death in berries of Shiraz and Chardonnay and its relationship with berry shrivel. *Aust. J. Grape Wine Res.* 19, 87–94. doi: 10.1111/ajgw.12010

Bonada, M., Sadras, V., Moran, M., and Fuentes, S. (2013b). Elevated temperature and water stress accelerate mesocarp cell death and shrivelling, and decouple sensory traits in Shiraz berries. *Irrig. Sci.* 31, 1317–1331. doi: 10.1007/s00271-013-0407-z

Bonnefoy, C., Quénel, H., Bonnardot, V., Barbeau, G., Madelin, M., Planchon, O., et al. (2013). Temporal and spatial analyses of temperature in a French wine-producing area: the Loire valley. *Int. J. Climatol.* 33, 1849–1862. doi: 10.1002/joc.3552

Bradford, M. A., McCulley, R. L., Crowther, T., Oldfield, E. E., Wood, S. A., and Fierer, N. (2019). Cross-biome patterns in soil microbial respiration predictable from evolutionary theory on thermal adaptation. *Nat. Ecol. Evol.* 3, 223–231. doi: 10.1038/s41559-018-0771-4

- Brillante, L., Belfiore, N., Gaiotti, F., Lovat, L., Sansone, L., Poni, S., et al. (2016). Comparing kaolin and pinolene to improve sustainable grapevine production during drought. *PLoS One* 11, e0156631. doi: 10.1371/journal.pone.0156631
- Buesa, I., Ballester, C., Mirás-Avalos, J. M., and Intrigliolo, D. S. (2020). Effects of leaning grapevine canopy to the West on water use efficiency and yield under Mediterranean conditions. *Agric. For. Meteorol.* 295, 108166. doi: 10.1016/j.agrformet.2020.108166
- Buesa, I., Caccavello, G., Basile, B., Merli, M. C., Poni, S., Chirivella, C., et al. (2018). Delaying berry ripening of bobal and tempranillo grapevines by late leaf removal in a semi-arid and temperate-warm climate under different water regimes. *Aust. J. Grape Wine Res.* 25, 70–82. doi: 10.1111/ajgw.12368
- Caffarra, A., and Eccel, E. (2010). Increasing the robustness of phenological models for vitis vinifera cv. Chardonnay. *Int. J. Biometeorol.* 54, 255–267. doi: 10.1007/s00484-009-0277-5
- Cai, W., and Cowan, T. (2013). Southeast Australia autumn rainfall reduction: A climate-change-induced poleward shift of ocean-atmosphere circulation. *J. Clim.* 26, 189–205. doi: 10.1175/JCLI-D-12-00035.1
- Cameron, W., Petrie, P. R., and Barlow, E. W. R. (2022). The effect of temperature on grapevine phenological intervals: Sensitivity of budburst to flowering. *Agric. For. Meteorol.* 315, 108841. doi: 10.1016/j.agrformet.2022.108841
- Cao, S., Xiao, Z., Jiranek, V., and Tyerman, S. D. (2019). The VvBAP1 gene is identified as a potential inhibitor of cell death in grape berries. *Funct. Plant Biol.* 46, 428–442. doi: 10.1071/FP18272
- Caravia, L., Collins, C., Petrie, P. R., and Tyerman, S. D. (2016). Application of shade treatments during Shiraz berry ripening to reduce the impact of high temperature. *Aust. J. Grape Wine Res.* 22, 422–437. doi: 10.1111/ajgw.12248
- Caravia, L., Collins, C., and Tyerman, S. D. (2015). Electrical impedance of Shiraz berries correlates with decreasing cell vitality during ripening. *Aust. J. Grape Wine Res.* 21, 430–438. doi: 10.1111/ajgw.12157
- Cataldo, E., Fucile, M., and Mattii, G. B. (2021). A review: Soil management, sustainable strategies and approaches to improve the quality of modern viticulture. *Agronomy* 11, 2359. doi: 10.3390/agronomy11112359
- Cataldo, E., Fucile, M., and Mattii, G. B. (2022). Biostimulants in viticulture: A sustainable approach against biotic and abiotic stresses. *Plants* 11, 162. doi: 10.3390/plants11020162
- Celette, F., Gaudin, R., and Gary, C. (2008). Spatial and temporal changes to the water regime of a Mediterranean vineyard due to the adoption of cover cropping. *Eur. J. Agron.* 29, 153–162. doi: 10.1016/j.eja.2008.04.007
- Chalmers, K. (2019). New arrivals in alternative varieties. *Aust. N.Z. Grapegrow. Winemak.* 665, 58–65.
- Charrier, G., Delzon, S., Domec, J. C., Zhang, L., Delmas, C. E., Merlin, I., et al. (2018). Drought will not leave your glass empty: Low risk of hydraulic failure revealed by long-term drought observations in world's top wine regions. *Sci. Adv.* 4, 6969. doi: 10.1126/sciadv.aao6969
- Chen, T., and Fluhr, R. (2018). Singlet oxygen plays an essential role in the root's response to osmotic stress. *Plant Physiol.* 177, 1717–1727. doi: 10.1104/pp.18.00634
- Clarke, H., Lucas, C., and Smith, P. (2013). Changes in Australian fire weather between 1973 and 2010. *Int. J. Climatol.* 33, 931–944. doi: 10.1002/joc.3480
- Clonch, C., Huynh, M., Goto, B., Levin, A., Selker, J., and Udell, C. (2021). High precision zero-friction magnetic dendrometer. *HardwareX* 10, e00248. doi: 10.1016/j.ohx.2021.e00248
- Cola, G., Failla, O., Maghradze, D., Megrelidze, L., and Mariani, L. (2017). Grapevine phenology and climate change in Georgia. *Int. J. Biometeorol.* 61, 761–773. doi: 10.1007/s00484-016-1241-9
- Colmer, T. D., and Greenway, H. (2011). Ion transport in seminal and adventitious roots of cereals during O<sub>2</sub> deficiency. *J. Exp. Bot.* 62, 39–57. doi: 10.1093/jxb/erq271
- Conde, A., Neves, A., Breia, R., Pimentel, D., Dinis, L. T., Bernardo, S., et al. (2018). Kaolin particle film application stimulates photoassimilate synthesis and modifies the primary metabolome of grape leaves. *J. Plant Physiol.* 223, 47–56. doi: 10.1016/j.jplph.2018.02.004
- Conde, A., Pimentel, D., Neves, A., Dinis, L. T., Bernardo, S., Correia, C. M., et al. (2016). Kaolin foliar application has a stimulatory effect on phenylpropanoid and flavonoid pathways in grape berries. *Front. Plant Sci.* 7, doi: 10.3389/fpls.2016.01150
- Culbert, J. A., Krstic, M. P., and Herderich, M. J. (2021). Development and utilization of a model system to evaluate the potential of surface coatings for protecting grapes from volatile phenols implicated in smoke taint. *Molecules* 26, 5197. doi: 10.3390/molecules26175197
- Curtin, D., Beare, M. H., and Hernandez-Ramirez, G. (2012). Temperature and moisture effects on microbial biomass and soil organic matter mineralization. *Soil Sci. Soc. Am. J.* 76, 2055–2067. doi: 10.2136/sssaj2012.0011
- Davies, W. J., Bacon, M. A., Stuart Thompson, D., Sobehi, W., and González Rodríguez, L. (2000). Regulation of leaf and fruit growth in plants growing in drying soil: exploitation of the plants' chemical signalling system and hydraulic architecture to increase the efficiency of water use in agriculture. *J. Exp. Bot.* 51, 1617–1626. doi: 10.1093/jxb/51.13.1617
- Dayer, S., Herrera, J. C., Dai, Z., Burrell, R., Lamarque, L. J., Delzon, S., et al. (2021). Nighttime transpiration represents a negligible part of water loss and does not increase the risk of water stress in grapevine. *Plant Cell Environ.* 44, 387–398. doi: 10.1111/pce.13923
- De Bei, R., Cozzolino, D., Sullivan, W., Cynkar, W., Fuentes, S., Damberg, R., et al. (2011). Non-destructive measurement of grapevine water potential using near infrared spectroscopy. *Aust. J. Grape Wine Res.* 17, 62–71. doi: 10.1111/j.1755-0238.2010.00117.x
- De Bei, R., Fuentes, S., Gilliam, M., Tyerman, S., Edwards, E., Bianchini, N., et al. (2016). VitiCanopy: A free computer app to estimate canopy vigor and porosity for grapevine. *Sensors* 16, 585. doi: 10.3390/s16040585
- De Bei, R., Wang, X., Papagiannis, L., Cocco, M., O'Brien, P., Zito, M., et al. (2019). Postveraison leaf removal does not consistently delay ripening in semillon and Shiraz in a hot Australian climate. *Am. J. Enol. Vitic.* 70, 398–410. doi: 10.5344/ajev.2019.18103
- De Bei, R., Wang, X., Papagiannis, L., Fuentes, S., Gilliam, M., Tyerman, S., et al. (2020). Shoot thinning of semillon in a hot climate did not improve yield and berry and wine quality. *Oeno One* 54, 469–484. doi: 10.20870/oeno-one.2020.54.3.2984
- Debnath, S., Paul, M., Rahaman, D. M., Debnath, T., Zheng, L., Baby, T., et al. (2021). Identifying individual nutrient deficiencies of grapevine leaves using hyperspectral imaging. *Remote Sens.* 13, 3317. doi: 10.3390/rs13163317
- Diago, M. P., Fernández-Novales, J., Gutiérrez, S., Marañón, M., and Tardaguila, J. (2018). Development and validation of a new methodology to assess the vineyard water status by on-the-go near infrared spectroscopy. *Front. Plant Sci.* 9, doi: 10.3389/fpls.2018.00059
- Dinis, L. T., Ferreira, H., Pinto, G., Bernardo, S., Correia, C. M., and Moutinho-Pereira, J. (2016). Kaolin-based, foliar reflective film protects photosystem II structure and function in grapevine leaves exposed to heat and high solar radiation. *Photosynthetica* 54, 47–55. doi: 10.1007/s11099-015-0156-8
- Doran, J. W. (2002). Soil health and global sustainability: translating science into practice. *Agric. Ecosyst. Environ.* 88, 119–127. doi: 10.1016/S0167-8809(01)00246-8
- Dry, I. B., Davies, C., Dunlevy, J. D., Smith, H. M., Thomas, M. R., Walker, A. R., et al. (2022). Development of new wine-, dried-and tablegrape scions and rootstocks for Australian viticulture: past, present and future. *Aust. J. Grape Wine Res.* 28, 177–195. doi: 10.1111/ajgw.12552
- Dry, I. B., and Thomas, M. R. (2015). Fast-tracking grape breeding for disease resistance. *Wine Vitic. J.* 30, 52–55.
- Duchêne, E., Huard, F., Dumas, V., Schneider, C., and Merdinoglu, D. (2010). The challenge of adapting grapevine varieties to climate change. *Clim. Res.* 41, 193–204. doi: 10.3354/cr00850
- Duchêne, E., and Schneider, C. (2005). Grapevine and climatic changes: A glance at the situation in Alsace. *Agron. Sustain. Dev.* 25, 93–99. doi: 10.1051/agro:2004057
- Du Jardin, P. (2015). Plant biostimulants: Definition, concept, main categories and regulation. *Sci. Hortic.* 196, 3–14. doi: 10.1016/j.scienta.2015.09.021
- Dungey, K. A., Hayasaka, Y., and Wilkinson, K. L. (2011). Quantitative analysis of glycoconjugate precursors of guaiaicol in smoke-affected grapes using liquid chromatography-tandem mass spectrometry based stable isotope dilution analysis. *Food Chem.* 126, 801–806. doi: 10.1016/j.foodchem.2010.11.094
- Edwards, E. J., and Clingeleffer, P. R. (2013). Interseasonal effects of regulated deficit irrigation on growth, yield, water use, berry composition and wine attributes of Cabernet sauvignon grapevines. *Aust. J. Grape Wine Res.* 19, 261–276. doi: 10.1111/ajgw.12027
- Elad, Y., Cytryn, E., Harel, Y. M., Lew, B., and Graber, E. R. (2011). The biochar effect: plant resistance to biotic stresses. *Phytopathol. Mediterr.* 50, 335–349. Available at: <https://www.jstor.org/stable/26556455>
- Etien, N., Daux, V., Masson-Delmotte, V., Mestre, O., Stievenard, M., Guillemain, M. T., et al. (2009). Summer maximum temperature in northern France over the past century: instrumental data versus multiple proxies (tree-ring isotopes, grape harvest dates and forest fires). *Clim. Change* 94, 429–456. doi: 10.1007/s10584-008-9516-8
- Falcão, L. D., Burin, V. M., Chaves, E. S., Vieira, H. J., Brighenti, E., Rosier, J. P., et al. (2010). Vineyard altitude and mesoclimate influences on the phenology and maturation of Cabernet-sauvignon grapes from Santa catarina state. *Oeno One* 44, 135–150. doi: 10.20870/oeno-one.2010.44.3.1470
- Filkov, A. I., Ngo, T., Matthews, S., Telfer, S., and Penman, T. D. (2020). Impact of australia's catastrophic 2019/20 bushfire season on communities and environment: retrospective analysis and current trends. *J. Safety Sci. Resil.* 1, 44–56. doi: 10.1016/j.jnlssr.2020.06.009
- Friend, A. P., and Trought, M. C. (2007). Delayed winter spur-pruning in new Zealand can alter yield components of merlot grapevines. *Aust. J. Grape Wine Res.* 13, 157–164. doi: 10.1111/j.1755-0238.2007.tb00246.x

- Fryda, L., and Visser, R. (2015). Biochar for soil improvement: Evaluation of biochar from gasification and slow pyrolysis. *Agriculture* 5, 1076–1115. doi: 10.3390/agriculture5041076
- Fuentes, S., and Gago, J. (2022). “Modern approaches to precision and digital viticulture,” in *Improving sustainable viticulture and winemaking practices*. Eds. J. Miguel Costa, S. Catarino, J. M. Escalona and P. Comuzzo (London: Academic Press), 125–145. doi: 10.1016/B978-0-323-85150-3.00015-3
- Fuentes, S., Sullivan, W., Tilbrook, J., and Tyerman, S. (2010). A novel analysis of grapevine berry tissue demonstrates a variety-dependent correlation between tissue vitality and berry shrivel. *Aust. J. Grape Wine Res.* 16, 327–336. doi: 10.1111/j.1755-0238.2010.00095.x
- Fuentes, S., and Tongson, E. (2018). Advances in smoke contamination detection systems for grapevine canopies and berries. *Wine Vitic. J.* 32, 36–39.
- Galarneau, E. R., Lawrence, D. P., Travadon, R., and Baumgartner, K. (2019). Drought exacerbates botryosphera dieback symptoms in grapevines and confounds host-based molecular markers of infection by *neofusicoccum parvum*. *Plant Dis.* 103, 1738–1745. doi: 10.1094/PDIS-09-18-1549-RE
- Gambetta, G. A., Manuck, C. M., Drucker, S. T., Shaghasi, T., Fort, K., Matthews, M. A., et al. (2012). The relationship between root hydraulics and scion vigour across vitis rootstocks: what role do root aquaporins play? *J. Exp. Bot.* 63, 6445–6455. doi: 10.1093/jxb/ers312
- García, L., Celette, F., Gary, C., Ripoché, A., Valdés-Gómez, H., and Metay, A. (2018). Management of service crops for the provision of ecosystem services in vineyards: A review. *Agric. Ecosyst. Environ.* 251, 158–170. doi: 10.1016/j.agee.2017.09.030
- García-Jaramillo, M., Meyer, K. M., Phillips, C. L., Acosta-Martínez, V., Osborne, J., Levin, A. D., et al. (2021). Biochar addition to vineyard soils: effects on soil functions, grape yield and wine quality. *Biochar* 3, 565–577. doi: 10.1007/s42773-021-00118-x
- Gatti, M., Galbignani, M., Garavani, A., Bernizzoni, F., Tombesi, S., Palliotti, A., et al. (2016). Manipulation of ripening via antitranspirants in cv. Barbera (Vitis vinifera L.). *Aust. J. Grape Wine Res.* 22, 245–255. doi: 10.1111/ajgw.12212
- Gatti, M., Garavani, A., Cantatore, A., Parisi, M. G., Bobeica, N. A., Merli, M. C., et al. (2015). Interactions of summer pruning techniques and vine performance in the white vitis vinifera cv. ortrugo. *Aust. J. Grape Wine Res.* 21, 80–89. doi: 10.1111/ajgw.12107
- Genesio, L., Miglietta, F., Baronti, S., and Vaccari, F. P. (2015). Biochar increases vineyard productivity without affecting grape quality: Results from a four years field experiment in Tuscany. *Agric. Ecosyst. Environ.* 201, 20–25. doi: 10.1016/j.agee.2014.11.021
- Giagnoni, L., Maienza, A., Baronti, S., Vaccari, F. P., Genesio, L., Taiti, C., et al. (2019). Long-term soil biological fertility, volatile organic compounds and chemical properties in a vineyard soil after biochar amendment. *Geoderma* 344, 127–136. doi: 10.1016/j.geoderma.2019.03.011
- Godden, P., and Muhlack, R. (2010). Trends in the composition of Australian wine. *Aust. N.Z. Grapegrow. Winemak.* 558, 47–61.
- Gomès, É., Maillot, P., and Duchêne, É. (2021). Molecular tools for adapting viticulture to climate change. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.633846
- González-Caballero, V., Sánchez, M. T., Fernández-Novales, J., López, M. I., and Pérez-Marín, D. (2012). On-vine monitoring of grape ripening using near-infrared spectroscopy. *Food Anal. Methods* 5, 1377–1385. doi: 10.1007/s12161-012-9389-3
- González, C. V., Fanzone, M. L., Cortés, L. E., Bottini, R., Lijavetzky, D. C., Ballaré, C. L., et al. (2015). Fruit-localized photoreceptors increase phenolic compounds in berry skins of field-grown vitis vinifera L. cv. Malbec. *Phytochemistry* 110, 46–57. doi: 10.1016/j.phytochem.2014.11.018
- Goode, J. (2012). Viticulture: fruity with a hint of drought. *Nature* 492 (7429), 351. doi: 10.1038/492351a
- Gouot, J. C., Smith, J. P., Holzapfel, B. P., Walker, A. R., and Barril, C. (2019). Grape berry flavonoids: A review of their biochemical responses to high and extreme high temperatures. *J. Exp. Bot.* 70, 397–423. doi: 10.1093/jxb/ery392
- Greer, D. H., Berry, J. A., and Björkman, O. (1986). Photoinhibition of photosynthesis in intact bean leaves: Role of light and temperature, and requirements for chloroplast-protein synthesis during recovery. *Planta* 168, 253–260. doi: 10.1007/BF00402971
- Greer, D. H., Rogiers, S. Y., and Steel, C. C. (2006). Susceptibility of Chardonnay grapes to sunburn. *Vitis* 45, 147–148. doi: 10.5073/vitis.2006.45.147-148
- Greer, D. H., and Weedon, M. M. (2012). Modelling photosynthetic responses to temperature of grapevine (Vitis vinifera cv. semillon) leaves on vines grown in a hot climate. *Plant Cell Environ.* 35, 1050–1064. doi: 10.1111/j.1365-3040.2011.02471.x
- Greer, D. H., and Weedon, M. M. (2013). The impact of high temperatures on vitis vinifera cv. semillon grapevine performance and berry ripening. *Front. Plant Sci.* 4. doi: 10.3389/fpls.2013.00491
- Greer, D. H., and Weedon, M. M. (2014). Does the hydrocooling of vitis vinifera cv. semillon vines protect the vegetative and reproductive growth processes and vine performance against high summer temperatures? *Funct. Plant Biol.* 41, 620–633. doi: 10.1071/FP13286
- Greer, D. H., and Weston, C. (2010). Heat stress affects flowering, berry growth, sugar accumulation and photosynthesis of vitis vinifera cv. semillon grapevines grown in a controlled environment. *Funct. Plant Biol.* 37, 206–214. doi: 10.1071/FP09209
- Hague, B. S. (2021). Seasonal climate summary for Australia and the southern hemisphere (summer 2018–19): extreme heat and flooding prominent. *J. South. Hemis. Earth Syst. Sci.* 71, 147–158. doi: 10.1017/ES20009
- Hall, A., and Jones, G. V. (2009). Effect of potential atmospheric warming on temperature-based indices describing Australian winegrape growing conditions. *Aust. J. Grape Wine Res.* 15, 97–119. doi: 10.1111/j.1755-0238.2008.00035.x
- Hannah, L., Roehrdanz, P. R., Ikegami, M., Shepard, A. V., Shaw, M. R., Tabor, G., et al. (2013). Climate change, wine, and conservation. *P.N.A.S.* 110, 6907–6912. doi: 10.1073/pnas.1210127110
- Hardie, W. J., and Considine, J. A. (1976). Response of grapes to water-deficit stress in particular stages of development. *Am. J. Enol. Vitic.* 27, 55–61.
- Hayman, P., Longbottom, M., McCarthy, M., and Thomas, D. (2012). *Managing grapevines during heatwaves*. GWRDC pp.1–6. Available at: [https://www.langhornecreek.com/wp-content/uploads/2017/08/HeatManging\\_grapevines\\_Factsheet\\_GWRDC.pdf](https://www.langhornecreek.com/wp-content/uploads/2017/08/HeatManging_grapevines_Factsheet_GWRDC.pdf) (Accessed 27 October 2022).
- Hranilovic, A., Gambetta, J. M., Jeffery, D. W., Grbin, P. R., and Jiranek, V. (2020). Lower-alcohol wines produced by metschnikowia pulcherrima and saccharomyces cerevisiae co-fermentations: The effect of sequential inoculation timing. *Int. J. Food Microbiol.* 329, 108651. doi: 10.1016/j.jfoodmicro.2020.108651
- Huang, J., Minasny, B., McBratney, A. B., Padarian, J., and Triantafyllis, J. (2018). The location- and scale-specific correlation between temperature and soil carbon sequestration across the globe. *Sci. Total Environ.* 615, 540–548. doi: 10.1016/j.scitotenv.2017.09.136
- Imbufe, A. U., Patti, A. F., Burrow, D., Surapaneni, A., Jackson, W. R., and Milner, A. D. (2005). Effects of potassium humate on aggregate stability of two soils from Victoria, Australia. *Geoderma* 125, 321–330. doi: 10.1016/j.geoderma.2004.09.006
- Jarvis, C., Barlow, E., Darbyshire, R., Eckard, R., and Goodwin, I. (2017). Relationship between viticultural climatic indices and grape maturity in Australia. *Int. J. Biometeorol.* 61, 1849–1862. doi: 10.1007/s00484-017-1370-9
- Jones, G. V. (2006). “Climate and terroir: Impacts of climate variability and change on wine,” in *Fine wine and terroir - the geoscience perspective*. Eds. R. W. Macqueen and L. D. Meinert (St. John's: Geological Association of Canada), 1–14.
- Jones, G. V., and Alves, F. (2012). Impact of climate change on wine production: a global overview and regional assessment in the douro valley of Portugal. *Int. J. Global Warming* 4, 383–406. doi: 10.1504/IJGW.2012.049448
- Jones, G. V., and Davis, R. E. (2000). Climate influences on grapevine phenology, grape composition, and wine production and quality for Bordeaux. *France. Am. J. Enol. Vitic.* 51, 249–261. doi: 10.5344/ajev.2000.51.3.249
- Jones, G. V., Edwards, E. J., Bonada, M., Sadras, V. O., Krstic, M. P., and Herdich, M. J. (2022). “Climate change and its consequences for viticulture,” in *Managing wine quality*. Ed. A. G. Reynolds (Sawston, UK: Woodhead Publishing), 727–778. doi: 10.1016/B978-0-08-102067-8.00015-4
- Jones, G. V., and Webb, L. B. (2010). Climate change, viticulture, and wine: challenges and opportunities. *J. Wine Res.* 21, 103–106. doi: 10.1080/09571264.2010.530091
- Jones, G. V. (2012). Climate, grapes and wine: Structure and suitability in a changing climate. *Acta Hort.* 931, 19–28. doi: 10.17660/ActaHortic.2012.931.1
- Jones, G. V., White, M. A., Cooper, O. R., and Storchmann, K. (2005). Climate change and global wine quality. *Clim. Change* 73, 319–343. doi: 10.1007/s10584-005-4704-2
- Keller, M. (2010). Managing grapevines to optimise fruit development in a challenging environment: a climate change primer for viticulturists. *Aust. J. Grape Wine Res.* 16, 56–69. doi: 10.1111/j.1755-0238.2009.00077.x
- Kennison, K. R., Wilkinson, K. L., Williams, H. G., Smith, J. H., and Gibberd, M. R. (2007). Smoke-derived taint in wine: Effect of postharvest smoke exposure of grapes on the chemical composition and sensory characteristics of wine. *J. Agric. Food Chem.* 55, 10897–10901. doi: 10.1021/jf072509k
- Kliever, W. M., and Torres, R. E. (1972). Effect of controlled day and night temperatures on grape coloration. *Am. J. Enol. Vitic.* 23, 71–77.
- Kolton, M., Graber, E. R., Tsehansky, L., Elad, Y., and Cytryn, E. (2017). Biochar-stimulated plant performance is strongly linked to microbial diversity and metabolic potential in the rhizosphere. *New Phytol.* 213, 1393–1404. doi: 10.1111/nph.14253
- Krasnow, M., Matthews, M., and Shackel, K. (2008). Evidence for substantial maintenance of membrane integrity and cell viability in normally developing grape (Vitis vinifera L.) berries throughout development. *J. Exp. Bot.* 59, 849–859. doi: 10.1093/jxb/erm372



- Krstic, M. P., Johnson, D. L., and Herderich, M. J. (2015). Review of smoke taint in wine: smoke-derived volatile phenols and their glycosidic metabolites in grapes and vines as biomarkers for smoke exposure and their role in the sensory perception of smoke taint. *Aust. J. Grape Wine Res.* 21, 537–553. doi: 10.1111/ajgw.12183
- Laird, D. A. (2008). The charcoal vision: A win-win-win scenario for simultaneously producing bioenergy, permanently sequestering carbon, while improving soil and water quality. *Agron. J.* 100, 178–181. doi: 10.2134/agronj2007.0161
- Lal, R. (2016). Soil health and carbon management. *Food Energy Secur.* 5, 212–222. doi: 10.1002/fes3.96
- Lecourieux, F., Kappel, C., Pieri, P., Charon, J., Pillet, J., Hilbert, G., et al. (2017). Dissecting the biochemical and transcriptomic effects of a locally applied heat treatment on developing Cabernet Sauvignon grape berries. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.00053
- Leolini, L., Moriondo, M., Romboli, Y., Gardiman, M., Costafreda-Aumedes, S., de Cortazar-Atauri, I. G., et al. (2019). Modelling sugar and acid content in sangiovese grapes under future climates: An Italian case study. *Clim. Res.* 78, 211–224. doi: 10.3354/cr01571
- Liu, S., Zeng, X., and Whitty, M. (2020). 3DBunch: A novel iOS-smartphone application to evaluate the number of grape berries per bunch using image analysis techniques. *IEEE Access* 8, 114663–114674. doi: 10.1109/ACCESS.2020.3003415
- Loescher, W. H., McCamant, T., and Keller, J. D. (1990). Carbohydrate reserves, translocation, and storage in woody plant roots. *HortSci* 25, 274–281.
- Longbottom, M. L., and Petrie, P. R. (2015). Role of vineyard practices in generating and mitigating greenhouse gas emissions. *Aust. J. Grape Wine Res.* 21, 522–536. doi: 10.1111/ajgw.12197
- Longo, R., Blackman, J. W., Antalick, G., Torley, P. J., Rogiers, S. Y., and Schmidtko, L. M. (2018). Harvesting and blending options for lower alcohol wines: A sensory and chemical investigation. *J. Sci. Food Agric.* 98, 33–42. doi: 10.1002/jsfa.8434
- López-Urrea, R., Sánchez, J. M., Montoro, A., Mañas, F., and Intrigliolo, D. S. (2020). Effect of using pruning waste as an organic mulching on a drip-irrigated vineyard evapotranspiration under a semi-arid climate. *Agric. For. Meteorol.* 291, 108064. doi: 10.1016/j.agrformet.2020.108064
- Lovisolo, C., Perrone, I., Carra, A., Ferrandino, A., Flexas, J., Medrano, H., et al. (2010). Drought-induced changes in development and function of grapevine (*Vitis* spp.) organs and in their hydraulic and non-hydraulic interactions at the whole-plant level: a physiological and molecular update. *Funct. Plant Biol.* 37, 98–116. doi: 10.1071/FP09191
- Mackie, K. A., Marhan, S., Ditterich, F., Schmidt, H. P., and Kandler, E. (2015). The effects of biochar and compost amendments on copper immobilization and soil microorganisms in a temperate vineyard. *Agric. Ecosyst. Environ.* 201, 58–69. doi: 10.1016/j.agee.2014.12.001
- Maes, W. H., and Steppe, K. (2019). Perspectives for remote sensing with unmanned aerial vehicles in precision agriculture. *Trends Plant Sci.* 24, 152–164. doi: 10.1016/j.tplants.2018.11.007
- Mahmud, K. P. (2016). *Factors influencing diurnal and seasonal fine root growth in grapevines* (Charles Sturt University: Wagga Wagga (NSW)).
- Malheiro, A. C., Campos, R., Fraga, H., Eiras-Dias, J., Silvestre, J., and Santos, J. A. (2013). Winegrape phenology and temperature relationships in the Lisbon wine region, Portugal. *OENO One* 47, 287–299. doi: 10.20870/oeno-one.2013.47.4.1558
- Mancha, L. A., Uriarte, D., and Prieto, M. D. H. (2021). Characterization of the transpiration of a vineyard under different irrigation strategies using sap flow sensors. *Water* 13, 2867. doi: 10.3390/w13202867
- Manfreda, S., McCabe, M. F., Miller, P. E., Lucas, R., Pajuelo Madrigal, V., Mallinis, G., et al. (2018). On the use of unmanned aerial systems for environmental monitoring. *Remote Sens.* 10, 641. doi: 10.3390/rs10040641
- Marin, D., Armengol, J., Carbonell-Bejerano, P., Escalona, J. M., Gramaje, D., Hernández-Montes, E., et al. (2021). Challenges of viticulture adaptation to global change: Tackling the issue from the roots. *Aust. J. Grape Wine Res.* 27, 8–25. doi: 10.1111/ajgw.12463
- Matese, A., Toscano, P., Di Gennaro, S. F., Genesio, L., Vaccari, F. P., Prmicierio, J., et al. (2015). Intercomparison of UAV, aircraft and satellite remote sensing platforms for precision viticulture. *Remote Sens.* 7, 2971–2990. doi: 10.3390/rs70302971
- May, P. (1994). *Using grapevine rootstocks: the Australian Perspective, 1st ed* (Adelaide, SA: Winetitles).
- Mayr, C. M., Parker, M., Baldock, G. A., Black, C. A., Pardon, K. H., Williamson, P. O., et al. (2014). Determination of the importance of in-mouth release of volatile phenol glycoconjugates to the flavor of smoke-tainted wines. *J. Agric. Food Chem.* 62, 2327–2336. doi: 10.1021/jf405327s
- Meitha, K., Konnerup, D., Colmer, T. D., Considine, J. A., Foyer, C. H., and Considine, M. J. (2015). Spatio-temporal relief from hypoxia and production of reactive oxygen species during bud burst in grapevine (*Vitis vinifera*). *Ann. Bot.* 116, 703–711. doi: 10.1093/aob/mcv123
- Meng, X., Li, B., Liu, J., and Tian, S. (2008). Physiological responses and quality attributes of table grape fruit to chitosan preharvest spray and postharvest coating during storage. *Food Chem.* 106, 501–508. doi: 10.1016/j.foodchem.2007.06.012
- Mittler, R., Zandalinas, S. I., Fichman, Y., and Van Breusegem, F. (2022). Reactive oxygen species signalling in plant stress responses. *Nat. Rev. Mol. Cell Biol.* 23, 663–679. doi: 10.1038/s41580-022-00499-2
- Monteiro, E., Gonçalves, B., Cortez, I., and Castro, I. (2022). The role of biostimulants as alleviators of biotic and abiotic stresses in grapevine: A review. *Plants* 11, 396. doi: 10.3390/plants11030396
- Morabito, C., Orozco, J., Tonel, G., Cavalletto, S., Meloni, G. R., Schubert, A., et al. (2022). Do the ends justify the means? Impact of drought progression rate on stress response and recovery in *Vitis vinifera*. *Physiol. Plantarum* 174, e13590. doi: 10.1111/jplp.13590
- Morales-Castilla, I., De Cortazar-Atauri, I., Cook, B. I., Lacombe, T., Parker, A., Van Leeuwen, C., et al. (2020). Diversity buffers winegrowing regions from climate change losses. *Proc. Natl. Acad. Sci.* 117, 2864–2869. doi: 10.1073/pnas.1906731117
- Moran, M. A., Bastian, S. E., Petrie, P. R., and Sadras, V. O. (2018). Late pruning impacts on chemical and sensory attributes of Shiraz wine. *Aust. J. Grape Wine Res.* 24, 469–477. doi: 10.1111/ajgw.12350
- Moran, M. A., Sadras, V. O., and Petrie, P. R. (2017). Late pruning and carry-over effects on phenology, yield components and berry traits in Shiraz. *Aust. J. Grape Wine Res.* 23, 390–398. doi: 10.1111/ajgw.12298
- Mosedale, J. R., Abernethy, K. E., Smart, R. E., Wilson, R. J., and Maclean, I. M. (2016). Climate Change impacts adaptive strategies: lessons grapevine. *Glob. Change Biol.* 22, 3814–3828. doi: 10.1111/gcb.13406
- Mundy, D. C., and Agnew, R. H. (2002). Grape marc in mulches: the potassium effect. *Aus. N.Z. Grapegrower Winemaker* 458, 59–62.
- Murray–Darling Basin Authority (2010). *Guide to the proposed basin plan: Overview* (Canberra: Murray–Darling Basin Authority).
- Nairn, J. R., and Fawcett, R. J. (2015). The excess heat factor: a metric for heatwave intensity and its use in classifying heatwave severity. *Int. J. Environ. Res. Public Health* 12, 227–253. doi: 10.3390/ijerph120100227
- Nguyen, L. T., Osanai, Y., Anderson, I. C., Bange, M. P., Braunack, M., Tissue, D. T., et al. (2018). Impacts of waterlogging on soil nitrification and ammonia-oxidizing communities in farming system. *Plant Soil* 426, 299–311. doi: 10.1007/s11104-018-3584-y
- Noestheden, M., Dennis, E. G., Romero-Montalvo, E., DiLabio, G. A., and Zandberg, W. F. (2018). Detailed characterization of glycosylated sensory-active volatile phenols in smoke-exposed grapes and wine. *Food Chem.* 259, 147–156. doi: 10.1016/j.foodchem.2018.03.097
- Oczkowski, E. (2016). The effect of weather on wine quality and prices: An Australian spatial analysis. *J. Wine Econ.* 11, 48–65. doi: 10.1017/jwe.2015.14
- OIV (2010) *Resolution OIV/VITI 333/2010, definition of vitivinicultural “terroir”*. Available at: <https://www.oiv.int/public/medias/379/viti-2010-1-en.pdf>.
- Oletic, D., Rosner, S., Zovko, M., and Bilas, V. (2020). Time-frequency features of grapevine’s xylem acoustic emissions for detection of drought stress. *Comput. Electron. Agric.* 178, 105797. doi: 10.1016/j.compag.2020.105797
- Ollat, N., Peccoux, A., Papura, D., Esmenjaud, D., Marguerit, E., Tandonnet, J. P., et al. (2015). “Rootstocks as a component of adaptation to environment,” in *Grapevine in a changing environment: a molecular and ecophysiological perspective*. Eds. H. Gerós, M. M. Chaves, H. M. Gil and S. Delrot (West Sussex: Wiley-Blackwell), 68–108. doi: 10.1002/9781118735985.ch4
- Ooi, M. K., Auld, T. D., and Denham, A. J. (2009). Climate change and bet-hedging: interactions between increased soil temperatures and seed bank persistence. *glob. Change Biol.* 15, 2375–2386. doi: 10.1111/j.1365-2486.2009.01887.x
- Ouyang, J., De Bei, R., and Collins, C. (2021). Assessment of canopy size using UAV-based point cloud analysis to detect the severity and spatial distribution of canopy decline. *Oeno One* 55, 253–266. doi: 10.20870/oeno-one.2021.55.1.3078
- Palliotti, A., Frioni, T., Tombesi, S., Sabbatini, P., Cruz-Castillo, J. G., Lanari, V., et al. (2017). Double-pruning grapevines as a management tool to delay berry ripening and control yield. *Am. J. Enol. Vitic.* 68, 412–421. doi: 10.5344/ajev.2017.17011
- Palliotti, A., Panara, F., Famiani, F., Sabbatini, P., Howell, G. S., Silvestroni, O., et al. (2013). Postveraison application of antitranspirant di-1-p-menthene to control sugar accumulation in sangiovese grapevines. *Am. J. Enol. Vitic.* 64, 378–385. doi: 10.5344/ajev.2013.13015
- Petrie, P. R., Brooke, S. J., Moran, M. A., and Sadras, V. O. (2017). Pruning after budburst to delay and spread grape maturity. *Aust. J. Grape Wine Res.* 23, 378–389. doi: 10.1111/ajgw.12303
- Petrie, P. R., and Sadras, V. O. (2008). Advancement of grapevine maturity in Australia between 1993 and 2006: putative causes, magnitude of trends and



- viticultural consequences. *Aust. J. Grape Wine Res.* 14, 33–45. doi: 10.1111/j.1755-0238.2008.00005.x
- Petrie, P., and Sadras, V. O. (2016). Quantifying the advancement and compression of vintage. *AWRI Tech. Rev.* 220, 9–11.
- Petrie, P. R., Wang, Y., Liu, S., Lam, S., Whitty, M. A., and Skewes, M. A. (2019). The accuracy and utility of a low cost thermal camera and smartphone-based system to assess grapevine water status. *Biosyst. Eng.* 179, 126–139. doi: 10.1016/j.biosystemseng.2019.01.002
- Pilati, S., Brazzale, D., Guella, G., Milli, A., Ruberti, C., Biasioli, F., et al. (2014). The onset of grapevine berry ripening is characterized by ROS accumulation and lipoxygenase-mediated membrane peroxidation in the skin. *BMC Plant Biol.* 14, 1–15. doi: 10.1186/1471-2229-14-87
- Pillet, J., Egert, A., Pieri, P., Lecourieux, F., Kappel, C., Charon, J., et al. (2012). VvGOLS1 and VvHsFA2 are involved in the heat stress responses in grapevine berries. *Plant Cell Physiol.* 53, 1776–1792. doi: 10.1093/pcp/pcs121
- Pipan, P., Hall, A., Rogiers, S. Y., and Holzapfel, B. P. (2021). Accuracy of interpolated versus in-vineyard sensor climate data for heat accumulation modelling of phenology. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.635299
- Pitt, T., Skewes, M., Stevens, R., Tan, J., Nicholas, P., and McCarthy, M. (2018). Longevity and sustained performance of rootstocks for Australian vineyards. *Wine Viticult. J.* 33, 47.
- Ponnamperuma, F. N. (1972). The chemistry of submerged soil. *Adv. Agron.* 24, 29–96. doi: 10.1016/S0065-2113(08)60633-1
- Power, A., Truong, V. K., Chapman, J., and Cozzolino, D. (2019). From the laboratory to the vineyard—evolution of the measurement of grape composition using NIR spectroscopy towards high-throughput analysis. *High Throughput* 8, 21. doi: 10.3390/ht8040021
- Reineke, A., and Thiéry, D. (2016). Grapevine insect pests and their natural enemies in the age of global warming. *J. Pest Sci.* 89, 313–328. doi: 10.1007/s10340-016-0761-8
- Ristic, R., Hranilovic, A., Li, S., Longo, R., Pham, D. T., Qesja, B., et al. (2016). Alcohol: Integrated strategies to moderate the alcohol content of wines. *Wine Viticult. J.* 31, 33.
- Rogiers, S. Y., and Clarke, S. J. (2013). Nocturnal and daytime stomatal conductance respond to root-zone temperature in 'Shiraz' grapevines. *Ann. Bot.* 111, 433–444. doi: 10.1093/aob/mcs298
- Rogiers, S. Y., Clarke, S. J., and Schmidtke, L. M. (2014). Elevated root-zone temperature hastens vegetative and reproductive development in Shiraz grapevines. *Aust. J. Grape Wine Res.* 20, 123–133. doi: 10.1111/ajgw.12053
- Rogiers, S. Y., Coetzee, Z. A., Walker, R. R., Deloire, A., and Tyerman, S. D. (2017). Potassium in the grape (*Vitis vinifera* L.) berry: transport and function. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.01629
- Rogiers, S. Y., and Fahey, D. J. (2019). Di-1-p-menthene reduces grape leaf and bunch transpiration. *Aust. J. Grape Wine Res.* 25, 134–141. doi: 10.1111/ajgw.12371
- Rogiers, S. Y., Greer, D. H., Hutton, R. J., and Landsberg, J. J. (2009). Does nighttime transpiration contribute to anisohydric behaviour in a *Vitis vinifera* cultivar? *J. Exp. Bot.* 60, 3751–3763. doi: 10.1093/jxb/erp217
- Rogiers, S. Y., Holzapfel, B. P., and Smith, J. P. (2011a). Sugar accumulation in roots of two grape varieties with contrasting response to water stress. *Ann. Appl. Biol.* 159, 399–413. doi: 10.1111/j.1744-7348.2011.00505.x
- Rogiers, S. Y., Smith, J. P., Holzapfel, B. P., and Hardie, W. J. (2011b). Soil temperature moderates grapevine carbohydrate reserves after bud break and conditions fruit set responses to photoassimilatory stress. *Funct. Plant Biol.* 38, 899–909. doi: 10.1071/FP10240
- Rosenzweig, C., Iglesias, A., Yang, X. B., Epstein, P. R., and Chivian, E. (2001). Climate change and extreme weather events—implications for food production, plant diseases, and pests. *Glob. Change Hum. Health* 2, 90–104. doi: 10.1023/A:1015086831467
- Ruperti, B., Botton, A., Populin, F., Eccher, G., Brilli, M., Quaggiotti, S., et al. (2019). Flooding responses on grapevine: A physiological, transcriptional, and metabolic perspective. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.00339
- Sadras, V. O., and McCarthy, M. G. (2007). Quantifying the dynamics of sugar concentration in berries of *Vitis vinifera* cv. Shiraz: a novel approach based on allometric analysis. *Aust. J. Grape Wine Res.* 13, 66–71. doi: 10.1111/j.1755-0238.2007.tb00236.x
- Sadras, V. O., and Moran, M. A. (2012). Elevated temperature decouples anthocyanins and sugars in berries of Shiraz and Cabernet franc. *Aust. J. Grape Wine Res.* 18, 115–122. doi: 10.1111/j.1755-0238.2012.00180.x
- Sadras, V. O., and Petrie, P. R. (2011). Climate shifts in south-eastern Australia: early maturity of Chardonnay, Shiraz and Cabernet sauvignon is associated with early onset rather than faster ripening. *Aust. J. Grape Wine Res.* 17, 199–205. doi: 10.1111/j.1755-0238.2011.00138.x
- Sadras, V., Petrie, P., Moran, M., Bastian, S., and Taylor, D. (2013). Decompressing harvest and preserving wine style in warming climates. *Aust. N.Z. Grapegrow. Winemak.* 594, 47.
- Saliba, A. J., Ovington, L. A., and Moran, C. C. (2013). Consumer demand for low-alcohol wine in an Australian sample. *Int. J. Wine Res.* 5, 1–8. doi: 10.2147/IJWR.S41448
- Salinari, F., Giosue, S., Tubiello, F. N., Rettori, A., Rossi, V., Spanna, F., et al. (2006). Downy mildew (*Plasmopara viticola*) epidemics on grapevine under climate change. *Glob. Change Biol.* 12, 1299–1307. doi: 10.1111/j.1365-2486.2006.01175.x
- Salvi, L., Brunetti, C., Cataldo, E., Niccolai, A., Centritto, M., Ferrini, F., et al. (2019). Effects of *Ascochylla nodosum* extract on *Vitis vinifera*: Consequences on plant physiology, grape quality and secondary metabolism. *Plant Physiol. Biochem.* 139, 21–32. doi: 10.1016/j.plaphy.2019.03.002
- Santos, J. A., Fraga, H., Malheiro, A. C., Moutinho-Pereira, J., Dinis, L. T., Correia, C., et al. (2020). A review of the potential climate change impacts and adaptation options for European viticulture. *Appl. Sci.* 10, 3092. doi: 10.3390/app10093092
- Sassu, A., Gambella, F., Ghiani, L., Mercenaro, L., Caria, M., and Pazzona, A. L. (2021). Advances in unmanned aerial system remote sensing for precision viticulture. *Sensors* 21, 956. doi: 10.3390/s21030956
- Sauter, M. (2013). Root responses to flooding. *Curr. Opin. Plant Biol.* 16, 282–286. doi: 10.1016/j.pbi.2013.03.013
- Schimel, J. P. (2018). Life in dry soils: effects of drought on soil microbial communities and processes. *Annu. Rev. Ecol. Evol. Syst.* 49, 409–432. doi: 10.1146/annurev-ecolsys-110617-062614
- Schmidt, H. P., Kammann, C., Niggli, C., Evangelou, M. W., Mackie, K. A., and Abiven, S. (2014). Biochar and biochar-compost as soil amendments to a vineyard soil: Influences on plant growth, nutrient uptake, plant health and grape quality. *Agric. Ecosyst. Environ.* 191, 117–123. doi: 10.1016/j.agee.2014.04.001
- Schultz, H. R. (2003). Differences in hydraulic architecture account for near-isohydric and anisohydric behaviour of two field-grown *Vitis vinifera* L. cultivars during drought. *Plant Cell Environ.* 26, 1393–1405. doi: 10.1046/j.1365-3040.2003.01064.x
- Schultz, H. R., and Jones, G. V. (2010). Climate induced historic and future changes in viticulture. *J. Wine Res.* 21, 137–145. doi: 10.1080/09571264.2010.530098
- Seng, K. P., Ang, L. M., Schmidtke, L. M., and Rogiers, S. Y. (2018). Computer vision and machine learning for viticulture technology. *IEEE Access* 6, 67494–67510. doi: 10.1109/ACCESS.2018.2875862
- Sharma, P., Jha, A. B., Dubey, R. S., and Pesarakli, M. (2012). Reactive oxygen species, oxidative damage, and antioxidative defense mechanism in plants under stressful conditions. *J. Bot.* 2012, 217037. doi: 10.1155/2012/217037
- Smith, K. A., Ball, T., Conen, F., Dobbie, K. E., Massheder, J., and Rey, A. (2018). Exchange of greenhouse gases between soil and atmosphere: interactions of soil physical factors and biological processes. *Eur. J. Soil Sci.* 69, 10–20. doi: 10.1111/ejss.12539
- Steel, C. C., Greer, L. A., Savocchia, S., and Samuelian, S. K. (2011). Effect of temperature on *Botrytis cinerea*, *Colletotrichum acutatum* and *Greeneria uvicola* mixed fungal infection of *Vitis vinifera* grape berries. *Vitis* 50, 69–71. doi: 10.5073/vitis.2011.50.69-71
- Steenwerth, K., and Belina, K. M. (2008). Cover crops enhance soil organic matter, carbon dynamics and microbiological function in a vineyard agroecosystem. *Appl. Soil Ecol.* 40, 359–369. doi: 10.1016/j.apsoil.2008.06.006
- Stevens, R. M., and Walker, R. R. (2002). Response of grapevines to irrigation-induced saline-sodic soil conditions. *Aust. J. Exp. Agric.* 42, 323–331. doi: 10.1071/EA00143
- Striegler, R. K., Howell, G. S., and Flore, J. A. (1993). Influence of rootstock on the response of seveal grapevines to flooding stress. *Am. J. Enol. Vitic.* 44, 313–319.
- Suddick, E. C., Steenwerth, K., Garland, G. M., Smart, D. R., and Six, J. (2011). "Discerning agricultural management effects on nitrous oxide emissions from conventional and alternative cropping systems: a California case study," in *Understanding greenhouse gas emissions from agricultural management*. Eds. L. Guo, A. Gunasekara and L. McConnell (Washington D.C: American Chemical Society), 203–226. doi: 10.1021/bk-2011-1072.ch012
- Šuklje, K., Zhang, X., Antalick, G., Clark, A. C., Deloire, A., and Schmidtke, L. M. (2016). Berry shriveling significantly alters Shiraz (*Vitis vinifera* L.) grape and wine chemical composition. *J. Agric. Food Chem.* 64, 870–880. doi: 10.1021/acs.jafc.5b05158
- Summerson, V., Gonzalez Viejo, C., Pang, A., Torrico, D. D., and Fuentes, S. (2021). Review of the effects of grapevine smoke exposure and technologies to assess smoke contamination and taint in grapes and wine. *Beverages* 7, 7. doi: 10.3390/beverages7010007

- Sweetman, C., Sadras, V. O., Hancock, R. D., Soole, K. L., and Ford, C. (2014). Metabolic effects of elevated temperature on organic acid degradation in ripening *vitis vinifera* fruit. *J. Exp. Bot.* 65, 5975–5988. doi: 10.1093/jxb/eru343
- Tardaguila, J., Stoll, M., Gutiérrez, S., Proffitt, T., and Diago, M. P. (2021). Smart applications and digital technologies in viticulture: A review. *Smart Agric. Tech.* 1, 100005. doi: 10.1016/j.atech.2021.100005
- Taskos, D., Stamatiadis, S., Yvin, J. C., and Jamois, F. (2019). Effects of an *Ascophyllum nodosum* (L.) Le Jol. extract on grapevine yield and berry composition of a Merlot vineyard. *Sci. Hortic.* 250, 27–32. doi: 10.1016/j.scienta.2019.02.030
- Tilbrook, J., and Tyerman, S. D. (2008). Cell death in grape berries: varietal differences linked to xylem pressure and berry weight loss. *Funct. Plant Biol.* 35, 173–184. doi: 10.1071/FP07278
- Tilbrook, J., and Tyerman, S. D. (2009). Hydraulic connection of grape berries to the vine: varietal differences in water conductance into and out of berries, and potential for backflow. *Funct. Plant Biol.* 36, 541–550. doi: 10.1071/FP09019
- Urhausen, S., Brien, S., Kapala, A., and Simmer, C. (2011). Climatic conditions and their impact on viticulture in the upper Moselle region. *Clim. Change* 109, 349–373. doi: 10.1007/s10584-011-0059-z
- Van der Hulst, L., Munguia, P., Culbert, J. A., Ford, C. M., Burton, R. A., and Wilkinson, K. L. (2019). Accumulation of volatile phenol glycoconjugates in grapes following grapevine exposure to smoke and potential mitigation of smoke taint by foliar application of kaolin. *Planta* 249, 941–952. doi: 10.1007/s00425-018-03079-x
- Van Leeuwen, C., and Destrac-Irvine, A. (2017). Modified grape composition under climate change conditions requires adaptations in the vineyard. *Oeno One* 51, 147–154. doi: 10.20870/oeno-one.2016.0.0.1647
- Van Leeuwen, C., Destrac-Irvine, A., Dubernet, M., Duchêne, E., Gowdy, M., Marguerit, E., et al. (2019). An update on the impact of climate change in viticulture and potential adaptations. *Agronomy* 9, 514. doi: 10.3390/agronomy9090514
- Van Leeuwen, C., Friant, P., Chone, X., Tregoat, O., Koundouras, S., and Dubourdieu, D. (2004). Influence of climate, soil, and cultivar on terroir. *Am. J. Enol. Vitic.* 55, 207–217. doi: 10.5344/ajev.2004.55.3.207
- Van Leeuwen, C., Schultz, H. R., de Cortazar-Atauri, I. G., Duchêne, E., Ollat, N., Pieri, P., et al. (2013). Why climate change will not dramatically decrease viticultural suitability in main wine-producing areas by 2050. *Proc. Natl. Acad. Sci.* 110, E3051–E3052. doi: 10.1073/pnas.1307927110
- Van Leeuwen, C., and Seguin, G. (2006). The concept of terroir in viticulture. *J. Wine Res.* 17, 1–10. doi: 10.1080/09571260600633135
- Viers, J. H., Williams, J. N., Nicholas, K. A., Barbosa, O., Kotzé, I., Spence, L., et al. (2013). Viticulture: pairing wine with nature. *Conserv. Lett.* 6, 287–299. doi: 10.1111/conl.12011
- Walker, R., and Clingeleffer, P. (2016). Potassium accumulation by grapevines and potassium-pH inter-relationships in grape juice and wine. *Aust. N.Z. Grapegrow. Winemak.* 626, 46.
- Walker, R. R., Read, P. E., and Blackmore, D. H. (2000). Rootstock and salinity effects on rates of berry maturation, ion accumulation and colour development in Shiraz grapes. *Aust. J. Grape Wine Res.* 6, 227–239. doi: 10.1111/j.1755-0238.2000.tb00183.x
- Webb, L. B., Whetton, P. H., and Barlow, E. W. R. (2011). Observed trends in winegrape maturity in Australia. *Glob. Change Biol.* 17, 2707–2719. doi: 10.1111/j.1365-2486.2011.02434.x
- Webb, L. B., Whetton, P. H., Bhend, J., Darbyshire, R., Briggs, P. R., and Barlow, E. W. R. (2012). Earlier wine-grape ripening driven by climatic warming and drying and management practices. *Nat. Clim. Change* 2, 259. doi: 10.1038/nclimate1417
- Webb, L., Whiting, J., Watt, A., Hill, T., Wigg, F., Dunn, G., et al. (2010). Managing grapevines through severe heat: A survey of growers after the 2009 summer heatwave in south-eastern Australia. *J. Wine Res.* 21, 147–165. doi: 10.1080/09571264.2010.530106
- Wheeler, S., Loveys, B., Ford, C., and Davies, C. (2009). The relationship between the expression of abscisic acid biosynthesis genes, accumulation of abscisic acid and the promotion of *vitis vinifera* l. berry ripening by abscisic acid. *Aust. J. Grape Wine Res.* 15, 195–204. doi: 10.1111/j.1755-0238.2008.00045.x
- Williams, A. A., Karoly, D. J., and Tapper, N. (2001). The sensitivity of Australian fire danger to climate change. *Clim. Change* 49, 171–191. doi: 10.1023/A:1010706116176
- Wolfe, D. W., Schwartz, M. D., Lakso, A. N., Otsuki, Y., Pool, R. M., and Shaulis, N. J. (2005). Climate change and shifts in spring phenology of three horticultural woody perennials in northeastern USA. *Int. J. Biometeorol.* 49, 303–309. doi: 10.1007/s00484-004-0248-9
- Xiao, Z., DeGaris, K. A., Baby, T., McLoughlin, S. J., Holzapfel, B. P., Walker, R. R., et al. (2020). Using rootstocks to lower berry potassium concentrations in 'Cabernet sauvignon' grapevines. *Vitis* 59, 117–126. doi: 10.5073/vitis.2020.59.117-126
- Xiao, Z., Liao, S., Rogiers, S. Y., Sadras, V. O., and Tyerman, S. D. (2018a). Effect of water stress and elevated temperature on hypoxia and cell death in the mesocarp of Shiraz berries. *Aust. J. Grape Wine Res.* 24 (4), 487–497. doi: 10.1111/ajgw.12363
- Xiao, Z., Rogiers, S. Y., Sadras, V. O., and Tyerman, S. D. (2018b). Hypoxia in grape berries: the role of seed respiration and lenticels on the berry pedicel and the possible link to cell death. *J. Exp. Bot.* 69, 2071–2083. doi: 10.1093/jxb/ery039
- Yang, T., Siddique, K. H., and Liu, K. (2020). Cropping systems in agriculture and their impact on soil health- a review. *Glob. Ecol. Conserv.* 23, e01118. doi: 10.1016/j.gecco.2020.e01118
- Zhang, Y., Chen, W., Smith, S. L., Riseborough, D. W., and Cihlar, J. (2005). Soil temperature in Canada during the twentieth century: Complex responses to atmospheric climate change. *J. Geophys. Res. Atmos.* 110 (D3), D03112. doi: 10.1029/2004JD004910

# The Roads to Haploid Embryogenesis

Kun Shen <sup>1,†</sup>, Mengxue Qu <sup>1,†</sup> and Peng Zhao <sup>1,2,\*</sup> <sup>1</sup> State Key Laboratory of Hybrid Rice, College of Life Sciences, Wuhan University, Wuhan 430072, China<sup>2</sup> Hubei Hongshan Laboratory, Wuhan 430070, China

\* Correspondence: pzha02000@whu.edu.cn

† These authors contribute equally to this work.

**Abstract:** Although zygotic embryogenesis is usually studied in the field of seed biology, great attention has been paid to the methods used to generate haploid embryos due to their applications in crop breeding. These mainly include two methods for haploid embryogenesis: in vitro microspore embryogenesis and in vivo haploid embryogenesis. Although microspore culture systems and maize haploid induction systems were discovered in the 1960s, little is known about the molecular mechanisms underlying haploid formation. In recent years, major breakthroughs have been made in in vivo haploid induction systems, and several key factors, such as the matrilineal (MTL), baby boom (BBM), domain of unknown function 679 membrane protein (DMP), and egg cell-specific (ECS) that trigger in vivo haploid embryo production in both the crops and Arabidopsis models have been identified. The discovery of these haploid inducers indicates that haploid embryogenesis is highly related to gamete development, fertilization, and genome stability in early embryos. Here, based on recent efforts to identify key players in haploid embryogenesis and to understand its molecular mechanisms, we summarize the different paths to haploid embryogenesis, and we discuss the mechanisms of haploid generation and its potential applications in crop breeding. Although these haploid-inducing factors could assist egg cells in bypassing fertilization to initiate embryogenesis or trigger genome elimination in zygotes after fertilization to form haploid embryos, the fertilization of central cells to form endosperms is a prerequisite step for haploid formation. Deciphering the molecular and cellular mechanisms for haploid embryogenesis, increasing the haploid induction efficiency, and establishing haploid induction systems in other crops are critical for promoting the application of haploid technology in crop breeding, and these should be addressed in further studies.

**Keywords:** parthenogenesis; haploid induction; single fertilization; genome elimination; crop breeding



**Citation:** Shen, K.; Qu, M.; Zhao, P. The Roads to Haploid Embryogenesis. *Plants* **2023**, *12*, 243. <https://doi.org/10.3390/plants12020243>

Academic Editor: Othmane Merah

Received: 14 November 2022

Revised: 19 December 2022

Accepted: 30 December 2022

Published: 5 January 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Although zygotic embryogenesis is the primary way to generate embryos in plant reproduction, there are still several alternative methods for forming haploid or diploid embryos. In addition to normal zygotic embryogenesis, in recent decades, a great deal of attention has been paid to the formation of haploid embryos due to their important applications in crop breeding. Haploid induction is an effective way to shorten breeding time and has been used in crop breeding for many years. Traditional breeding requires 7–8 generations to obtain ideal homozygous plants during the cross, while haploid technology can shorten the breeding time to 2–3 generations, which will greatly save on breeding time and reduce the cost of breeding [1]. Haploid embryogenesis is the core of haploid breeding technology.

According to the existing methods of generating haploid embryos, a haploid induction system can be divided into two categories: in vitro and in vivo methods. Microspore embryogenesis is the primary in vitro method used to induce haploid embryos. It was first developed in 1964 and has been employed in over 75 species. The in vitro methods for haploid embryogenesis are mainly achieved through the suitable in vitro cell culture system, which promotes haploid microspores to be reprogrammed and enter into embryogenic pathways to generate haploid embryos. However, these in vitro culture systems have not been established for many plant species, and the mechanisms of microspore embryogenesis

remain largely unknown. The existing *in vivo* haploid induction systems can be further divided into CENH3-mediated haploid embryogenesis, parental factor-induced haploid embryogenesis, and transcriptional factor-triggered haploid embryogenesis. The haploid inducer line in maize was first discovered in 1959, but the mechanisms underlying haploid formation have long remained a mystery. The recent discoveries of several key molecular players in maize and *Arabidopsis* haploid induction have greatly aided us in understanding the mechanisms of haploid embryogenesis, demonstrating that defects in gamete development and fertilization are the fundamental mechanisms for haploid generation.

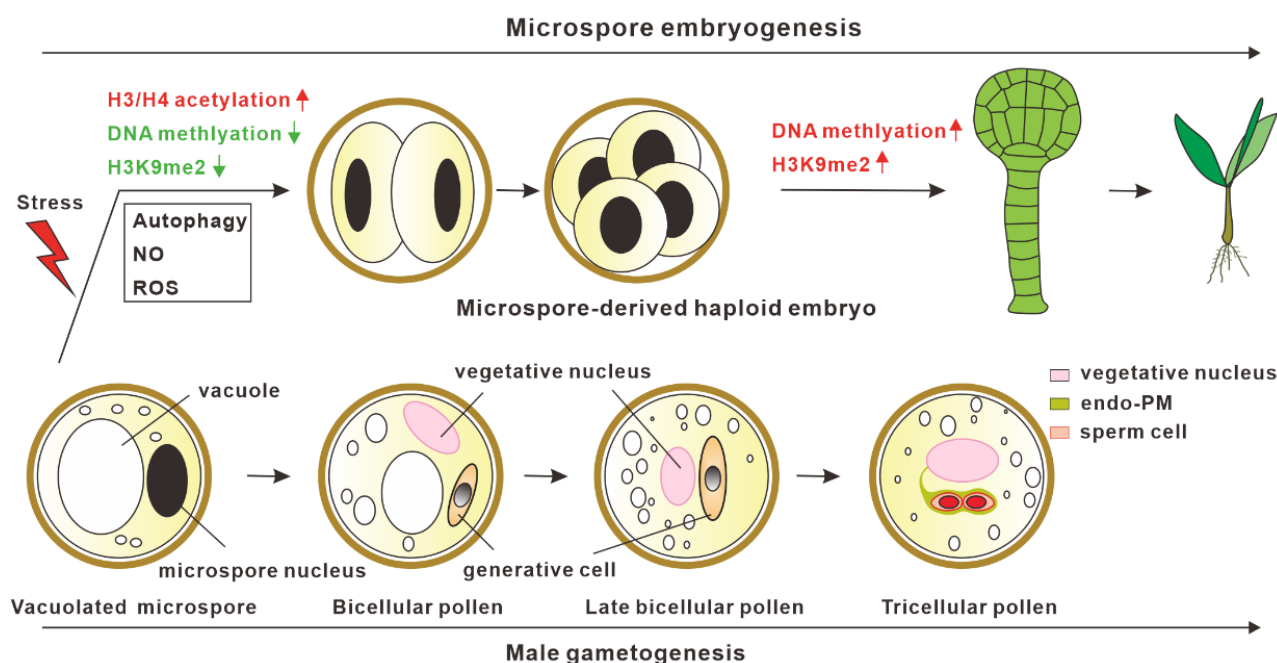
In this review, we summarize the recent advances in the methods for generating haploid embryogenesis, and we primarily focus on microspore embryogenesis, CENH3-mediated haploid embryogenesis, and parental factor-induced and transcription factor-triggered haploid embryogenesis. We also discuss the mechanisms underlying these methods and the potential applications of haploid embryogenesis in crop breeding.

### 1. Microspore Embryogenesis

Microspore embryogenesis is a type of embryogenesis in which haploid microspores after stress treatment undergo cell reprogramming and shift into embryogenic pathways to generate haploid embryos [2–8]. The haploid embryos then can be automatically doubled or chemically treated to produce double haploids, thus reducing the time required to obtain homozygous plants [9,10]. Among these microspore embryogenesis systems, exine-dehisced *Brassica napus* microspores treated by physical stress can induce polarization and develop into typical embryos with the differentiation of an embryo proper and suspensor [11]. However, microspore embryogenesis induction systems have not been well established in many other plants, such as *Arabidopsis thaliana* and *Solanum lycopersicum*, indicating that the embryogenic potential of microspores or signals for triggering cell fate reprogramming into embryogenic pathways may vary in different species.

In microspore embryogenesis, the reprogramming of microspores into the embryo developmental pathway after stress treatment is the critical step for haploid embryo generation. Epigenetic mechanisms, including H3/ H4 deacetylation, DNA methylation, and H3K9me2, are reported to be involved in microspore embryogenesis [12]. The inhibition of histone deacetylases (HDAC) activities by the chemical inhibitor Trichostatin A (TSA) can efficiently promote microspore embryogenesis in both *B. napus* and *A. thaliana*. Suppression of HDAC activities by TSA leads to the hyperacetylation of histones H3 and H4, which results in the upregulation of genes related to cell wall remodeling, cell division, and embryogenesis [13]. In addition to *B. napus* and *A. thaliana*, TSA treatment has been efficient at improving the rate of microspore embryogenesis in several other plants, including pakchoi [14], wheat [15,16], and barley [17], indicating that the HDAC-dependent mechanism in microspore embryogenesis appears to be conserved in different plants. Besides suppressing HDAC activities, the inhibition of DNA methyltransferases catalytic activities by the inhibitors 5-azacytidine (AC) and 2'-deoxy-5-azacytidine (DAC) can remarkably increase the frequency of microspore embryogenesis [18], suggesting that decreased DNA methylation levels are also responsible for microspore embryogenesis, which is likely achieved through increased chromatin accessibility for transcription activation. Similar results in promoting microspore embryogenesis were also observed when the microspores were treated with BIX-01294, which can efficiently inhibit the activities of H3K9me2 methyltransferase [19]. In summary, a low DNA methylation level, low H3K9me2 level, and high acetylation level are critical for the initiation of microspore embryogenesis, whereas, after reprogramming, microspore embryogenesis is accompanied by increased DNA methylation levels and H3K9me2 levels [19,20] at later developmental stages, which may promote haploid embryo differentiation (Figure 1).





**Figure 1.** Model for in vitro microspore embryogenesis. After stress treatment, vacuolated microspores switch from normal male gametogenesis into the embryogenic pathway. Epigenetic modifications, autophagy, reactive oxygen species (ROS), and nitric oxide (NO) are involved in promoting the initiation of microspore embryogenesis.

Besides epigenetic modifications, autophagy, reactive oxygen species (ROS), and nitric oxide (NO) have been shown to be involved in the initiation process of microspore embryogenesis (Figure 1) [12,21,22]. Recent studies have revealed that autophagy-mediated cytoplasm clearance is not only critical for promoting pollen germination [23] and pollen tube growth [24], but it is also important for microspore embryogenesis [25–27]. Two independent studies have demonstrated that the cell death of microspores after stress treatment is accompanied by the activation of autophagy, and blocking autophagy prevented the cell death of microspores and increased the frequency of microspore embryogenesis in *B. napus*, indicating that enhanced autophagic activities play a role in preventing microspore embryogenesis [27]. A similar role of autophagy in *Hordeum vulgare* microspore embryogenesis was also found. The suppression of autophagosome formation by 3-methyladenine or the inhibition of autophagic body degradation in the vacuoles by E-64 can promote microspore embryogenesis [26]. In addition, ROS and NO were also shown to play important roles in response to stress-induced microspore cell death and microspore reprogramming in barley microspore embryogenesis [28]. Stress-treated microspore embryonic suspension exhibited high ROS levels, high NO signals, and enhanced microspore cell deaths [28,29]. Treating microspores with  $MnCl_2$  ( $O_2^-$  scavenger), ascorbate ( $H_2O_2$  scavenger), and cP-TIO (NO scavenger) led to reduced cell death and increased embryogenesis initiation efficiency. Hence, investigating these mechanisms will not only help us understand the mechanism for microspore embryogenesis initiation, but also provide an opportunity to improve haploid induction.

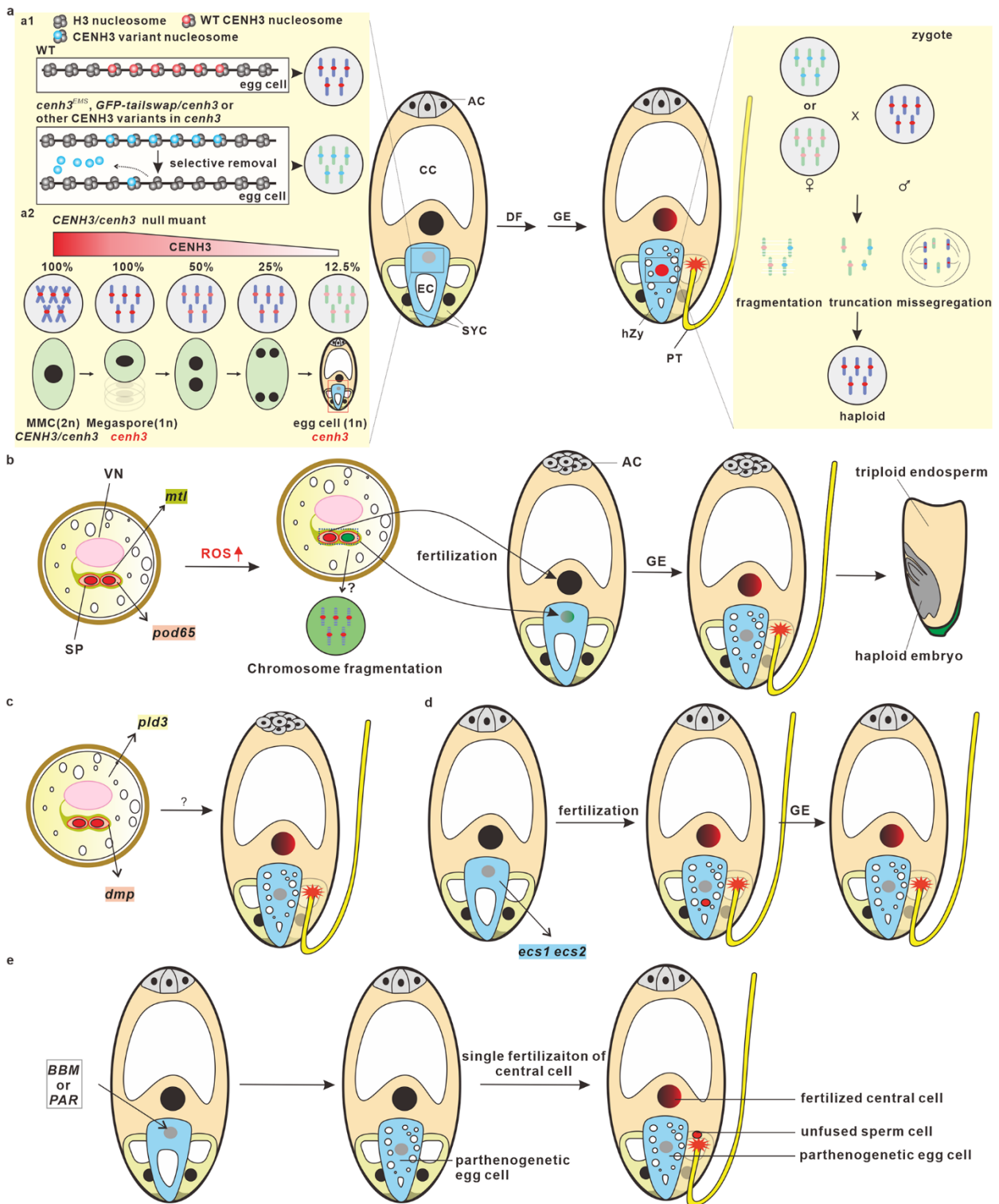
## 2. CENH3-Mediated Haploid Embryogenesis

CENH3 (CENP-A in humans, Cse4p in *Saccharomyces cerevisiae*, and HTR12 in *Arabidopsis*) [30,31] is the centromere-specific histone H3 variant which contains a variable N-terminal domain (NTD) and a conserved histone fold domain (HFD). The core functions of CENH3 primarily include two aspects: recruiting CENH3/H4 reloading factor to nucleosomes and providing a platform for kinetochore binding and assembling. The NTD of CENH3 is essential for CENH3 loading onto the centromeres of meiotic chromo-

somes, rather than its deposition and function in mitotic nuclei [32,33]. Recent studies have revealed that both the quality (stability of CENH3 on nucleosomes and recognizability of CENH3 by CENH3/H4 reloading factors or kinetochore complex proteins) and quantity (loading amount of CENH3 nucleosomes on centromeres) of CENH3 in centromeres are important for CENH3's function [34–39], and defects in CENH3 will lead to chromosome elimination.

Based on the characteristics of CENH3, GFP-CENH3 and GFP-tailswap (GFP fused with the N-terminal tail of a CENH3 variant, with NTD replaced by conserved Histone3.3 NTD) were designed to rescue the developmental defects in a *cenh3-1* mutant. Expressing GFP-tailswap in a *cenh3-1* mutant can partially rescue the embryonic lethal phenotype, but it is accompanied by severe male sterility [33,40]. More importantly, expressing GFP-tailswap or GFP-CENH3 in a *cenh3-1* mutant can produce aneuploid and haploid plants due to chromosome elimination. When GFP-tailswap/*cenh3-1* pistils were pollinated by wild-type (WT) pollen grains, over 30% of the progenies were haploids, though with paternal genomes. Haploid-induced rates (HIR) slightly decreased when GFP-tailswap expressing pollen grains were used for the cross with the WT plants. Expressing CENH3 from *B. rapa*, *Lepidium oleraceum*, or *Z. mays* in the Arabidopsis *cenh3-1* mutant can also lead to haploid formation, which is similar to the expression GFP-CENH3 or GFP-tailswap in *cenh3-1* [41]. In addition, point mutations in CENH3 can also induce haploids during the outcrossing with WT plants [39,42,43]. Under this scenario, the CENH3-mediated haploid induction system appears to be conserved in different plants.

The potential mechanisms underlying CENH3-mediated haploid induction have been summarized and discussed, and several hypotheses and theories have been proposed in recent reviews [36,38,44–52]. CENH3-induced haploids are the result of post-fertilization genome elimination [53]. Based on recent published results, fertilization is normally completed to form zygotes, but uniparental genomes will be eliminated at early embryonic developmental stages, and eventually they will form the haploid mature embryos. Mutations or modifications in CENH3 may impair the recruitment of CENH3 to centromeres or reduce their stability, resulting in the selective dispossession of CENH3 variants from centromeres in egg cells and zygotes [54]. In a heterozygous *cenh3* null mutant, the CENH3 amount on the centromeres of the *cenh3* embryo sac are significantly diluted during post-meiotic cell division, prior to gamete formation [55]. The inconsistency of the centromere strength caused by CENH3 among the two parental genomes in zygotes thereby leads to the delay of CENH3 reloading and/or kinetochore assembly in uniparental genomes, resulting in chromosome missegregation, truncation, or fragmentation. Chromosomal irregularities can produce micronuclei or lead to the missegregation of chromosomes [56]. The chromosomal truncation and fragmentation may form micronuclei, which may re-emerge with the nucleus to produce aneuploids or eventually degrade to produce haploid embryos (Figure 2a).



**Figure 2.** Model of in vivo haploid embryogenesis. (a) CENH3-mediated haploid embryogenesis is accomplished through post-fertilization genome elimination. This model shows CENH3-mediated maternal genome elimination after fertilization. In a *cenh3<sup>EMS</sup>* mutant, the *GFP-tailswap/cenh3* mutant or *cenh3* mutants which express the CENH3 variant from other species, selectively eviction of CENH3 variants in mature egg cells and the zygotes produce “weak” centromeres (a1). In a heterozygous model of a *cenh3* null mutant, the quantity of CENH3 on the centromeres is significantly diluted after meiosis, which also generates “weak” centromeres in *cenh3* egg cells (a2). After being pollinated by wild-type pollen grains, the maternal genome (gray), which contains “weak” centromeres,

rather than the normal paternal genome (red), is selectively eliminated in zygotes and early embryos. (b) Mutations in *MTL* and *POD* induces haploid embryogenesis through ROS-triggered spermatid chromosome fragmentation. *MTL* is localized at the endo-PM of vegetative cells while *POD65* is localized in sperm cells. An ROS level increase leads to spermatid chromosome fragmentation. Sperm cells (green) with fragmented chromosomes can fertilize egg cells, but the paternal genomes will be eliminated after fertilization (gray). (c) Vegetative cell-expressed *PLD3* and sperm cell-expressed *DMP* may induce maternal haploid embryogenesis. (d) Mutation in egg cell-specifically expressed *ECS1/2* induces haploid embryogenesis, which may be due to the karyogamy defect and genome elimination after fertilization. (e) The ectopic expression of *BBM* and *PAR* in egg cells induces haploid embryogenesis through parthenogenesis. WT, wild-type; MMC, megaspore mother cell; AC, antipodal cell; CC, central cell; EC, egg cell; SYC, synergid cell; hZy, haploid zygote; DF, double fertilization; GE, genome elimination; PT, pollen tube.

### 3. Parental Factor-Mediated Haploid Embryogenesis

Although the paternal inducer line of maize (Stock6), which can be used to induce maternal haploids, was discovered in 1959 [57], the specific gene responsible for haploid generation was not identified until recently. In this section, we will summarize the advances in identifying both maternal and parental factors that could induce haploid embryos and discuss the mechanism underlying these parental factors in haploid embryogenesis.

#### 3.1. Paternal Players in Haploid Embryogenesis

Four paternal factors (*MTL/PLA1/NLD*, *DMP*, *PLD3*, and *POD65*) have been identified from maize (Figure 2b,c). *RMZM2G47124*, the first-identified maternal haploid-inducing gene cloned from *qhir1*, was named *MATRILINEAL (MTL)* [58], *PHOSPHOLIPASE A1 (PLA1)* [59], and *NOT LIKE DAD (NLD)* [60] by three different groups. *MTL/PLA1/NLD* (hereafter referred to as *MTL*) belongs to the phospholipase family and is an enzyme that hydrolyzes the phospholipids that function in membrane remodeling [61–63]. *MTL* is localized at the pollen endo-plasma membrane, a special plasma membrane that originates in the plasma membrane of vegetative cells and closely surrounds two sperm cells [64]. A 4 bp insertion leads to a frame-shift mutation in the *MTL*, which results in seed abortion and haploid induction. The second-identified maternal haploid inducer gene is *ZmDMP*, which encodes a DUF679 domain membrane protein and is specifically expressed in sperm cells. *ZmDMP* mutation induced an HIR of 0.1~0.3%, but it significantly increased HIR at a five-to-six-fold in combination of *MTL* mutation [65]. In addition to maize, recent studies have demonstrated that the role of *DMP* in haploid induction is conserved between monocots and eudicots, which has been achieved in *A. thaliana*, *B. napus* [66,67], *B. oleracea* [68], *Medicago truncatula* [69], *Nicotiana tabacum* [67,70], *S. lycopersicum* [71], and *S. tuberosum* [72]. In addition to *MTL* and *DMP*, vegetative cell-expressed *ZmPLD3* and sperm cell-expressed *ZmPOD65* were also shown to be able to induce haploids in maize [73,74]. *ZmPLD3* belongs to phospholipase D (PLD) family, and it is localized in the ER, plastids, the Golgi apparatus, and the cytosol of vegetative cells [73], whereas *ZmPOD65* encodes a peroxidase (POD) protein and is highly expressed in pollen at the tricellular stage [74].

#### 3.2. Maternal Factor in Haploid Embryogenesis

In addition to the paternal factors in haploid embryogenesis, recent reports have demonstrated that egg cell-expressed maternal factors can also be used to induce haploid embryos. *EGG CELL-SPECIFIC1/2 (ECS1/2)* encodes egg cell-specifically expressed aspartic proteases, which are secreted to the synergid cell region upon fertilization to avoid polytubey by degrading the pollen tube attractant *LURE1* [75]. Recently, two independent studies demonstrated that the mutation of *ECS1* and *ECS2* can also induce haploid embryogenesis [76,77]. Unfused sperm nuclei were observed in zygotes and early embryos, suggesting that karyogamy defects occurred in the sperm–egg fertilization, and the haploids from the *ecs1 ecs2* mutant progenies may have resulted from the post-fertilization



genomic elimination [77]. In summary, ECS-mediated haploid induction is likely caused by pseudogamy and potential post-fertilization genomic elimination (Figure 2d).

### 3.3. Synergistic Effects on Haploid Embryogenesis

Since a low efficiency of haploid embryogenesis is observed in most inducer lines, efficiency has become a major barrier for its application in crop breeding. To improve the efficiency of haploid embryogenesis, the combination of *MTL* and *DMP* [65] or *MTL* and *PLD3* [73] was performed to test whether they could improve haploid production. The mutation of *PLD3* or *DMP* in *mtl* mutant background could significantly increase its haploid induction rate, but the HIR was still lower than the expected. A *dmp-mtl-pld3* triple mutant was also created to test whether it could increase the HIR. However, *dmp-mtl-pld3* triple-homozygous plants could not be obtained, which was likely due to the pollen developmental defect or the fertilization defect in triple mutants. Hence, it is still worth testing other combinations of haploid inducer genes to improve the efficiency of haploid induction in further studies.

### 3.4. Mechanism for Haploid Embryogenesis

Two mechanisms have been proposed to explain the haploid formation in maize. The first is a single fertilization-induced haploid and the second is post-fertilization genome elimination. The former may produce haploid embryos through the parthenogenesis of the egg cell while the central cell fertilizes normally to form the endosperm. In the latter, double fertilization occurs normally, but the zygote undergoes uniparental genome elimination that results in haploid embryo formation.

The mechanisms for haploid embryogenesis are primarily focused on *MTL*-induced haploid embryogenesis, and whether it is conserved among different inducers, such as *DMP*, remains largely unknown. Several recent studies have demonstrated that *MTL*-induced haploids may form through post-fertilization genome elimination, rather than single fertilization-mediated parthenogenesis. The markers, including B chromosomes and CENH3-YFP derived from the paternal genome, were detected in haploid progenies [78], suggesting that the egg cells were fertilized successfully, and that uniparental genome elimination occurred during haploid induction. In addition, when WT pistils were pollinated by *mtl* mutant pollen grains (which were carrying the *Cas9* and *gRNA* expression cassette) [79], genome-edited haploids without the *Cas9* expression cassette were detected in the progenies. Since the CRISPR/*Cas9* system only exists in the paternal genome, this result strongly suggests that the paternal genome is transmitted to the egg cell upon fertilization and is eventually eliminated after fertilization. A multi-omics analysis of *mtl* pollen grains revealed that ROS signals were involved in post-fertilization genome elimination [74]. Elevated ROS levels may cause DNA damage in pollen from the *mtl* mutant [80], which may lead to chromosome fragmentation in sperm cells and then induce genome elimination after fertilization. Spermatid chromosome fragmentation in the CAU5 haploid inducer line was detected through single nucleus sequencing [81]. In addition, the in vitro treatment of pollen grains with ROS-inducing reagents can also result in sperm DNA fragmentation and lead to the formation of haploids when pollinated to the WT plants. Studies on the sperm cell-expressed peroxidase gene *ZmPOD65* further confirmed the role of ROS in haploid embryogenesis [74]. Peroxidases widely exist in the plant kingdom [82], and they convert hydrogen peroxide ( $H_2O_2$ ) into  $H_2O$  during the POD catalytic reaction [83]. Therefore, *POD65* may be functional in the removal of  $H_2O_2$  in sperm cells, and *POD65* mutation can cause  $H_2O_2$  to burst in sperm cells, which leads to sperm DNA fragmentation and, eventually, haploid production.

In summary, vegetative cell-expressed *MTL* and sperm cell-expressed *POD65* are involved in the regulation of ROS levels, as well as the expression of ROS-related genes in pollen grains. Mutations in *MTL* and *POD65* do not appear to impair fertilization, but the zygotes generated from the cross between *mtl* or *pod65* mutants and WT plants will undergo genome elimination to form haploid embryos (Figure 2b). An *MTL* mutation may

cause an imbalance in the hydrolyzed phospholipids (such as PC) between sperm cells and vegetative cells, resulting in the over-accumulation of PCs in sperm cells, which disrupts mitochondrial homeostasis and leads to increased ROS levels [84]. Elevated ROS levels in sperm cells will induce DNA damage and impair the expression of ROS-related genes, leading to the chromatin fragmentation. The fragmented chromatins of sperm cells are then eliminated after fertilization, leading to the formation of haploid embryos. This genome elimination mechanism is different from that in CENH3-mediated haploid embryogenesis. In the *MTL*- or *POD65*-induced haploid embryos, chromosome fragmentation occurs before fertilization and the fragmented paternal genome cannot function properly after fertilization and is eventually eliminated, whereas the CENH3-induced post-fertilization genome elimination is largely due to the incompatibility of the parental genome.

#### 4. Transcription Factors Triggered Haploid Embryogenesis

Normal zygotic embryogenesis is triggered by the fusion of a sperm cell and an egg cell. However, recent reports have demonstrated that egg cells can bypass fertilization to trigger embryogenesis by expressing the transcriptional factors BABY BOOM (BBM) or PARTHENOGENESIS (PAR). BBM, an AINTEGUMENTA-like (AIL) transcription factor, belongs to the APETALA2/ETHYLENE RESPONSE FACTOR (AP2/ERF) family, which was previously shown to play a critical role in inducing or enhancing somatic embryogenesis [85–92]. In addition to triggering somatic embryogenesis, recent studies have demonstrated that the ectopic expression of *BBM* in egg cells can trigger egg cells to initiate embryogenesis in multiple sexual plants [93–99], indicating the conserved role of BBM in promoting embryogenesis. For example, in rice, the ectopic expression of *BBM1* in egg cells can produce embryo-like structures without fertilization [96]. An expression pattern analysis revealed that *OsBBM1* is expressed in sperm cells, and only the paternal allele of *OsBBM1* is expressed in 2.5 HAP (hours after pollination) zygotes. Both the maternal and paternal transcripts could be detected in 6.5 HAP zygotes, possibly due to the auto-activation ability of BBM. Recently, the ectopic expression of *BnBBM* in the egg cells of *A. thaliana*, *B. napus*, and *S. lycopersicon* also bypassed fertilization for embryogenesis [99], although how *BnBBM* triggered the egg cells to initiate the embryo program remains unknown. Besides *BBM*, *PAR* has also been reported to be able to induce parthenogenesis in sexual plants (Figure 2e). *PAR* encodes a zinc finger domain protein with an EAR motif, which may function as a transcription factor in dandelion (*Taraxacum officinale*). The ectopic expression of *ToPAR* in an egg cell can also promote that egg cell to produce embryo-like structures without fertilization in sexually reproductive lettuce [100]. There are two *PAR* homologs, named *DAZ3* and *TREE1*, in the *Arabidopsis* genome. It is worth investigating whether the homologs of *PAR* can also trigger egg cells to initiate embryogenesis in other plants, especially in crops.

#### 5. Application of Haploid Embryogenesis

##### 5.1. Haploid Breeding

The main application of haploid embryogenesis is to generate haploid plants which can accelerate the gain of homozygous plants and efficiently shorten breeding times. Most haploid induction systems have been tested in multiple crops (Table 1). The CENH3-mediated haploid induction system can be used to generate both maternal and paternal haploids, whereas the other haploid induction systems discussed above can only produce maternal haploids. The CENH3-mediated paternal haploid induction method can also be used to introduce a nuclear genome of interest into the targeted cytoplasm, such as the cytoplasmic male sterile (CMS) line. Therefore, the CENH3-mediated haploid induction system is also a useful tool for the establishment or improvement of the CMS line for hybrid seed production. However, CENH3-mediated haploid induction systems have not been successfully used in crop breeding [49], primarily due to the low HIR in crops and the complex steps required to construct CENH3 induction lines. In addition, the *MTL*-mediated haploid induction system has only been validated in monocots [58–60,79,80,101–104], while

the DMP-mediated haploid induction system has been confirmed in both monocots and eudicots [65–72,105]. Hence, it is worth optimizing the haploid induction systems to promote their application in crop breeding in future studies.

**Table 1.** List of in vivo haploid systems tested in various species.

Gene	Species	Type of Induction System	Reference
CENH3	<i>Arabidopsis thaliana</i>	GFP-CENH3; GFP-tailswap	[53,106]
	<i>Arabidopsis thaliana</i>	Point mutation	[39,42,43]
	<i>Arabidopsis thaliana</i>	BnCENH3; LoCENH3; ZmCENH3	[41]
	<i>Arabidopsis thaliana</i>	CRISPR/cas9 mutant	[39,79]
	<i>Zea mays</i>	GFP-tailswap	[107]
	<i>Zea mays</i>	CRISPR/cas9 mutant	[55]
	<i>Solanum lycopersicum</i>	Point mutation	[49]
	<i>Oryza sativa</i>	Point mutation	[49]
	<i>Cucumis sativus</i> L. <i>Cucumis melo</i> L.	Point mutation Point mutation	[49] [49]
MTL/PLA1/NLD	<i>Zea mays</i>	CRISPR/cas9 mutant	[58–60,79,102]
	<i>Oryza sativa</i>	TILLING	[101]
	<i>Triticum aestivum</i>	CRISPR/cas9 mutant	[80,103]
	<i>Setaria italica</i>	CRISPR/cas9 mutant	[104]
DMP	<i>Zea mays</i>	CRISPR/cas9 mutant	[65]
	<i>Arabidopsis thaliana</i>	CRISPR/cas9 mutant	[105]
	<i>Brassica napus</i>	CRISPR/cas9 mutant	[66,67]
	<i>Brassica oleracea</i>	CRISPR/cas9 mutant	[68]
	<i>Nicotiana tabacum</i>	CRISPR/cas9 mutant	[67,70]
	<i>Medicago truncatula</i>	CRISPR/cas9 mutant	[69]
	<i>Solanum lycopersicum</i> <i>Solanum tuberosum</i> L.	CRISPR/cas9 mutant CRISPR/cas9 mutant	[71] [72]
PLD3	<i>Zea mays</i>	CRISPR/cas9 mutant	[73]
POD65	<i>Zea mays</i>	CRISPR/cas9 mutant	[74]
ECS1/ECS2	<i>Arabidopsis thaliana</i>	T-DNA mutant	[76,77]
BBM	<i>Pennisetu glaucum</i>	PsASGR-BBML	[93]
	<i>Zea mays</i>	PsASGR-BBML	[94]
	<i>Nicotiana tabacum</i>	PsASGR-BBML	[97]
	<i>Oryza sativa</i>	PsASGR-BBML	[94]
	<i>Oryza sativa</i>	pDD45::SiBBM1-3	[98]
	<i>Oryza sativa</i>	pDD45::OsBBM	[96]
	<i>Ceratopteris richardii</i>	35S::BnBBM	[95]
	<i>Arabidopsis thaliana</i>	pDD45::BnBBM	[99]
	<i>Brassica Napus</i> <i>Solanum Lycopersicon</i>	pDD45::BnBBM pDD45::BnBBM	[99] [99]
PAR	<i>Lactuca sativa</i>	pEC1::ToPAR	[100]

## 5.2. Genome Editing

In addition to haploid breeding, a haploid induction system combined with CRISPR/Cas9 technology could be used for genome editing in crops that are resistant to genetic transformation. Gene-editing strategies, named haploid induction editing technology (HI-EDIT) and haploid inducer mediated genome editing (IMGE), have been developed based on in vivo haploid induction systems, respectively [79,108]. In the HI-EDIT system, CRISPR-Cas9 technology was combined with the MTL-mediated or CENH3-mediated haploid induction system to create a one-step genome editing method. Targeted genome loci could be efficiently edited in the haploid progenies and steadily inherited in the next generation, and at the same time, the Cas9 construct could be completely removed through genome elimination. This method is based on the post-fertilization genome elimination of the haploid induction systems. In maize, when pollinated by *mtl* pollen grains that carry the Cas9

construct, five of six maize germplasms received a more than 3% editing efficiency increase in the haploids via the HI-EDIT method. In *Arabidopsis*, a CENH3 haploid induction system, in combination with CRISPR-Cas9 technology, was developed for the HI-EDIT method. The results demonstrated that when *Arabidopsis Landsberg erecta* pollen grains pollinated to the *Arabidopsis cenH3-1* mutant (*Columbia* ecotype) expressing CENH3 from maize and the *Cas9* construct [109], 16.9% of the targeted genes in the haploid progenies were edited [79]. Thus, the HI-Edit method works effectively in both monocots and eudicots, which can provide a power tool for wide applications in genome-editing technology of the commercial variety. Similarly, Baobao Wang et al. [108] developed an approach, named IMGE, for genome editing. They introduced a CRISPR/Cas9 cassette into the CAU5 inducer line and tested its ability for genome editing in haploids, and they reported an approximately 4.1% increase in editing efficiency in the haploids. Hence, HI-EDIT and IMGE can be efficiently used for genome editing in crops that are resistant to genetic transformation, and their use avoids the interference of transgenes on crop traits through uniparental chromosome elimination.

### 5.3. Heterosis Fixation

Another application of haploid embryogenesis is for heterosis fixation, when used in combination with the *MiMe* (*Mitosis instead of Meiosis*) system [110–112]. In *MiMe* system, the combination of the mutations in the genes which are responsible for abolishing meiotic recombination, separating the sister chromatids and skipping the second division during meiosis, respectively, will shift meiosis into a mitosis-like division [113,114]. Since *MiMe* gametes are diploid, the introduction of a haploid induction system into the *MiMe* background can produce clonal propagation seeds with hybrid genotypes [96,102,115]. For example, *osd1–pair1–rec8–mtl* quadruple mutants (named *Fix*, for *Fixation of hybrids*) were used for heterosis fixation in rice [102]. Approximately 6.2% of the *Fix* diploid progenies with the same genotype as the mother were obtained, indicating that the *Fix* system can produce clonal seeds and fix heterozygosity in the F<sub>1</sub> generation of hybrid rice. Recent results have demonstrated that the heterotic phenotypes and synthetic apomixis traits of these clonal seeds could be stably transmitted to the next generation [116]. Similarly, the *MiMe* system, in combination with the ectopic expression of *BBM1* in egg cells, was also successfully used for heterosis fixation in rice, and clonal seeds were obtained from 11% and 29% of the progeny of the two transgenic lines of the *MiMe* plus *BBM1-ee* (ectopic expression of *OsBBM1* in egg cells) constructs [96], respectively. In further studies, more attention should be paid to increasing the efficiency of clonal seeds.

## 6. Conclusions and Perspectives

There are two alternative roads for generating haploid embryos, as described above. Microspore embryogenesis can not only be applied for haploid breeding, but it also is an ideal *in vitro* system for investigating cell fate determination and plant embryogenesis. Epigenetic reprogramming is critical for the initiation of microspore embryogenesis, which appears to be similar to somatic embryogenesis. It worth comparing the mechanisms behinds the different roads to embryogenesis, including zygotic embryogenesis, microspore embryogenesis, and somatic embryogenesis, in further studies. In addition, as the single-cell epigenome sequencing technologies develop, profiling DNA modifications and histone modifications during the initial stages of microspore embryogenesis will greatly aid us in understanding the mechanisms behind microspore reprogramming and haploid embryogenesis initiation.

The studies of *in vivo* haploid embryogenesis in flowering plants greatly promote the use of haploid technology in plant breeding, especially for maize. Based on the current *in vivo* haploid induction system discussed above, haploid embryogenesis may be induced by uniparental centromere defects, sperm chromosome fragmentation, and the ectopic expression of transcriptional factors related to embryo initiation in egg cells. These genes responsible for haploid induction are closely related to plant reproduction and, in particular, gamete development and fertilization. For example, *MTL*, *POD65*, *DMP*, and



*PLD3* are involved male gametophyte development, *ECS* is involved in egg cell development, and *BBM* and *PAR* are involved in embryogenesis initiation. In addition to haploid embryogenesis, sperm-central cell fertilization to form endosperm is necessary for haploid generation. Thus, basic studies related to the molecular and cellular mechanisms for gamete development, fertilization, and embryogenesis initiation will greatly aid us in establishing or optimizing haploid induction systems.

At present, there are still several questions that remain to be answered about these haploid induction systems. First, a high haploid induction rate is critical for their applications in plant breeding. Presently, *in vivo* haploid induction systems are typically accompanied by a high frequency of seed abortion. Reducing seed abortion and increasing haploid induction efficiency are key points for the improvement of *in vivo* haploid induction systems. Second, haploid induction systems have been established for only a few species, primarily for grass and Arabidopsis. Expanding these haploid induction systems to other important economic crops must be investigated with great care. Thus, more attention should be paid to the mechanisms of haploid embryogenesis, which will not only aid us in understanding haploid embryo formation, but it will also help us to optimize these haploid induction systems and expand the applications of haploid technology in crop breeding.

**Author Contributions:** Writing—original draft preparation, K.S., M.Q. and P.Z.; writing—review and editing, K.S. and P.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Major Project of Hubei Hongshan Laboratory, grant number 2022hszd017; the Science and Technology Department of Hubei Province, grant number 2022CFA071 and the National Natural Science Foundation of China, grant number 31970340.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Meng, D.; Liu, C.; Chen, S.; Jin, W. Haploid induction and its application in maize breeding. *Mol. Breed.* **2021**, *41*, 20. [[CrossRef](#)]
- Zhang, F.L.; Takahata, Y. Inheritance of microspore embryogenic ability in *Brassica crops*. *Theor. Appl. Genet.* **2001**, *103*, 254–258. [[CrossRef](#)]
- Satpute, G.K.; Long, H.; Seguí-Simarro, J.M.; Risueño, M.C.; Testillano, P.S. Cell architecture during gametophytic and embryogenic microspore development in *Brassica napus* L. *Acta Physiol. Plant.* **2005**, *27*, 665–674. [[CrossRef](#)]
- Shariatpanahi, M.E.; Bal, U.; Heberle-Bors, E.; Touraev, A. Stresses applied for the re-programming of plant microspores towards *in vitro* embryogenesis. *Physiol. Plant.* **2006**, *127*, 519–534. [[CrossRef](#)]
- Tang, X.C.; He, Y.Q.; Wang, Y.; Sun, M.X. The role of arabinogalactan proteins binding to Yariv reagents in the initiation, cell developmental fate, and maintenance of microspore embryogenesis in *Brassica napus* L. cv. Topas. *J. Exp. Bot.* **2006**, *57*, 2639–2650. [[CrossRef](#)]
- Germanà, M.A. Anther culture for haploid and doubled haploid production. *Plant Cell Tissue Organ Cult.* **2010**, *104*, 283–300. [[CrossRef](#)]
- Prem, D.; Solis, M.T.; Barany, I.; Rodriguez-Sanz, H.; Risueno, M.C.; Testillano, P.S. A new microspore embryogenesis system under low temperature which mimics zygotic embryogenesis initials, expresses auxin and efficiently regenerates doubled-haploid plants in *Brassica napus*. *BMC Plant Biol.* **2012**, *12*, 127. [[CrossRef](#)]
- Rivas-Sendra, A.; Campos-Vega, M.; Calabuig-Serna, A.; Seguí-Simarro, J.M. Development and characterization of an eggplant (*Solanum melongena*) doubled haploid population and a doubled haploid line with high androgenic response. *Euphytica* **2017**, *213*, 89. [[CrossRef](#)]
- Testillano, P.; Georgiev, S.; Mogensen, H.L.; Coronado, M.J.; Dumas, C.; Risueno, M.C.; Matthys-Rochon, E. Spontaneous chromosome doubling results from nuclear fusion during *in vitro* maize induced microspore embryogenesis. *Chromosoma* **2004**, *112*, 342–349. [[CrossRef](#)]
- Pintos, B.; Manzanera, J.A.; Bueno, M.A. Antimitotic agents increase the production of doubled-haploid embryos from cork oak anther culture. *J. Plant Physiol.* **2007**, *164*, 1595–1604. [[CrossRef](#)]
- Tang, X.; Liu, Y.; He, Y.; Ma, L.; Sun, M.X. Exine dehiscing induces rape microspore polarity, which results in different daughter cell fate and fixes the apical-basal axis of the embryo. *J. Exp. Bot.* **2013**, *64*, 215–228. [[CrossRef](#)] [[PubMed](#)]
- Testillano, P.S. Microspore embryogenesis: Targeting the determinant factors of stress-induced cell reprogramming for crop improvement. *J. Exp. Bot.* **2019**, *70*, 2965–2978. [[CrossRef](#)] [[PubMed](#)]
- Li, H.; Soriano, M.; Cordewener, J.; Muino, J.M.; Riksen, T.; Fukuoka, H.; Angenent, G.C.; Boutilier, K. The histone deacetylase inhibitor trichostatin A promotes totipotency in the male gametophyte. *Plant Cell* **2014**, *26*, 195–209. [[CrossRef](#)] [[PubMed](#)]

14. Zhang, L.; Zhang, Y.; Gao, Y.; Jiang, X.; Zhang, M.; Wu, H.; Liu, Z.; Feng, H. Effects of histone deacetylase inhibitors on microspore embryogenesis and plant regeneration in Pakchoi (*Brassica rapa* ssp. *chinensis* L.) *Sci. Hort.* **2016**, *209*, 61–66. [[CrossRef](#)]
15. Jiang, F.; Ryabova, D.; Diedhiou, J.; Hucl, P.; Randhawa, H.; Marillia, E.F.; Foroud, N.A.; Eudes, F.; Kathiria, P. Trichostatin A increases embryo and green plant regeneration in wheat. *Plant Cell Rep.* **2017**, *36*, 1701–1706. [[CrossRef](#)]
16. Castillo, A.M.; Valero-Rubira, I.; Burrell, M.A.; Allue, S.; Costar, M.A.; Valles, M.P. Trichostatin A Affects Developmental Reprogramming of Bread Wheat Microspores towards an Embryogenic Route. *Plants* **2020**, *9*, 1442. [[CrossRef](#)]
17. Pandey, P.; Daghma, D.S.; Houben, A.; Kumlehn, J.; Melzer, M.; Rutten, T. Dynamics of post-translationally modified histones during barley pollen embryogenesis in the presence or absence of the epi-drug trichostatin A. *Plant Reprod.* **2017**, *30*, 95–105. [[CrossRef](#)] [[PubMed](#)]
18. Nowicka, A.; Juzon, K.; Krzewska, M.; Dziurka, M.; Dubas, E.; Kopec, P.; Zielinski, K.; Zur, I. Chemically-induced DNA de-methylation alters the effectiveness of microspore embryogenesis in triticale. *Plant Sci.* **2019**, *287*, 110189. [[CrossRef](#)]
19. Berenguer, E.; Barany, I.; Solis, M.T.; Perez-Perez, Y.; Risueno, M.C.; Testillano, P.S. Inhibition of Histone H3K9 Methylation by BIX-01294 Promotes Stress-Induced Microspore Totipotency and Enhances Embryogenesis Initiation. *Front. Plant Sci.* **2017**, *8*, 1161. [[CrossRef](#)]
20. Solis, M.T.; Rodriguez-Serrano, M.; Meijon, M.; Canal, M.J.; Cifuentes, A.; Risueno, M.C.; Testillano, P.S. DNA methylation dynamics and *MET1a-like* gene expression changes during stress-induced pollen reprogramming to embryogenesis. *J. Exp. Bot.* **2012**, *63*, 6431–6444. [[CrossRef](#)]
21. Dwivedi, S.L.; Britt, A.B.; Tripathi, L.; Sharma, S.; Upadhyaya, H.D.; Ortiz, R. Haploids: Constraints and opportunities in plant breeding. *Biotechnol. Adv.* **2015**, *33*, 812–829. [[CrossRef](#)] [[PubMed](#)]
22. Hale, B.; Ferrie, A.M.R.; Chellamma, S.; Samuel, J.P.; Phillips, G.C. Androgenesis-Based Doubled Haploidy: Past, Present, and Future Perspectives. *Front. Plant Sci.* **2021**, *12*, 751230. [[CrossRef](#)] [[PubMed](#)]
23. Zhao, P.; Zhou, X.M.; Zhao, L.L.; Cheung, A.Y.; Sun, M.X. Autophagy-mediated compartmental cytoplasmic deletion is essential for tobacco pollen germination and male fertility. *Autophagy* **2020**, *16*, 2180–2192. [[CrossRef](#)] [[PubMed](#)]
24. Yan, H.; Zhuang, M.; Xu, X.; Li, S.; Yang, M.; Li, N.; Du, X.; Hu, K.; Peng, X.; Huang, W.; et al. Autophagy and its mediated mitochondrial quality control maintain pollen tube growth and male fertility in Arabidopsis. *Autophagy* **2022**, 1–16. [[CrossRef](#)]
25. Corral-Martínez, P.; Parra-Vega, V.; Seguí-Simarro, J.M. Novel features of *Brassica napus* embryogenic microspores revealed by high pressure freezing and freeze substitution: Evidence for massive autophagy and excretion-based cytoplasmic cleaning. *J. Exp. Bot.* **2013**, *64*, 3061–3075. [[CrossRef](#)]
26. Barany, I.; Berenguer, E.; Solis, M.T.; Perez-Perez, Y.; Santamaria, M.E.; Crespo, J.L.; Risueno, M.C.; Diaz, I.; Testillano, P.S. Autophagy is activated and involved in cell death with participation of cathepsins during stress-induced microspore embryogenesis in barley. *J. Exp. Bot.* **2018**, *69*, 1387–1402. [[CrossRef](#)] [[PubMed](#)]
27. Berenguer, E.; Minina, E.A.; Carneros, E.; Bárány, I.; Bozhkov, P.V.; Testillano, P.S. Suppression of Metacaspase- and Autophagy-Dependent Cell Death Improves Stress-Induced Microspore Embryogenesis in *Brassica napus*. *Plant Cell Physiol.* **2021**, *61*, 2097–2110. [[CrossRef](#)]
28. Rodriguez-Serrano, M.; Barany, I.; Prem, D.; Coronado, M.J.; Risueno, M.C.; Testillano, P.S. NO, ROS, and cell death associated with caspase-like activity increase in stress-induced microspore embryogenesis of barley. *J. Exp. Bot.* **2012**, *63*, 2007–2024. [[CrossRef](#)]
29. Zur, I.; Dubas, E.; Golemić, E.; Szechynska-Hebda, M.; Golebiowska, G.; Wedzony, M. Stress-related variation in antioxidative enzymes activity and cell metabolism efficiency associated with embryogenesis induction in isolated microspore culture of triticale (*x Triticosecale* Wittm.). *Plant Cell Rep.* **2009**, *28*, 1279–1287. [[CrossRef](#)]
30. Wieland, G.; Orthaus, S.; Ohndorf, S.; Diekmann, S.; Hemmerich, P. Functional complementation of human centromere protein A (CENP-A) by Cse4p from *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **2004**, *24*, 6620–6630. [[CrossRef](#)]
31. Talbert, P.B.; Masuelli, R.; Tyagi, A.P.; Comai, L.; Henikoff, S. Centromeric localization and adaptive evolution of an *Arabidopsis* histone H3 variant. *Plant Cell* **2002**, *14*, 1053–1066. [[CrossRef](#)] [[PubMed](#)]
32. Lermontova, I.; Koroleva, O.; Rutten, T.; Fuchs, J.; Schubert, V.; Moraes, I.; Koszegi, D.; Schubert, I. Knockdown of CENH3 in *Arabidopsis* reduces mitotic divisions and causes sterility by disturbed meiotic chromosome segregation. *Plant J.* **2011**, *68*, 40–50. [[CrossRef](#)] [[PubMed](#)]
33. Ravi, M.; Shibata, F.; Ramahi, J.S.; Nagaki, K.; Chen, C.; Murata, M.; Chan, S.W. Meiosis-specific loading of the centromere-specific histone CENH3 in *Arabidopsis thaliana*. *PLoS Genet.* **2011**, *7*, e1002121. [[CrossRef](#)] [[PubMed](#)]
34. Sanei, M.; Pickering, R.; Kumke, K.; Nasuda, S.; Houben, A. Loss of centromeric histone H3 (CENH3) from centromeres precedes uniparental chromosome elimination in interspecific barley hybrids. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, E498–E505. [[CrossRef](#)] [[PubMed](#)]
35. Fukagawa, T.; Earnshaw, W.C. The centromere: Chromatin foundation for the kinetochore machinery. *Dev. Cell* **2014**, *30*, 496–508. [[CrossRef](#)]
36. Ishii, T.; Karimi-Ashtiyani, R.; Houben, A. Haploidization via Chromosome Elimination: Means and Mechanisms. *Annu. Rev. Plant Biol.* **2016**, *67*, 421–438. [[CrossRef](#)]
37. Comai, L.; Maheshwari, S.; Marimuthu, M.P.A. Plant centromeres. *Curr. Opin. Plant Biol.* **2017**, *36*, 158–167. [[CrossRef](#)]
38. Wang, N.; Dawe, R.K. Centromere Size and Its Relationship to Haploid Formation in Plants. *Mol. Plant* **2018**, *11*, 398–406. [[CrossRef](#)]

39. Kuppuppu, S.; Ron, M.; Marimuthu, M.P.A.; Li, G.; Huddleson, A.; Siddeek, M.H.; Terry, J.; Buchner, R.; Shabek, N.; Comai, L.; et al. A variety of changes, including CRISPR/Cas9-mediated deletions, in CENH3 lead to haploid induction on outcrossing. *Plant Biotechnol. J.* **2020**, *18*, 2068–2080. [[CrossRef](#)]
40. Ravi, M.; Kwong, P.N.; Menorca, R.M.; Valencia, J.T.; Ramahi, J.S.; Stewart, J.L.; Tran, R.K.; Sundaresan, V.; Comai, L.; Chan, S.W. The rapidly evolving centromere-specific histone has stringent functional requirements in *Arabidopsis thaliana*. *Genetics* **2010**, *186*, 461–471. [[CrossRef](#)]
41. Maheshwari, S.; Tan, E.H.; West, A.; Franklin, F.C.; Comai, L.; Chan, S.W. Naturally occurring differences in CENH3 affect chromosome segregation in zygotic mitosis of hybrids. *PLoS Genet.* **2015**, *11*, e1004970. [[CrossRef](#)] [[PubMed](#)]
42. Karimi-Ashtiyani, R.; Ishii, T.; Niessen, M.; Stein, N.; Heckmann, S.; Gurushidze, M.; Banaei-Moghaddam, A.M.; Fuchs, J.; Schubert, V.; Koch, K.; et al. Point mutation impairs centromeric CENH3 loading and induces haploid plants. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 11211–11216. [[CrossRef](#)] [[PubMed](#)]
43. Kuppuppu, S.; Tan, E.H.; Nguyen, H.; Rodgers, A.; Comai, L.; Chan, S.W.; Britt, A.B. Point Mutations in Centromeric Histone Induce Post-zygotic Incompatibility and Uniparental Inheritance. *PLoS Genet.* **2015**, *11*, e1005494. [[CrossRef](#)]
44. Britt, A.B.; Kuppuppu, S. CenH3: An Emerging Player in Haploid Induction Technology. *Front. Plant Sci.* **2016**, *7*, 357. [[CrossRef](#)]
45. Watts, A.; Kumar, V.; Bhat, S.R. Centromeric histone H3 protein: From basic study to plant breeding applications. *J. Plant Biochem. Biotechnol.* **2016**, *25*, 339–348. [[CrossRef](#)]
46. Ren, J.; Wu, P.; Trampe, B.; Tian, X.; Lubberstedt, T.; Chen, S. Novel technologies in doubled haploid line development. *Plant Biotechnol. J.* **2017**, *15*, 1361–1370. [[CrossRef](#)] [[PubMed](#)]
47. Kursel, L.E.; Malik, H.S. The cellular mechanisms and consequences of centromere drive. *Curr. Opin. Cell Biol.* **2018**, *52*, 58–65. [[CrossRef](#)]
48. Comai, L.; Tan, E.H. Haploid Induction and Genome Instability. *Trends Genet.* **2019**, *35*, 791–803. [[CrossRef](#)]
49. Kalinowska, K.; Chamas, S.; Unkel, K.; Demidov, D.; Lermontova, I.; Dresselhaus, T.; Kumlehn, J.; Dunemann, F.; Houben, A. State-of-the-art and novel developments of in vivo haploid technologies. *Theor. Appl. Genet.* **2019**, *132*, 593–605. [[CrossRef](#)]
50. Wang, S.; Jin, W.; Wang, K. Centromere histone H3- and phospholipase-mediated haploid induction in plants. *Plant Methods* **2019**, *15*, 42. [[CrossRef](#)]
51. Karimi-Ashtiyani, R. Centromere engineering as an emerging tool for haploid plant production: Advances and challenges. *Doubled Haploid Technol.* **2021**, *2289*, 3–22.
52. Zhou, J.; Liu, Y.; Guo, X.; Birchler, J.A.; Han, F.; Su, H. Centromeres: From chromosome biology to biotechnology applications and synthetic genomes in plants. *Plant Biotechnol. J.* **2022**, *20*, 2051–2063. [[CrossRef](#)] [[PubMed](#)]
53. Ravi, M.; Chan, S.W. Haploid plants produced by centromere-mediated genome elimination. *Nature* **2010**, *464*, 615–618. [[CrossRef](#)] [[PubMed](#)]
54. Marimuthu, M.P.A.; Maruthachalam, R.; Bondada, R.; Kuppuppu, S.; Tan, E.H.; Britt, A.; Chan, S.W.L.; Comai, L. Epigenetically mismatched parental centromeres trigger genome elimination in hybrids. *Sci. Adv.* **2021**, *7*, eabk1151. [[CrossRef](#)]
55. Wang, N.; Gent, J.I.; Dawe, R.K. Haploid induction by a maize cenH3 null mutant. *Sci. Adv.* **2021**, *7*, eabe2299. [[CrossRef](#)]
56. Crasta, K.; Ganem, N.J.; Dagher, R.; Lantermann, A.B.; Ivanova, E.V.; Pan, Y.; Nezi, L.; Protopopov, A.; Chowdhury, D.; Pellman, D. DNA breaks and chromosome pulverization from errors in mitosis. *Nature* **2012**, *482*, 53–58. [[CrossRef](#)]
57. Coe, E.H. A Line of Maize with High Haploid Frequency. *Am. Nat.* **1959**, *93*, 381–382. [[CrossRef](#)]
58. Kelliher, T.; Starr, D.; Richbourg, L.; Chintamanani, S.; Delzer, B.; Nuccio, M.L.; Green, J.; Chen, Z.; McCuiston, J.; Wang, W.; et al. MATRILINEAL, a sperm-specific phospholipase, triggers maize haploid induction. *Nature* **2017**, *542*, 105–109. [[CrossRef](#)]
59. Liu, C.; Li, X.; Meng, D.; Zhong, Y.; Chen, C.; Dong, X.; Xu, X.; Chen, B.; Li, W.; Li, L.; et al. A 4-bp Insertion at ZmPLA1 Encoding a Putative Phospholipase A Generates Haploid Induction in Maize. *Mol. Plant* **2017**, *10*, 520–522. [[CrossRef](#)]
60. Gilles, L.M.; Khaled, A.; Laffaire, J.B.; Chaignon, S.; Gendrot, G.; Laplaige, J.; Berges, H.; Beydon, G.; Bayle, V.; Barret, P.; et al. Loss of pollen-specific phospholipase NOT LIKE DAD triggers gynogenesis in maize. *EMBO J.* **2017**, *36*, 707–717. [[CrossRef](#)]
61. Wang, G.; Ryu, S.; Wang, X. Plant Phospholipases: An Overview. *Methods Mol. Biol.* **2012**, *861*, 123–137. [[PubMed](#)]
62. Aloulou, A.; Rahier, R.; Arhab, Y.; Noiriél, A.; Abousalham, A. Phospholipases: An overview. *Lipases Phospholipases* **2018**, *1835*, 69–105.
63. Ali, U.; Lu, S.; Fadlalla, T.; Iqbal, S.; Yue, H.; Yang, B.; Hong, Y.; Wang, X.; Guo, L. The functions of phospholipases and their hydrolysis products in plant growth, development and stress responses. *Prog. Lipid Res.* **2022**, *86*, 101158. [[CrossRef](#)] [[PubMed](#)]
64. Gilles, L.M.; Calhau, A.R.M.; La Padula, V.; Jacquier, N.M.A.; Lionnet, C.; Martinant, J.P.; Rogowsky, P.M.; Widiez, T. Lipid anchoring and electrostatic interactions target NOT-LIKE-DAD to pollen endo-plasma membrane. *J. Cell Biol.* **2021**, *220*, e202010077. [[CrossRef](#)] [[PubMed](#)]
65. Zhong, Y.; Liu, C.; Qi, X.; Jiao, Y.; Wang, D.; Wang, Y.; Liu, Z.; Chen, C.; Chen, B.; Tian, X.; et al. Mutation of ZmDMP enhances haploid induction in maize. *Nat. Plants* **2019**, *5*, 575–580. [[CrossRef](#)]
66. Li, Y.; Li, D.; Xiao, Q.; Wang, H.; Wen, J.; Tu, J.; Shen, J.; Fu, T.; Yi, B. An in planta haploid induction system in *Brassica napus*. *J. Integr. Plant Biol.* **2022**, *64*, 1140–1144. [[CrossRef](#)]
67. Zhong, Y.; Wang, Y.; Chen, B.; Liu, J.; Wang, D.; Li, M.; Qi, X.; Liu, C.; Boutilier, K.; Chen, S. Establishment of a dmp based maternal haploid induction system for polyploid *Brassica napus* and *Nicotiana tabacum*. *J. Integr. Plant Biol.* **2022**, *64*, 1281–1294. [[CrossRef](#)]
68. Zhao, X.; Yuan, K.; Liu, Y.; Zhang, N.; Yang, L.; Zhang, Y.; Wang, Y.; Ji, J.; Fang, Z.; Han, F.; et al. In vivo maternal haploid induction based on genome editing of DMP in *Brassica oleracea*. *Plant Biotechnol. J.* **2022**, *20*, 2242–2244. [[CrossRef](#)]



69. Wang, N.; Xia, X.; Jiang, T.; Li, L.; Zhang, P.; Niu, L.; Cheng, H.; Wang, K.; Lin, H. In planta haploid induction by genome editing of DMP in the model legume *Medicago truncatula*. *Plant Biotechnol. J.* **2022**, *20*, 22–24. [[CrossRef](#)]
70. Zhang, X.; Zhang, L.; Zhang, J.; Jia, M.; Cao, L.; Yu, J.; Zhao, D. Haploid induction in allotetraploid tobacco using *DMPs* mutation. *Planta* **2022**, *255*, 98. [[CrossRef](#)]
71. Zhong, Y.; Chen, B.; Wang, D.; Zhu, X.; Li, M.; Zhang, J.; Chen, M.; Wang, M.; Riksen, T.; Liu, J.; et al. In vivo maternal haploid induction in tomato. *Plant Biotechnol. J.* **2022**, *20*, 250–252. [[CrossRef](#)] [[PubMed](#)]
72. Zhang, J.; Yin, J.; Luo, J.; Tang, D.; Zhu, X.; Wang, J.; Liu, Z.; Wang, P.; Zhong, Y.; Liu, C.; et al. Construction of homozygous diploid potato through maternal haploid induction. *aBIOTECH* **2022**, *3*, 163–168. [[CrossRef](#)] [[PubMed](#)]
73. Li, Y.; Lin, Z.; Yue, Y.; Zhao, H.; Fei, X.; Li, E.; Liu, C.; Chen, S.; Lai, J.; Song, W. Loss-of-function alleles of *ZmPLD3* cause haploid induction in maize. *Nat. Plants* **2021**, *7*, 1579–1588. [[CrossRef](#)]
74. Jiang, C.; Sun, J.; Li, R.; Yan, S.; Chen, W.; Guo, L.; Qin, G.; Wang, P.; Luo, C.; Huang, W.; et al. A reactive oxygen species burst causes haploid induction in maize. *Mol. Plant* **2022**, *15*, 943–955. [[CrossRef](#)]
75. Yu, X.; Zhang, X.; Zhao, P.; Peng, X.; Chen, H.; Bleckmann, A.; Bazhenova, A.; Shi, C.; Dresselhaus, T.; Sun, M.X. Fertilized egg cells secrete endopeptidases to avoid polytubey. *Nature* **2021**, *592*, 433–437. [[CrossRef](#)]
76. Mao, Y.; Nakel, T.; Serbes, I.E.; Tekleyohans, D.G.; Joshi, S.; Baum, T.; Groß-Hardt, R. ECS1 and ECS2 regulate polyspermy and suppress the formation of haploid plants by promoting double fertilization. *BioRxiv* **2022**. [[CrossRef](#)]
77. Zhang, X.; Shi, C.; Li, S.; Zhang, B.; Dresselhaus, T.; Sun, M.-x. A female in vivo haploid-induction system via mutagenesis of egg cell-specific peptidases. *Res. Sq.* **2022**. [[CrossRef](#)]
78. Zhao, X.; Xu, X.; Xie, H.; Chen, S.; Jin, W. Fertilization and uniparental chromosome elimination during crosses with maize haploid inducers. *Plant Physiol.* **2013**, *163*, 721–731. [[CrossRef](#)] [[PubMed](#)]
79. Kelliher, T.; Starr, D.; Su, X.; Tang, G.; Chen, Z.; Carter, J.; Wittich, P.E.; Dong, S.; Green, J.; Burch, E.; et al. One-step genome editing of elite crop germplasm during haploid induction. *Nat. Biotechnol.* **2019**, *37*, 287–292. [[CrossRef](#)] [[PubMed](#)]
80. Sun, G.; Geng, S.; Zhang, H.; Jia, M.; Wang, Z.; Deng, Z.; Tao, S.; Liao, R.; Wang, F.; Kong, X.; et al. Matrilineal empowers wheat pollen with haploid induction potency by triggering postmitosis reactive oxygen species activity. *New Phytol.* **2022**, *233*, 2405–2414. [[CrossRef](#)]
81. Li, X.; Meng, D.; Chen, S.; Luo, H.; Zhang, Q.; Jin, W.; Yan, J. Single nucleus sequencing reveals spermatid chromosome fragmentation as a possible cause of maize haploid induction. *Nat. Commun.* **2017**, *8*, 991. [[CrossRef](#)] [[PubMed](#)]
82. Cosio, C.; Dunand, C. Specific functions of individual class III peroxidase genes. *J. Exp. Bot.* **2009**, *60*, 391–408. [[CrossRef](#)] [[PubMed](#)]
83. Yoshida, K.; Kaothien, P.; Matsui, T.; Kawaoka, A.; Shinmyo, A. Molecular biology and application of plant peroxidase genes. *Appl. Microbiol. Biotechnol.* **2003**, *60*, 665–670. [[CrossRef](#)]
84. Ruban, A.; Houben, A. Highly reactive chemicals meet haploidization. *Mol. Plant* **2022**, *15*, 937–939. [[CrossRef](#)] [[PubMed](#)]
85. Boutilier, K.; Offringa, R.; Sharma, V.K.; Kieft, H.; Ouellet, T.; Zhang, L.; Hattori, J.; Liu, C.M.; van Lammeren, A.A.; Miki, B.L.; et al. Ectopic expression of *BABY BOOM* triggers a conversion from vegetative to embryonic growth. *Plant Cell* **2002**, *14*, 1737–1749. [[CrossRef](#)] [[PubMed](#)]
86. El Ouakfaoui, S.; Schnell, J.; Abdeen, A.; Colville, A.; Labbe, H.; Han, S.; Baum, B.; Laberge, S.; Miki, B. Control of somatic embryogenesis and embryo development by AP2 transcription factors. *Plant Mol. Biol.* **2010**, *74*, 313–326. [[CrossRef](#)] [[PubMed](#)]
87. Irikova, T.; Grozeva, S.; Denev, I. Identification of *BABY BOOM* and *LEAFY COTYLEDON* genes in sweet pepper (*Capsicum annuum* L.) genome by their partial gene sequences. *Plant Growth Regul.* **2012**, *67*, 191–198. [[CrossRef](#)]
88. Salvo, S.A.; Hirsch, C.N.; Buell, C.R.; Kaeppler, S.M.; Kaeppler, H.F. Whole transcriptome profiling of maize during early somatic embryogenesis reveals altered expression of stress factors and embryogenesis-related genes. *PLoS ONE* **2014**, *9*, e111407. [[CrossRef](#)]
89. Florez, S.L.; Erwin, R.L.; Maximova, S.N.; Guiltinan, M.J.; Curtis, W.R. Enhanced somatic embryogenesis in *Theobroma cacao* using the homologous *BABY BOOM* transcription factor. *BMC Plant Biol.* **2015**, *15*, 121. [[CrossRef](#)]
90. Lutz, K.A.; Martin, C.; Khairzada, S.; Maliga, P. Steroid-inducible *BABY BOOM* system for development of fertile *Arabidopsis thaliana* plants after prolonged tissue culture. *Plant Cell Rep.* **2015**, *34*, 1849–1856. [[CrossRef](#)]
91. Horstman, A.; Li, M.; Heidmann, I.; Weemen, M.; Chen, B.; Muino, J.M.; Angenent, G.C.; Boutilier, K. The *BABY BOOM* Transcription Factor Activates the LEC1-ABI3-FUS3-LEC2 Network to Induce Somatic Embryogenesis. *Plant Physiol.* **2017**, *175*, 848–857. [[CrossRef](#)] [[PubMed](#)]
92. Jha, P.; Kumar, V. *BABY BOOM* (BBM): A candidate transcription factor gene in plant biotechnology. *Biotechnol. Lett.* **2018**, *40*, 1467–1475. [[CrossRef](#)] [[PubMed](#)]
93. Conner, J.A.; Mookkan, M.; Huo, H.; Chae, K.; Ozias-Akins, P. A parthenogenesis gene of apomict origin elicits embryo formation from unfertilized eggs in a sexual plant. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 11205–11210. [[CrossRef](#)] [[PubMed](#)]
94. Conner, J.A.; Podio, M.; Ozias-Akins, P. Haploid embryo production in rice and maize induced by *PsASGR-BBML* transgenes. *Plant Reprod.* **2017**, *30*, 41–52. [[CrossRef](#)]
95. Bui, L.T.; Pandzic, D.; Youngstrom, C.E.; Wallace, S.; Irish, E.E.; Szovenyi, P.; Cheng, C.L. A fern *AINTEGUMENTA* gene mirrors *BABY BOOM* in promoting apogamy in *Ceratopteris richardii*. *Plant J.* **2017**, *90*, 122–132. [[CrossRef](#)]
96. Khanday, I.; Skinner, D.; Yang, B.; Mercier, R.; Sundaresan, V. A male-expressed rice embryogenic trigger redirected for asexual propagation through seeds. *Nature* **2019**, *565*, 91–95. [[CrossRef](#)]




97. Zhang, Z.; Conner, J.; Guo, Y.; Ozias-Akins, P. Haploidy in Tobacco Induced by *PsASGR-BBML* Transgenes via Parthenogenesis. *Genes* **2020**, *11*, 1072. [[CrossRef](#)]
98. Chahal, L.S.; Conner, J.A.; Ozias-Akins, P. Phylogenetically Distant *BABY BOOM* Genes from *Setaria italica* Induce Parthenogenesis in Rice. *Front. Plant Sci.* **2022**, *13*, 863908. [[CrossRef](#)]
99. Chen, B.; Maas, L.; Figueiredo, D.; Zhong, Y.; Reis, R.; Li, M.; Horstman, A.; Riksen, T.; Weemen, M.; Liu, H.; et al. *BABY BOOM* regulates early embryo and endosperm development. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2201761119. [[CrossRef](#)]
100. Underwood, C.J.; Vijverberg, K.; Rigola, D.; Okamoto, S.; Oplaat, C.; Camp, R.; Radoeva, T.; Schauer, S.E.; Fierens, J.; Jansen, K.; et al. A *PARTHENOGENESIS* allele from apomictic dandelion can induce egg cell division without fertilization in lettuce. *Nat. Genet.* **2022**, *54*, 84–93. [[CrossRef](#)]
101. Yao, L.; Zhang, Y.; Liu, C.; Liu, Y.; Wang, Y.; Liang, D.; Liu, J.; Sahoo, G.; Kelliher, T. *OsMATL* mutation induces haploid seed formation in indica rice. *Nat. Plants* **2018**, *4*, 530–533. [[CrossRef](#)] [[PubMed](#)]
102. Wang, C.; Liu, Q.; Shen, Y.; Hua, Y.; Wang, J.; Lin, J.; Wu, M.; Sun, T.; Cheng, Z.; Mercier, R.; et al. Clonal seeds from hybrid rice by simultaneous genome engineering of meiosis and fertilization genes. *Nat. Biotechnol.* **2019**, *37*, 283–286. [[CrossRef](#)] [[PubMed](#)]
103. Liu, C.; Zhong, Y.; Qi, X.; Chen, M.; Liu, Z.; Chen, C.; Tian, X.; Li, J.; Jiao, Y.; Wang, D.; et al. Extension of the in vivo haploid induction system from diploid maize to hexaploid wheat. *Plant Biotechnol. J.* **2020**, *18*, 316–318. [[CrossRef](#)] [[PubMed](#)]
104. Cheng, Z.; Sun, Y.; Yang, S.; Zhi, H.; Yin, T.; Ma, X.; Zhang, H.; Diao, X.; Guo, Y.; Li, X.; et al. Establishing in planta haploid inducer line by edited *SiMTL* in foxtail millet (*Setaria italica*). *Plant Biotechnol. J.* **2021**, *19*, 1089–1091. [[CrossRef](#)]
105. Zhong, Y.; Chen, B.; Li, M.; Wang, D.; Jiao, Y.; Qi, X.; Wang, M.; Liu, Z.; Chen, C.; Wang, Y.; et al. A DMP-triggered in vivo maternal haploid induction system in the dicotyledonous *Arabidopsis*. *Nat. Plants* **2020**, *6*, 466–472. [[CrossRef](#)]
106. Ravi, M.; Marimuthu, M.P.; Tan, E.H.; Maheshwari, S.; Henry, I.M.; Marin-Rodriguez, B.; Urtecho, G.; Tan, J.; Thornhill, K.; Zhu, F.; et al. A haploid genetics toolbox for *Arabidopsis thaliana*. *Nat. Commun.* **2014**, *5*, 5334. [[CrossRef](#)] [[PubMed](#)]
107. Kelliher, T.; Starr, D.; Wang, W.; McCuiston, J.; Zhong, H.; Nuccio, M.L.; Martin, B. Maternal Haploids Are Preferentially Induced by *CENH3-tailswap* Transgenic Complementation in Maize. *Front. Plant Sci.* **2016**, *7*, 414. [[CrossRef](#)] [[PubMed](#)]
108. Wang, B.; Zhu, L.; Zhao, B.; Zhao, Y.; Xie, Y.; Zheng, Z.; Li, Y.; Sun, J.; Wang, H. Development of a Haploid-Inducer Mediated Genome Editing System for Accelerating Maize Breeding. *Mol. Plant* **2019**, *12*, 597–602. [[CrossRef](#)]
109. Feng, C.; Yuan, J.; Bai, H.; Liu, Y.; Su, H.; Liu, Y.; Shi, L.; Gao, Z.; Birchler, J.A.; Han, F. The deposition of *CENH3* in maize is stringently regulated. *Plant J.* **2020**, *102*, 6–17. [[CrossRef](#)]
110. Jacquier, N.M.A.; Gilles, L.M.; Pyott, D.E.; Martinant, J.P.; Rogowsky, P.M.; Widiez, T. Puzzling out plant reproduction by haploid induction for innovations in plant breeding. *Nat. Plants* **2020**, *6*, 610–619. [[CrossRef](#)]
111. Ozias-Akins, P.; Conner, J.A. Clonal Reproduction through Seeds in Sight for Crops. *Trends Genet.* **2020**, *36*, 215–226. [[CrossRef](#)] [[PubMed](#)]
112. Underwood, C.J.; Mercier, R. Engineering Apomixis: Clonal Seeds Approaching the Fields. *Annu. Rev. Plant Biol.* **2022**, *73*, 201–225. [[CrossRef](#)] [[PubMed](#)]
113. d’Erfurth, I.; Jolivet, S.; Froger, N.; Catrice, O.; Novatchkova, M.; Mercier, R. Turning meiosis into mitosis. *PLoS Biol.* **2009**, *7*, e1000124. [[CrossRef](#)] [[PubMed](#)]
114. Mieulet, D.; Jolivet, S.; Rivard, M.; Cromer, L.; Vernet, A.; Mayonove, P.; Pereira, L.; Droc, G.; Courtois, B.; Guiderdoni, E.; et al. Turning rice meiosis into mitosis. *Cell Res.* **2016**, *26*, 1242–1254. [[CrossRef](#)] [[PubMed](#)]
115. Marimuthu, M.P.; Jolivet, S.; Ravi, M.; Pereira, L.; Davda, J.N.; Cromer, L.; Wang, L.; Nogue, F.; Chan, S.W.; Siddiqi, I.; et al. Synthetic clonal reproduction through seeds. *Science* **2011**, *331*, 876. [[CrossRef](#)]
116. Liu, C.; He, Z.; Zhang, Y.; Hu, F.; Li, M.; Liu, Q.; Huang, Y.; Wang, J.; Zhang, W.; Wang, C.; et al. Synthetic apomixis enables stable transgenerational transmission of heterotic phenotypes in hybrid rice. *Plant Commun.* **2022**, *4*, 100470. [[CrossRef](#)] [[PubMed](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

## CURRENT REVIEW

# Unveiling the Diversity: Plant Parasitic Nematode Effectors and Their Plant Interaction Partners

Sapinder Bali and Cynthia Gleason<sup>†</sup> 

Department of Plant Pathology, Washington State University, Pullman, WA 99164, U.S.A.

Accepted for publication 20 October 2023.

**Root-knot and cyst nematodes are two groups of plant parasitic nematodes that cause the majority of crop losses in agriculture. As a result, these nematodes are the focus of most nematode effector research. Root-knot and cyst nematode effectors are defined as secreted molecules, typically proteins, with crucial roles in nematode parasitism. There are likely hundreds of secreted effector molecules exuded through the nematode stylet into the plant. The current research has shown that nematode effectors can target a variety of host proteins and have impacts that include the suppression of plant immune responses and the manipulation of host hormone signaling. The discovery of effectors that localize to the nucleus indicates that the nematodes can directly modulate host gene expression for cellular reprogramming during feeding site formation. In addition, plant peptide mimicry by some nematode effectors highlights the sophisticated strategies the nematodes employ to manipulate host processes. Here we describe research on the interactions between nematode effectors and host proteins that will provide insights into the molecular mechanisms underpinning plant–nematode interactions. By identifying the host proteins and pathways that are targeted by root-knot and cyst nematode effectors, scientists can gain a better understanding of how nematodes establish feeding sites and subvert plant immune responses. Such information will be invaluable for future engineering of nematode-resistant crops, ultimately fostering advancements in agricultural practices and crop protection.**

**Keywords:** cyst nematodes, effectors, *Globodera*, *Heterodera*, immunity, *Meloidogyne*, resistance, root-knot nematodes

Plant parasitic nematodes cause considerable losses in agriculture. This review will focus on two major groups of sedentary endoparasitic nematodes: root-knot nematodes (RKNs, *Meloidogyne* spp.) and cyst nematodes (CNs, *Globodera* spp. and *Heterodera* spp.). Both RKNs and CNs have generally similar life cycles, in that they will hatch from the

egg stage in the soil as a migratory, nonfeeding second stage juvenile (J2). The J2s will freely move in the soil until they penetrate host roots, migrate to the vasculature, and establish their feeding sites. The nematodes possess a hollow, needle-like structure called the stylet, which they use for feeding from their feeding sites. Both RKNs and CNs are stylet-bearing nematodes in Clade 12, as based on small subunit rDNA sequences (Bird et al. 2015; Quist et al. 2015). Despite having a common bacterivore ancestor millions of years ago, they are currently placed in different families within this clade. One of the obvious differences between RKNs and CNs is in their feeding sites. RKNs will choose a cluster of five to eight plant cells that grow into large, multinucleated cells called “giant cells” (Caillaud et al. 2008; Perry et al. 2009). CNs craft a feeding site referred to as a syncytium, formed through the fusion of a solitary plant cell with adjacent cells, followed by endoreduplication (de Almeida Engler and Gheysen 2013; Grundler et al. 1998). Because RKNs and CNs remain sedentary upon establishing their feeding sites, these sites are crucial to their survival in the plant. Both RKNs and CNs will undergo three additional molts to reach the adult life stage. In the case of RKNs, the females lay eggs on the outside of the root surface in a gelatinous matrix called an egg mass. CNs retain eggs inside the body of the female. When the CN female dies, the cuticle undergoes melanization and fortification, culminating in the formation of protective egg-containing cysts. In general, RKN- and CN-infected plants may present symptoms like stunting, wilting, and reduced growth, and overall, infections can lead to significant crop yield losses unless they are controlled (FAO 2019).

Plant parasitic nematode effectors are secreted molecules that promote parasitism by suppressing plant immune responses and/or facilitating the formation and maintenance of the feeding sites (Molloy et al. 2023; Rosso et al. 2011; Vieira and Gleason 2019). Nematodes can secrete effectors from their amphids and/or cuticle (Jones and Robertson 1997). For example, the RKN effector MiMIF2 is secreted by the *Meloidogyne incognita* hypodermis to the cuticle. This effector is translocated into the host cells, where it interacts with the host annexins AnnAt1 and AnnAt4 (Zhao et al. 2019). Because annexins are involved in Ca<sup>2+</sup> signaling (Davies 2014), MiMIF2 might be interacting with annexins to modulate calcium signaling and interfere with host immune responses (Zhao et al. 2019). Although MiMIF2 is an interesting example of an effector secreted from the cuticle, most characterized effectors are proteins that are synthesized in the subventral or dorsal glands and then secreted via the stylet during the root invasion and infection (Vieira and Gleason 2019).

There have been many reviews published that thoroughly discuss effectors in nematode parasitism (Ali et al. 2017; Evesvan den Akker 2021; Mejias et al. 2019; Perrine-Walker 2019; Rutter et al. 2022; Sato et al. 2019; Vieira and Gleason 2019).

<sup>†</sup>Corresponding author: C. Gleason; [cynthia.gleason@wsu.edu](mailto:cynthia.gleason@wsu.edu)

**Funding:** Funding was provided by the National Institute of Food and Agriculture (2019-67013-29350), the Northwest Potato Research Consortium, and the U.S. Department of Agriculture Multistate Research Project W4186.

The author(s) declare no conflict of interest.



The author(s) have dedicated the work to the public domain under the Creative Commons CC0 “No Rights Reserved” license by waiving all of his or her rights to the work worldwide under copyright law, including all related and neighboring rights, to the extent allowed by law, 2024.

Although several nematode effectors have been reported so far, only a few of their host target proteins have been identified. In this review, we will focus on characterized RKN and CN effectors with known targets in plants (Tables 1 and 2). These interactions predominantly relate to plant defense and immune responses, which can be attributed to well-established immune assays to study effectors (Eves-van den Akker 2021; Molloy et al. 2023). However, there are also effectors related to altering host hormonal and developmental pathways, with potential intersections with immune responses.

## Effector-Mediated Suppression of Basal Plant Defenses

The basal immune response, known as pathogen-associated molecular pattern (PAMP)-triggered immunity (PTI), is an inducible immune response that nematodes must effectively control to successfully infect plants (Holbein et al. 2019; Mendy et al. 2017; Sato et al. 2019). The PAMPs derived from nematodes (i.e., NAMPs) are recognized by membrane-bound pattern-recognition receptors, such as the leucine-rich repeat (LRR) receptor-like kinase called NLR1 in *Arabidopsis* (Mendy et al. 2017) (Fig. 1). Ascaroside 18 is a secreted nematode pheromone that is abundant in both RKNs and CNs, and it acts as a NAMP

to trigger basal immune responses (Manohar et al. 2020). Recently, AtNLR1 was identified as the receptor that directly binds ascaroside 18 and activates PTI (Huang et al. 2023).

One of the hallmarks of PTI is the rapid production of reactive oxygen species (ROS), a phenomenon predominantly governed by membrane-bound NADPH-oxidases. None of the currently characterized nematode effectors interact directly with NADPH-oxidases, but evidence suggests that ROS levels must be finely calibrated to prevent cellular damage and, instead, direct the host metabolism toward infections (Chopra et al. 2021). There are nematode effectors that help achieve a nuanced regulation of ROS levels and their associated defense functions. This is accomplished by effector interactions with plant proteins implicated in ROS accumulation or ROS-mediated signaling cascades (Holbein et al. 2019). For example, the *Meloidogyne graminicola* effector MgMO289 targets a novel rice heavy metal-associated plant protein 04 (OsHPP04), a copper-binding metallochaperone. Both OsHPP04 and MgMO289 stimulate the activity of the cytosolic copper/zinc superoxide dismutase 2, which decreases superoxide radical accumulation in the roots (Song et al. 2021). Another effector shown to affect ROS scavenging and help manage oxidative stress is the *Meloidogyne javanica* effector MjTTL5 (Lin et al. 2016). This effector interacts with *Arabidopsis* FTRc (ferredoxin: thioredoxin reductase catalytic subunit), a key protein that regulates the oxidative stress response in plants

**Table 1.** Summary of characterized cyst nematode effectors with their plant targets

Effector ID <sup>a</sup>	Protein identity <sup>b</sup>	Interaction partner in the host <sup>c</sup>	Reference
Effector-mediated suppression of basal plant defenses			
Hs10A06	Pioneer protein	AtSPDS2 (spermidine synthase 2)	Hewezi et al. 2010
Gr29D09	29D09-like effector family protein	StHXK7 (hexokinase 7); StHXK1 (hexokinase 1)	Chen et al. 2024; Yeh 2021
Hg30C02	Pioneer protein	AtPR2 ( $\beta$ -1,3-endoglucanase)	Hamamouch et al. 2012
NMAS1	Pioneer protein	ATG8 (AuTophagy-related protein 8)	Chen et al. 2023
HaGLAND5	G16B09 family protein	Pyruvate dehydrogenase subunit	Yang et al. 2019b
Hs25A01	25A01-like effector family protein	AUF1 (auxin up-regulated F-box protein 1), CHS (chalcone synthase) and eIF-2bs (translation initiation factor eIF-2 $\beta$ subunit)	Pogorelko et al. 2016
HsSNARE1	t-SNARE domain-containing protein	AtPR1 (pathogenesis-related protein 1) and AtSNAP2 (soluble NSF attachment protein)	Zhao and Liu 2023
HgSLP-1	t-SNARE domain-containing protein	Rhg1 $\alpha$ -SNAP (alpha soluble NSF attachment protein)	Bekal et al. 2015
Effectors and effector-triggered immunity (ETI)			
GrSPRYSEC-15	SPRYSEC family protein	NRC2 (NLR required for cell-death 2) and NRC3 (NLR required for cell-death 3)	Derevnina et al. 2021
Hs4E02	Pioneer protein	RD21A (Responsive to dehydration 21)	Pogorelko et al. 2019
GrVAP1	Venom Allergen-like protein	PLCP-Rcr3pim (papain-like cysteine protease)	Lozano-Torres et al. 2012
GpRbp-1	SPRYSEC-like protein	GPA2 ( <i>R</i> gene) and UPL3 (ubiquitin E3 ligase)	Sacco et al. 2009
GrSPRYSEC-19	SPRYSEC family protein	SISW5 ( <i>R</i> gene)	Postma et al. 2012
Effectors and the nematode feeding site			
Hs2D01	Pioneer protein	HAESA (leucine-rich repeat receptor-like kinase)	Verma et al. 2022
GpSPRY-414-2	SPRYSEC-type	StCLASP (CLIP-associated protein)	Mei et al. 2018
Hs19C07	Pioneer protein	LAX3 (Like Auxin3)	Lee et al. 2011
Hg27D09	Pioneer protein	GmIPK-2 (Interacting plant kinase 2)	Yang et al. 2022
HgCBP	Cellulose-binding protein	Pectin methylesterase	Hewezi et al. 2008
Effectors in the nucleus			
Hs10A07	Pioneer protein	IPK (Interacting plant kinase) and IAA16 (indole-3-acetic acid inducible 16) transcription factor	Hewezi et al. 2015
HsGLAND4	DNA-binding protein	LTP (lipid transfer protein genes)	Barnes et al. 2018
Hs30D08	Pioneer protein	SMU2 (Suppressor of mec-8 and unc-52-2)	Verma et al. 2018
Hs32E03	Pioneer protein	AtHDT1 (Histone deacetylase) and AtFKBP53 (FK506-binding protein 53)	Vijayapalani et al. 2018
HaVAP2	Venom Allergen-like protein 2	HvCLP (CYPRO4-like protein)	Luo et al. 2019
Effectors mimicking host peptides			
Hs4F01	Plant annexins mimic	Oxidoreductase (2OG-Fe (II) oxygenase family)	Patel et al. 2010
CLE-like effectors	CLAVATA3 (CLV3)/ESR (CLE) hormone mimics	StCLV2 (CLAVATA2), BAM1 (Barely any meristem 1) and BAM2 (Barely any meristem 2)	Chen et al. 2015; Guo et al. 2011, 2015

<sup>a</sup> Hs, *Heterodera schachtii*; Hg, *Heterodera glycines*; Ha, *Heterodera avenae*; Gr, *Globodera rostochiensis*; Gp, *Globodera pallida*.

<sup>b</sup> Protein identity corresponds to the known complete or partial annotation of the effector proteins; the identity is still unknown for pioneer proteins.

<sup>c</sup> Interaction partners in the host correspond to the categories based on their targeted plant processes.

(Lin et al. 2016). Yet another *M. javanica* effector is MjNEROS, which was shown to suppress ROS accumulation by localizing to the plastid and interacting with an iron-sulfur protein (ISP) (Stojilković et al. 2022). MjNEROS interferes with ISP function in the electron transport chain to suppress ROS accumulation and other defense responses (Stojilković et al. 2022). In *M. incognita*, the effector MiPDI1 interacts with the tomato stress-associated protein 12 (SAP12), a redox-dependent sensor involved in biotic and abiotic responses (Zhao et al. 2020). This interaction may help protect the nematode from oxidative stress and modulate SAP12-mediated defense responses. Intriguingly, a report found that the *M. incognita* C-type lectin effector MiCTL1a could interact with plant catalase, an antioxidant enzyme (Zhao et al. 2021). Transient expression of *MiCTL1a* could suppress an elicitor-triggered ROS burst, indicating that MiCTL1a interacting with host catalases may help modulate H<sub>2</sub>O<sub>2</sub> levels to promote nematode virulence (Zhao et al. 2021).

Because ROS are often linked to plant defense responses, it was surprising that CNs activate NADPH oxidases RbohD and RbohF to produce ROS during infections. Moreover, the NADPH-oxidase mutants are more resistant to CNs, indicating that ROS are required for successful cyst nematode parasitism (Siddique et al. 2014). The data stand as an intriguing deviation from the prevalent association of ROS with plant defenses (Heller and Tudzynski 2011). However, CNs are likely using effectors to tune ROS levels in order to stimulate specific downstream responses, including auxin accumulation, that help promote cyst infections (Chopra et al. 2021). A case in point is the effector from *Heterodera schachtii* Hs10A06, which was shown to interact with *Arabidopsis* spermidine synthase (AtSPDS2), an enzyme involved in polyamine biosynthesis

(Hewezi et al. 2010; Mejias et al. 2019). Nematode-infected plants exhibited increased polyamine oxidase (PAO) gene expression and PAO activity. The authors concluded that when Hs10A06 is secreted into the plant, it causes an increase in spermidine content and a subsequent increase in PAO activity (Hewezi et al. 2010). Because the catabolism of spermidine by PAO leads to the production of H<sub>2</sub>O<sub>2</sub>, the effector may be stimulating low levels of ROS to promote infection, or it may be controlling the ROS levels from reaching cytotoxic levels by increasing antioxidant gene expression (Hewezi et al. 2010). Another CN effector is the *Globodera rostochiensis* 29D09 (Jones et al. 2009), which interacts with the mitochondria-associated hexokinase 7 (StHXX7). When *Gr29D09* is transiently expressed in *Nicotiana benthamiana* leaves, it suppresses flg22-ROS production. It does so by either enhancing StHXX7 expression or stabilizing the StHXX7 protein on the mitochondria, thereby stimulating the cellular antioxidant potential and effectively controlling ROS levels within the cell (Yeh 2021). More recently Chen et al. (2024) showed that *Gr29D09* also interacts with potato hexokinase 1 (StHXX1), which is homologous to StHXX7. Transient expression of *StHXX7* in *N. benthamiana* leaves also suppressed flg22-induced ROS and defense gene expression (Chen et al. 2024).

In addition to impacting ROS levels, effectors also target other proteins integral to the immune responses of host plants. For instance, the CN effector Hg30C02 was shown to interact with PR2, an *Arabidopsis* β-1,3-endoglucanase (Hamamouch et al. 2012). This interaction may control defense-related callose formation or prevent the release of β-1,3-glucan oligosaccharides that function as damage associated molecular patterns. Another example from CNs is the effector Manipulator of Au-

**Table 2.** Summary of characterized root-knot nematode effectors with their plant targets

Effector ID <sup>a</sup>	Protein identity <sup>b</sup>	Interaction partner in the host <sup>c</sup>	Reference
Effector-mediated suppression of basal plant defenses			
MgMO237	Pioneer protein	OsGSC (1,3-β-glucan synthase component), OsCRRSP55 (Cysteine-rich repeat secretory protein 55), and OsBetvI (Pathogenesis-related BetvI family protein)	Chen et al. 2018
MgMO289	Pioneer protein	OsHPP04 (Heavy metal-associated plant protein 04)	Song et al. 2021
MjNEROS	Pioneer protein	ISP (iron-sulfur protein)	Stojilković et al. 2022
MjTTL5	Pioneer protein	AtFTRc (ferredoxin:thioredoxin reductase catalytic subunit)	Lin et al. 2016
MiPDI1	Protein disulfide isomerase-like	SISAP12 (stress-associated protein 12)	Zhao et al. 2020
MiCTL1a	C-type lectin	Catalase	Zhao et al. 2021
MiMIF2	Macrophage migration inhibitory factor-like	AnnAt1 (annexin 1), AnnAt4 (annexin 4)	Zhao et al. 2019
Mc1194	Pioneer protein	RD21A (Responsive to dehydration 21)	Davies et al. 2015
MiMSP32	Pioneer protein	OPR2 (12-oxophytodienoate reductase 2)	Verhoeven et al. 2023
MaMsp4	Pioneer protein	B-g11 (β-galactosidase 11), sam2 (S-adenosyl methionine decarboxylase 2), pec (pectinesterase), eth-pct (ethanolamine-phosphate cytidyltransferase)	Przybylska et al. 2023
Effectors and the nematode feeding site			
Minc00344	Pioneer protein	GmHub10 (kinesin light chain-related protein 2)	Godinho Mendes et al. 2021
MjNULG1a	Pioneer protein	GmHub10 (kinesin light chain-related protein 2)	Godinho Mendes et al. 2021
Mi8D05	Pioneer protein	TIP2 (tonoplast intrinsic protein 2)	Xue et al. 2013
Effectors in the nucleus			
MiEFF18	Pioneer protein	Smd1 (spliceosomal protein)	Mejias et al. 2021
MiEFF1	Pioneer protein with nuclear localization signal (NLS)	USP (universal stress protein) and GAPCs (cytosolic glyceraldehyde-3-phosphate dehydrogenases)	Truong et al. 2021
Mg16820	Pioneer protein	DIP1 (Dehydration stress-inducible protein 1)	Naalden et al. 2018
Effectors mimicking host peptides			
MiRALF1, MiRALF2	Rapid alkalization factor (RALF) mimic	FERONIA receptor kinase	Zhang et al. 2020
Mi16D10	CLAVATA3 (CLV3)/ESR (CLE) hormone mimic	SCL (SCARECROW-like transcription factors)	Huang et al. 2006

<sup>a</sup> Mi, *Meloidogyne incognita*; Mg, *Meloidogyne graminicola*; Mj, *Meloidogyne javanica*.

<sup>b</sup> Protein identity corresponds to the known complete or partial annotation of the effector proteins; the identity is still unknown for pioneer proteins.

<sup>c</sup> Interaction partners in the host correspond to the plant proteins that are targeted by the effector proteins.



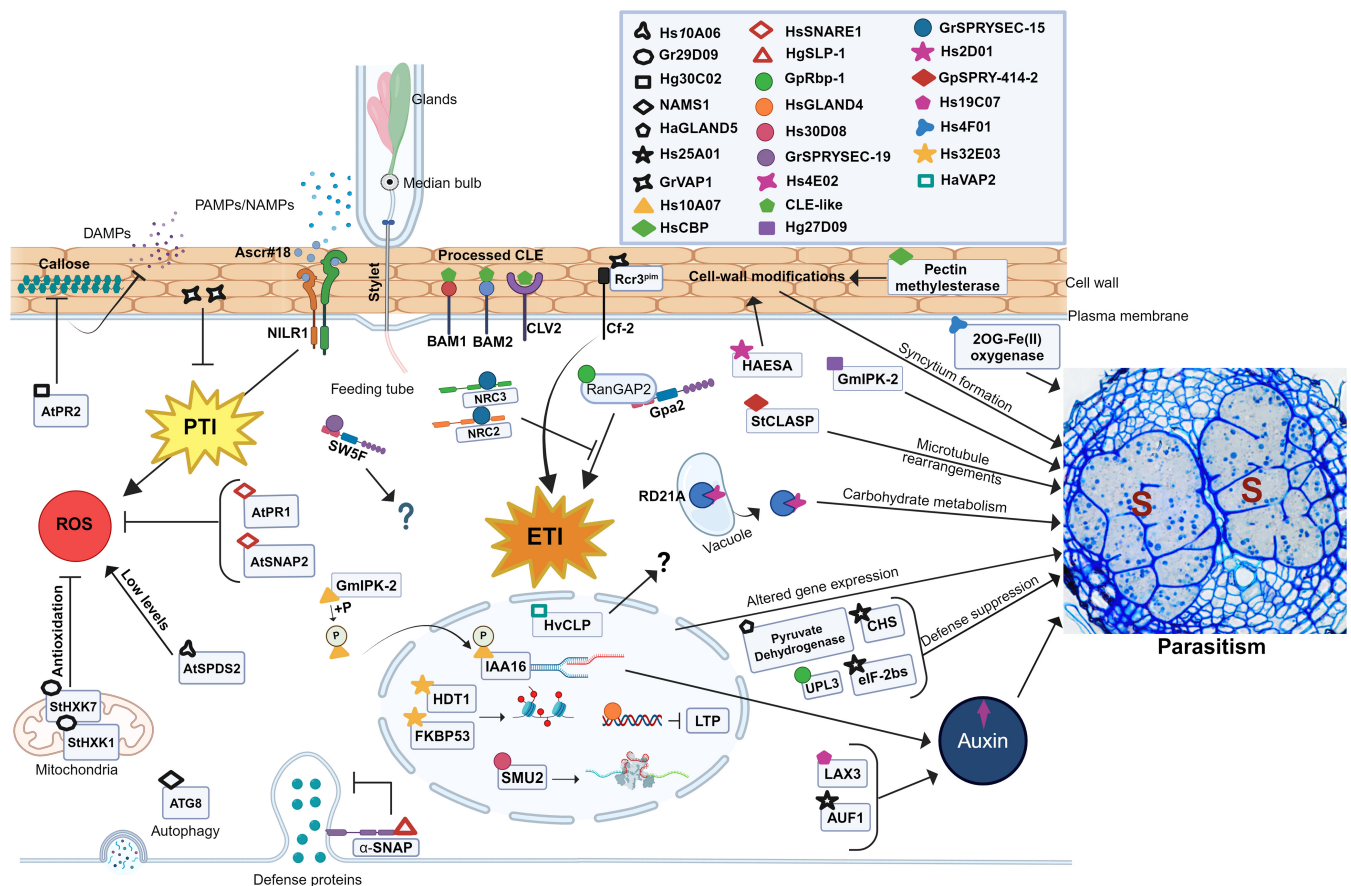
tophagy System 1 (NMAS1). This effector interacts with the AuTophagy-related protein 8 (ATG8) (Chen et al. 2023). It is hypothesized that the NMAS1 binds to ATG8 to inhibit defense-related selective autophagy. Thus, the NMAS1-ATG8 interaction might facilitate infections by orchestrating the degradation of cell-surface immune receptors, thereby altering basal immune responses, and contributing to the ability of the nematode to evade host defenses (Chen et al. 2023).

CNs contain the “G16B09 family” of effectors, which includes 2D01, 16B09, 2D01, and GLAND5 (Cotton et al. 2014; Thorpe et al. 2014; Yang et al. 2019a, b). Within this family, the HaGLAND5 effector can interact with the *Arabidopsis* pyruvate dehydrogenase subunit. This interaction may interfere with salicylic acid (SA)-mediated plant defenses, which thereby can thereby help promote nematode parasitism (Yang et al. 2019b).

Exogenous application of the hormone jasmonic acid (JA) can enhance plant resistance to RKNs (Fujimoto et al. 2011; Gleason et al. 2016). Interestingly, the *M. incognita* effector MiMSP32 was shown to interact with 12-oxophytodienoate reductase 2 (OPR2), an enzyme involved in converting a 12-OPDA derivative to JA (Chini et al. 2018). Thus, this RKN effector seems to be altering OPR2-dependent host defenses to promote nematode parasitism (Verhoeven et al. 2023).

Although many of the examples provided describe effectors converging on one host target, there are also examples of

a single effector interacting with multiple host proteins. For example, the *Meloidogyne arenaria* effector MaMSP4 interacts with several host proteins, including  $\beta$ -galactosidase 11, pectinesterase, *S*-adenosyl methionine decarboxylase 2, and ethanolamine-phosphate cytidylyltransferase (Przybylska et al. 2023).  $\beta$ -galactosidase 11 and pectinesterase are known to play a role in cell wall modifications, and *S*-adenosyl methionine decarboxylase 2 and ethanolamine-phosphate cytidylyltransferase are involved in polyamine-mediated plant defenses (Przybylska et al. 2023). The *H. schachtii* effector Hs25A01 can interact with *Arabidopsis* AUXIN UP-REGULATED F-BOX PROTEIN1 (AUF1), a chalcone synthase (CHS), and the translation initiation factor eIF-2 $\beta$  subunit (eIF-2bs) (Pogorelko et al. 2016). The effector Hs25A01 may be involved in regulating auxin levels through interaction with AUF1, CHS, and eIF-2bs, interfering with production of defense compounds, and quelling plant defenses by suppressing translation of defense proteins (Pogorelko et al. 2016). The *M. graminicola* effector MgMO237 targets three distinct rice defense-related proteins: 1,3- $\beta$  glucan synthase component (OsGSC), cysteine-rich repeat secretory protein 55 (OsCRRSP 55), and pathogenesis-related BetvI family protein (OsBetvI) (Chen et al. 2018). This effector’s interaction with these defense-related proteins suppresses several hallmarks of PTI in the host, including defense gene expression, callose deposition, and the ROS burst. In *H. schachtii*, a



**Fig. 1.** Overview of the characterized cyst nematode (CN) effectors and their plant targets. CN effectors are typically secreted by the nematode stylet into the apoplast or directly into the host cell cytoplasm. In some interactions, the effectors are targeting plant proteins involved in pathogen-associated molecular pattern (PAMP)-triggered immunity (PTI) or effector-triggered immunity (ETI). Many effectors have plant targets involved in controlling reactive oxygen species (ROS) levels or their downstream responses. Some effectors contain nuclear localization signals and move to the nucleus of the feeding site. They may carry plant targets to the nucleus or bind to plant targets within the nucleus to modulate host responses. A subset of effectors, such as the CLE-like effectors and the cyst annexin-like effector (Hs4F01) can mimic plant proteins/peptides. A localized auxin accumulation is modulated by effectors targeting the transcriptional regulation of auxin-responsive genes and auxin transport proteins. Although CN effector research is summarized here, some effectors may be secreted by both CNs and root-knot nematode (RKNs). Cyst nematode syncytium photo provided by Dr. Mirosław Sobczak. The figure was created in BioRender.com. S, syncytium; DAMPs, damage associated molecular patterns; NAMPS, nematode associated molecular patterns.

SNARE domain-containing effector HsSNARE1 can bind to two plant targets: AtPR1 and AtSNAP2. Expression of HsSNARE1 in transgenic *Arabidopsis* suppressed plant expression of *AtPR1* (pathogenesis-related protein 1) and *AtSHMT4* (serine hydroxymethyltransferase isoform 4) (Zhao and Liu 2023). The overexpression of *AtPR1*, *AtSNAP2*, and *AtSHMT4* enhanced resistance against *H. schachtii* (Zhao and Liu 2023; Zhao et al. 2022).

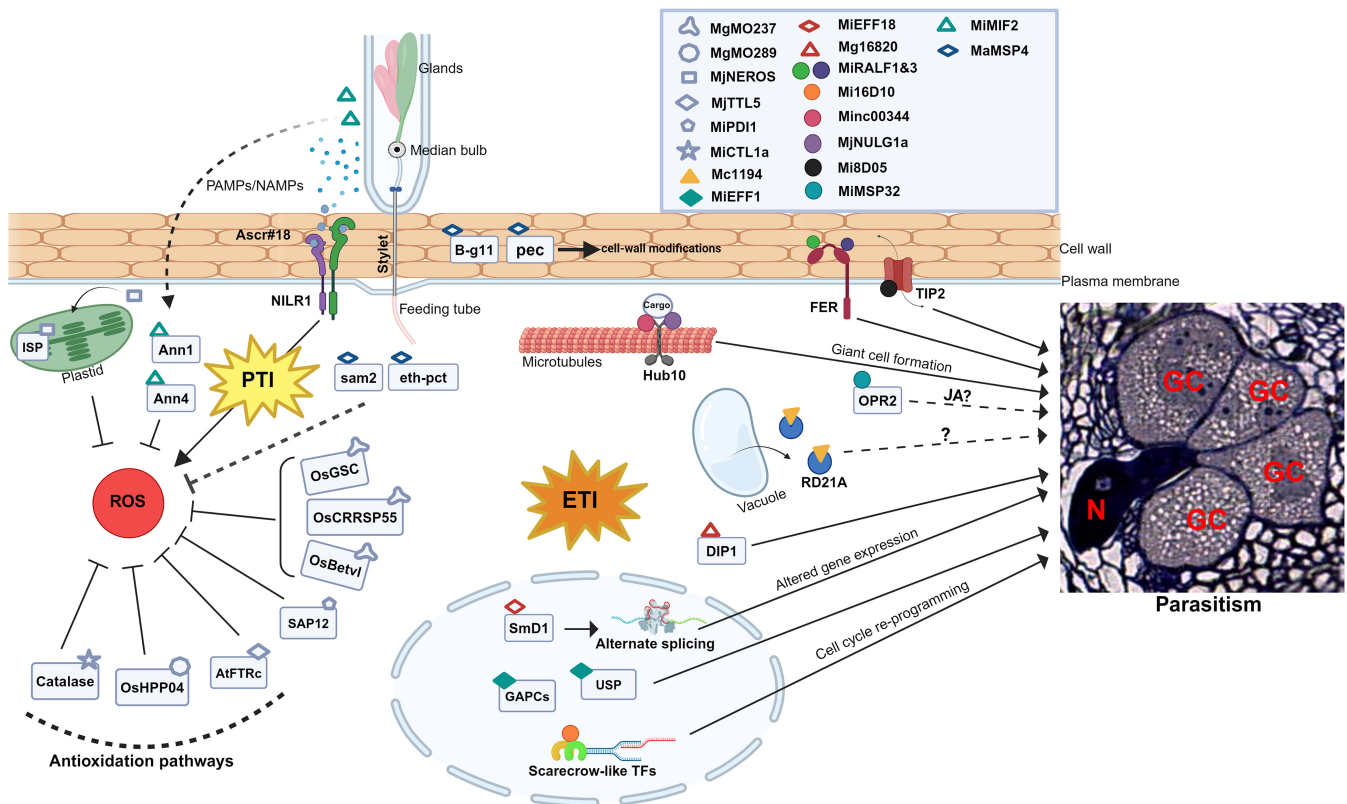
The HsSNARE effector is interesting, because SNAREs are proteins that are involved in vesicle fusion to the plasma membrane, and work in plant-pathogen interactions indicates that some SNAREs may be controlling the trafficking of immune proteins (Yun et al. 2023). The *Heterodera glycines* effector HgSLP-1, which contains a tSNARE domain, co-purified with the soybean resistance protein Rhg1  $\alpha$ -SNAP (Bekal et al. 2015). The role of HgSLP-1 is still speculative, but it may be binding to and sequestering the  $\alpha$ -SNAP to prevent membrane fusion of vesicles containing defense-related or other proteins. However, considering that virulent populations of SCN have fewer copies of HgSLP-1, this effector may also function as an avirulence protein (Bekal et al. 2015). RKN and CN effector targets in basal defenses are summarized in Figures 1 and 2.

## Effectors and Effector-Triggered Immunity (ETI)

Plants contain resistance (*R*) genes that encode intracellular nucleotide-binding LRR (NLR) receptors (Lolle et al. 2020). The recognition of pathogen effectors, either directly or indirectly, by the cognate *R*-protein can trigger a strong immune response, and this is referred to as ETI. In the case of sedentary endoparasitic

nematodes, ETI is typically characterized by a hypersensitive response (HR) around the head of the nematode, presumably to block feeding and prevent successful infection (Williamson and Kumar 2006). To counteract ETI, nematodes have effectors that prevent *R* gene-mediated defense activation. For example, the SPRYSEC family of effectors from CNs have been studied for their ability to function as suppressors of the plant immune responses (Ali et al. 2015; Diaz-Granados et al. 2016). More specifically, members of the *G. rostochiensis* SPRYSEC effector family can suppress the HR mediated by specific *R* proteins containing coiled-coil nucleotide-binding LRR (CC-NB-LRR) domains (Ali et al. 2015, 2017; Diaz-Granados et al. 2016; Mei et al. 2015; Postma et al. 2012). One of the SPRYSEC effector members, GrSPRYSEC-19, can specifically interact with SW5F, a member of the SW5 *R*-gene cluster in tomato (Rehman et al. 2009). Curiously, the interaction between SPRYSEC-19 and SW5F did not suppress cell death or disease resistance; SPRYSEC-19 can only suppress defense-related cell death mediated by some CC-NB-LRR resistance proteins (Postma et al. 2012). The cyst nematode SPRYSEC family has undergone a large expansion and functional diversification, which may help add novel versatility to the SPRYSEC effectors that enables interactions with one or more host targets and disruption of plant immune responses (Diaz-Granados et al. 2016).

Only a few nematode resistance genes have been cloned so far, and among them, the cyst nematode *R* gene *Gpa2* is well characterized (Van Der Vossen et al. 2000; Zheng et al. 2021). In the potato cyst nematode *Globodera pallida*, the SPRYSEC effector called GpRbp-1 is involved in *Gpa2*-dependent HR when transiently expressed in *N. benthamiana* leaves (Sacco et al. 2009). The GpRbp-1 effector directly interacts with the RanGTP-



**Fig. 2.** Overview of the characterized root-knot nematode (RKN) effectors and their plant target proteins. RKN effectors secreted by the nematode stylet are deposited into the apoplast or into the host cell cytoplasm. A cohort of effectors are known to suppress pathogen-associated molecular pattern (PAMP)-triggered immunity (PTI), and some of these effectors help orchestrate an intricate manipulation of host reactive oxygen species (ROS) levels and their subsequent downstream signaling cascades. Specific RKN effectors were shown to enter the plant cell nucleus. Their roles are postulated to be modulating gene expression, protecting nucleic acids, or affecting gene splicing to alter host responses. The rapid alkalization factors (RALF)-like and 16D10 effectors are host mimics that alter plant developmental and physiological processes to promote successful parasitism. The figure was created in BioRender.com. GC, giant cell; N, nematode.

binding protein RanGAP2. The CC domain of GPA2 also binds to RanGAP2, and it is hypothesized that RanGAP2 is monitored by GPA2, and the interaction between GpRbp-1 and RanGAP2 triggers the GPA2-mediated HR (Sacco et al. 2009).

Interestingly, GpRbp-1 has also been shown to interact with a potato homolog of the Homology to E6-AP C-Terminus (HECT)-type ubiquitin E3 ligase, UPL3, whose function is to transfer ubiquitin from the E2 ligase to the target protein (Diaz-Granados et al. 2020). UPL3 is known to play a role in plant immune responses because mutations in UPL3 abrogated SA-mediated defenses (Furniss et al. 2018). The interaction between the effector GpRbp-1 and UPL3 points to the nematode manipulating proteasome-dependent ubiquitination to promote nematode parasitism, potentially through changes in SA-dependent host immunity (Diaz-Granados et al. 2020).

Most characterized effectors are involved in host defense suppression, but in some cases, the effector is recognized by the plant to trigger plant defenses. One such case is the venom allergen-like effector VAP1 from the CNs. In the susceptible host, VAPs are thought to play roles in the suppression of basal immune responses by binding to an apoplastic papain-like cysteine protease (PLCP) called Rcr3<sup>pimm</sup> (Lozano-Torres et al. 2014). However, if the plasma membrane immune receptor Cf-2 is also present, it can detect the GrVAP1-induced perturbations of the Rcr3<sup>pimm</sup>, triggering an HR in the tomato (Lozano-Torres et al. 2012). Interestingly, like VAP1, *H. schachtii* effector Hs4E02 also was shown to interact with a PLCP. Hs4E02 interacts with the *Arabidopsis* PCLP called Responsive to Dehydration 21A (RD21A), and although this interaction does not affect RDA21 enzymatic activity, the effector relocates RD21A from the vacuole to the nucleus and cytoplasm (Pogorelko et al. 2019). The mislocalized protease promotes parasitism, perhaps by altering the carbohydrate metabolism of the host (Pogorelko et al. 2019). Like 4E02 in CNs, the RKN effector Mc1194 can also interact with RD21A. *Arabidopsis* mutants of RD21A showed enhanced susceptibility to *Meloidogyne chitwoodi* infection (Davies et al. 2015). Although the underlying mechanism of action of this effector is unknown, its interaction with a PCLP underscores the importance of PCLP in plant immunity (Misas-Villamil et al. 2016).

The most well characterized *R* genes are the sensor NLRs, which encode proteins with the canonical TIR (toll/interleukin-1)-type or CC-type NB-LRR architecture. *Gpa2* in potato, which provides resistance against certain populations of *G. pallida* (Gartner et al. 2021), and *Mil.2* in tomato (Kaloshian and Teixeira 2019), which confers resistance against certain species of RKNs, are both CC-NB-LRR-type sensor NLRs. The NLR immune network also has “helper” NLRs that help translate the recognition of the effector signal by the sensor into the signaling pathway leading to HR. These NLR helpers are referred to as NRCs (NB-LRR protein required for HR-associated cell death) and are part of a widely expanded family in the Solanaceae (Wu et al. 2017). Evaluation of the CN effector GrSPRYSEC-15 showed that the effector could suppress the cell death induced by autoactive NRC2 and NRC3 (Derevnina et al. 2021). SPRYSEC-15 binds the nucleotide-binding domain of NRC2 and NRC3 and alters their function, resulting in suppressed host immune responses (Derevnina et al. 2021). Recently, it was shown that the binding of SPRYSEC-15 to NRC2 inhibited the oligomerization of the NRC2 proteins (Contreras et al. 2023). Inhibiting NRC2 oligomerization subsequently prevented the activation of downstream immune signaling (Contreras et al. 2023). RKN and CN effector targets in ETI are summarized in Figures 1 and 2.

## Effectors and the Nematode Feeding Site

Sedentary endoparasitic nematodes must establish a feeding site within plant roots in order to survive. Thus, it makes

sense that some nematode effectors are playing roles in feeding site formation and maintenance. For example, the RKN effector 8D05 was shown to interact with plant aquaporin tonoplast intrinsic protein 2 (TIP-2) that is predicted to facilitate giant cell formation by regulating cellular turgor pressure via water and solute transport (Xue et al. 2013). The *H. schachtii* effector Hs2D01 was shown to interact with the intracellular kinase domain of HAESA (HAE), an LRR receptor-like kinase (Verma et al. 2022). HAE has documented roles in cell wall modifications for abscission and lateral root emergence, and the interaction between Hs2D01 and HAESA may modulate its activity to promote infections, perhaps by regulating the expression of cell wall-modifying enzymes needed for syncytial formation (Verma et al. 2022).

Both CNs and RKNs secrete cellulose-binding proteins (CBPs) (Ding et al. 1998; Gao et al. 2003, 2004). Interestingly, the CBPs do not have enzymatic activity but have roles in RKN and CN pathogenesis. In the case of the CNs, CBP can bind to and enhance the activity of an *Arabidopsis* pectin, methylesterase (Hewezi et al. 2008). As a result, there is an increase in the levels of methylesterified pectin in the cell walls, which may allow for cell wall-modifying enzymes to have increased access to cell wall polymers and accelerate the development of the syncytium (Hewezi et al. 2008).

The SPRYSEC effector from *G. pallida*, GpSPRY-414-2, was shown to interact with a potato cytoplasmic linker protein (CLIP)-associated protein (StCLASP). The StCLASPs regulate the dynamics of microtubules (Mei et al. 2018). During syncytial formation, the microtubule network undergoes rearrangement (de Almeida Engler et al. 2010); GpSPRY-414-2 may be binding to the CLASPs to target the microtubules to regulate the development of syncytia. CLASPs are also involved in auxin and brassinosteroid signaling (Ambrose et al. 2007; Ruan et al. 2018), both of which are hormones involved in plant–nematode interactions (Gheysen and Mitchum 2019). GpSPRY-414-2 may be influencing these hormone signaling pathways to help establish successful infections.

Auxin is an important hormone needed to invoke the transcriptional and developmental changes required for CN and RKN feeding site formation (Abril-Urias et al. 2023; Gheysen and Mitchum 2019). Nematode effectors are known to manipulate at least one mechanism linked to the nematode-induced changes in auxin homeostasis. In the case of CNs, the *H. schachtii* effector Hs19C07 has been shown to bind to the auxin importer (LAX3), and this may result in the increased auxin levels and flow in the developing syncytia (Lee et al. 2011).

In some cases, the same plant protein is targeted by more than one nematode effector during infections, perhaps indicating the importance of specific targeted host processes for successful nematode parasitism. For example, two RKN effectors, Minc00344 and MjnNULG1a, were shown to interact with the soybean Hub10 protein (Godinho Mendes et al. 2021). In *Arabidopsis*, Hub10 is a kinesin light chain-related protein 2 involved in moving cargo along microtubules in plant cells for plant defenses and developmental responses, including pollen tube growth and cell wall deposition (Ganguly et al. 2020; Mukhtar et al. 2011). Interestingly, bacterial and oomycete pathogens also target Hub10 in *Arabidopsis*, indicating that diverse pathogens are focusing on this “immunity hub” protein to promote their infections (Mukhtar et al. 2011).

In the CN–plant interaction, two CN effectors (27D09 and 10A07) are targeting “interacting plant kinases” (IPKs) in order to exploit the phosphorylation activity of the host kinases. The *H. glycines* effector Hg27D09 interacts with the plant kinase GmIPK-2 (Yang et al. 2022); overexpression of *GmIPK-2* enhanced plant susceptibility to CNs, but the specific downstream signaling processes leading to the enhanced nematode suscepti-



bility are not clear. GmPK-2 is also targeted by the CN effector Hs10A07 (Hewezi et al. 2015). The phosphorylated Hs10A07 translocates to the nucleus, where it binds to the IAA16 transcription factor and regulates auxin responses (Hewezi et al. 2015). RKN and CN effector targets in feeding site formation are summarized in Figures 1 and 2.

## Effectors in the Nucleus

Some nematode effectors have been shown to localize to plant nuclei during nematode infections (Chen et al. 2017; Harris et al. 2023; Hewezi et al. 2015; Jaouannet et al. 2012; Molloy et al. 2023; Tytgat et al. 2004; Zhang et al. 2015). One example is the CN effector 10A07, described earlier, which translocates to the nucleus, but only after host-mediated phosphorylation. Another CN effector called Hs10A07 relocates to the nucleus, where it binds to the auxin-responsive protein IAA16 (INDOLE-3-ACETIC ACID INDUCIBLE16) and modulates the transcription of specific auxin response genes in the developing syncytium (Hewezi et al. 2015). The cereal cyst nematode effector HaVAP2, a homolog of VAP1, also localizes to the nucleus. HaVAP2 physically interacts with a *Hordeum vulgare* CYPRO4-like protein (HvCLP) in the nucleus, although the role of HvCLP in plant–nematode interactions has not been investigated (Luo et al. 2019). Another example of a nuclear-localized effector is GLAND4, found in both *H. glycines* and *H. schachtii*. GLAND4 translocates into the host cell nucleus and binds to DNA in the promoters of two lipid transfer protein (*LTP*) genes. Because the overexpression of the two *LTP* genes decreases the host's susceptibility to nematodes, it is hypothesized that GLAND4 is a nuclear effector involved in *LTP* gene suppression (Barnes et al. 2018). The CN effector Hs30D08 is also targeted to the nucleus. Once in the nucleus, Hs30D08 interacts with SMU2 (homolog of *suppressor of mec-8 and unc-52 2*). The SMU protein is a member of the spliceosome complex that splices out introns during mRNA processing and mediates alternative splicing (Chung et al. 2009). The interaction between SMU2 and Hs30D08 may influence host gene expression and the regulation of genes that are critical for infection (Verma et al. 2018). Similarly, the *M. incognita* effector MiEFF18 localizes to the plant nucleus and interacts with the plant ribonucleosomal protein SmD1 (Mejias et al. 2021). SmD1 is essential for the functioning of the spliceosome complex, and its interaction with MiEFF18 modulates the alternate splicing of plant genes to promote nematode infection. Another RKN effector called MiEFF1 was shown to be secreted in planta and targeted to the giant cell nuclei (Jaouannet et al. 2012). It interacts with both the universal stress protein (USP) and cytosolic glyceraldehyde-3-phosphate dehydrogenases (GAPDHs) in *Arabidopsis* (Truong et al. 2021). MiEFF1 may be modulating the GAPCs in the giant cells to promote feeding site formation and maintenance, perhaps through a role as a negative defense regulator (Truong et al. 2021). Mi16D10 is an RKN effector that also moves to the nucleus, and it binds to a host transcription factor (Huang et al. 2006). The *M. graminicola* effector Mg16820 is secreted to the apoplast and cytoplasm (Naalden et al. 2018). However, during giant cell formation, Mg16820 is found in the cytoplasm and nucleus. Because Mg16820 is small (7.5 kilodaltons), it may be diffusing into the nucleus rather than being imported into it (Naalden et al. 2018). Mg16820 interacts with DIP1 (dehydration stress-inducible protein 1), a protein with roles in abscisic acid responses (Naalden et al. 2018).

Another mechanism of host cell reprogramming is through the epigenetic regulation of susceptibility genes. The *H. schachtii* effector Hs32E03 localizes to the nucleus, and it is also the first nematode effector that exerts influence by modulating the acetylation of histones in plants (Vijayapalani et al. 2018). The histone modifications are mediated through the direct interaction of

Hs32E03 with two proteins in the plant nucleus: *Arabidopsis* histone deacetylase HDT1 and the FK506-binding protein FKBP53 (Vijayapalani et al. 2018). By binding and inhibiting the histone deacetylase, Hs32E03 causes chromatin changes that alter plant rRNA levels. The impact of Hs32E03 is dose dependent; at low levels, it promotes parasitism, while higher levels lead to silencing of rDNA and an impediment of nematode development (Castroverde 2018; Vijayapalani et al. 2018). Nucleus-targeted RKN and CN effectors and their targets are summarized in Figures 1 and 2.

## Effectors Mimicking Host Peptides

Some nematode effectors have been shown to mimic plant host proteins to subvert plant developmental or defense pathways. For example, the *H. schachtii* effector Hs4F01, a mimic of plant annexin, interacts specifically with a member of the plant oxidoreductase of the 2OG-Fe (II) oxygenase family (Patel et al. 2010). The role of the 2OG-Fe (II) oxygenase in plant–nematode interactions is unclear, but downy mildew research suggests that plant 2OG-Fe (II) oxygenase is required for pathogen susceptibility (Van Damme et al. 2008). RKN effectors MiRALF1 and MiRALF3 are peptide mimics of the highly conserved plant rapid alkalization factors (RALFs) that are involved in plant growth, development, and defense (Zhang et al. 2023). Similar to plant RALFs, the MiRALFs can interact with the receptor-like kinase FERONIA (FER). Data suggest that the nematode MiRALF peptides block the FER receptor to regulate nematode parasitism (Zhang et al. 2020). Equally intriguing are the CLE-like effectors found in CNs that resemble plant CLAVATA3 (CLV3)/ESR (CLE) hormones. The CLE-like effectors are secreted into the host cell by the nematode. They are then processed and secreted into the apoplast by the plant's endogenous machinery (Chen et al. 2015). The perception of processed nematode CLE-like peptides activates downstream signaling processes needed for the feeding site formation and maintenance (Kiyohara and Sawa 2012). Processed CLE-like effectors from the cyst nematodes (potato cyst, beet cyst, soybean cyst) have been shown to bind to receptors CLAVATA2-like (StCLV2), BAM1 and BAM2 (Chen et al. 2015; Guo et al. 2011, 2015), although the interaction data require further verification (Mitchum and Liu 2022; Shinohara and Matsubayashi 2015). Interestingly, the RKN effector Mi16D10 also exhibits sequence similarity to the CLE protein family (Huang et al. 2006); 16D10 stimulates root proliferation and differentiation when ectopically expressed in the host by interacting with two GRAS family SCARECROW-like (SCL) transcription factors (Huang et al. 2006). *AtSCL* genes may be involved in reprogramming the cell cycle to promote the giant cell formation. RKN and CN effectors acting as host mimics and their host targets are summarized in Figures 1 and 2.

## Concluding Remarks and Future Directions

Plant parasitic nematodes have evolved an array of effectors that manipulate various aspects of plant physiology, immunity, and development to establish successful infections. These effectors target a wide range of host proteins and pathways, from suppressing PTI and ETI to influencing hormone signaling, modifying cell walls, and even altering nuclear processes, such as gene expression (Figs. 1 and 2; Tables 1 and 2). The work on nematode effectors underscores the importance of host–parasite coevolution and the intricate arms race between nematodes and their plant hosts. Despite significant progress, many questions remain unanswered in the field of nematode effector research. Future studies should focus on uncovering the precise molecular mechanisms by which effectors target and manipulate host proteins, as well as the downstream signaling events that result from



these interactions. Additionally, investigating the genetic diversity of nematode populations and the variations in their effector repertoires will provide insights into how different nematode species and populations within a species can adapt to different hosts. Understanding the full scope of nematode effectors and their interactions with host plants holds promise for developing novel strategies to mitigate nematode-induced crop losses. Targeting specific effectors or their host targets could potentially lead to the development of resistant crops or innovative control measures. Ultimately, a comprehensive understanding of nematode effectors and their roles in parasitism will contribute to more sustainable and resilient agricultural practices.

## Acknowledgments

We thank Itsuhiro Ko (Washington State University) for help in figure preparation using BioRender. We acknowledge Mirosław Sobczak (Department of Botany, Warsaw University of Life Sciences Warsaw, Poland) for letting us use the cyst nematode syncytium photo. We apologize to researchers whose work was not mentioned or cited in this review owing to space and cited literature limitations.

## Literature Cited

- Abril-Urias, P., Ruiz-Ferrer, V., Cabrera, J., Olmo, R., Silva, A. C., Díaz-Manzano, F. E., Domínguez-Figueroa, J., Martínez-Gómez, Á., Gómez-Rojas, A., Moreno-Risueno, M. Á., Fenoll, C., and Escobar, C. 2023. Divergent regulation of auxin responsive genes in root-knot and cyst nematodes feeding sites formed in *Arabidopsis*. *Front. Plant Sci.* 14:1024815.
- Ali, M. A., Azeem, F., Abbas, A., Joyia, F. A., Li, H., and Dababat, A. A. 2017. Transgenic strategies for enhancement of nematode resistance in plants. *Front. Plant Sci.* 8:750.
- Ali, S., Magne, M., Chen, S., Obradovic, N., Jamshaid, L., Wang, X., Bélair, G., and Moffett, P. 2015. Analysis of *Globodera rostochiensis* effectors reveals conserved functions of SPRYSEC proteins in suppressing and eliciting plant immune responses. *Front. Plant Sci.* 6:623.
- Ambrose, J. C., Shoji, T., Kotzer, A. M., Pighin, J. A., and Wasteneys, G. O. 2007. The *Arabidopsis* CLASP gene encodes a microtubule-associated protein involved in cell expansion and division. *Plant Cell* 19:2763-2775.
- Barnes, S. N., Wram, C. L., Mitchum, M. G., and Baum, T. J. 2018. The plant-parasitic cyst nematode effector GLAND4 is a DNA-binding protein. *Mol. Plant Pathol.* 19:2263-2276.
- Bekal, S., Domier, L. L., Gonfa, B., Lakhssassi, N., Meksem, K., and Lambert, K. N. 2015. A SNARE-like protein and biotin are implicated in soybean cyst nematode virulence. *PLoS One* 10:e0145601.
- Bird, D. M., Jones, J. T., Opperman, C. H., Kikuchi, T., and Danchin, E. G. J. 2015. Signatures of adaptation to plant parasitism in nematode genomes. *Parasitology* 142:S71-S84.
- Caillaud, M.-C., Dubreuil, G., Quentin, M., Perfus-Barbeoch, L., Lecomte, P., de Almeida Engler, J., Abad, P., Rosso, M.-N., and Favery, B. 2008. Root-knot nematodes manipulate plant cell functions during a compatible interaction. *J. Plant Physiol.* 165:104-113.
- Castroverde, C. D. M. 2018. Worming into the plant chromatin: A nematode effector influences host histone acetylation. *Plant Cell* 30:2883-2884.
- Chen, J., Chen, S., Xu, C., Yang, H., Achom, M., and Wang, X. 2023. A key virulence effector from cyst nematodes targets host autophagy to promote nematode parasitism. *New Phytol.* 237:1374-1390.
- Chen, J., Hu, L., Sun, L., Lin, B., Huang, K., Zhuo, K., and Liao, J. 2018. A novel *Meloidogyne graminicola* effector, MgMO237, interacts with multiple host defence-related proteins to manipulate plant basal immunity and promote parasitism. *Mol. Plant Pathol.* 19:1942-1955.
- Chen, J., Lin, B., Huang, Q., Hu, L., Zhuo, K., and Liao, J. 2017. A novel *Meloidogyne graminicola* effector, MgGPP, is secreted into host cells and undergoes glycosylation in concert with proteolysis to suppress plant defenses and promote parasitism. *PLoS Pathog.* 13:e1006301.
- Chen, S., Lang, P., Chronis, D., Zhang, S., De Jong, W. S., Mitchum, M. G., and Wang, X. 2015. In planta processing and glycosylation of a nematode CLAVATA3/ENDOSPERM SURROUNDING REGION-like effector and its interaction with a host CLAVATA2-like receptor to promote parasitism. *Plant Physiol.* 167:262-272.
- Chen, S., Tran, T. T. T., Yeh, A. Y.-C., Yang, H., Chen, J., Yang, Y., and Wang, X. 2024. The *Globodera rostochiensis* Gr29D09 effector with a role in defense suppression targets the potato hexokinase 1 protein. *Mol. Plant-Microbe Interact.* 37:25-35.
- Chini, A., Monte, I., Zamarreño, A. M., Hamberg, M., Lassueur, S., Reymond, P., Weiss, S., Stintzi, A., Schaller, A., Porzel, A., García-Mina, J. M., and Solano, R. 2018. An OPR3-independent pathway uses 4,5-didehydrojasmonate for jasmonate synthesis. *Nat. Chem. Biol.* 14:171-178.
- Chopra, D., Hasan, M. S., Matera, C., Chitambo, O., Mendy, B., Mahlitz, S.-V., Naz, A. A., Szumski, S., Janakowski, S., Sobczak, M., Mithöfer, A., Kyndt, T., Grundler, F. M. W., and Siddique, S. 2021. Plant parasitic cyst nematodes redirect host indole metabolism via NADPH oxidase-mediated ROS to promote infection. *New Phytol.* 232:318-331.
- Chung, T., Wang, D., Kim, C.-S., Yadegari, R., and Larkins, B. A. 2009. Plant SMU-1 and SMU-2 homologues regulate pre-mRNA splicing and multiple aspects of development. *Plant Physiol.* 151:1498-1512.
- Contreras, M. P., Pai, H., Selvaraj, M., Toghiani, A., Lawson, D. M., Tumtas, Y., Duggan, C., Yuen, E. L. H., Stevenson, C. E. M., Harant, A., Maqbool, A., Wu, C.-H., Bozkurt, T. O., Kamoun, S., and Derevnina, L. 2023. Resurrection of plant disease resistance proteins via helper NLR bioengineering. *Sci Adv.* 9:eadg3861.
- Cotton, J. A., Lillley, C. J., Jones, L. M., Kikuchi, T., Reid, A. J., Thorpe, P., Tsai, I. J., Beasley, H., Blok, V., Cock, P. J. A., Eves-van den Akker, S., Holroyd, N., Hunt, M., Mantelin, S., Naghra, H., Pain, A., Palomares-Rius, J. E., Zarowiecki, M., Berriman, M., Jones, J. T., and Urwin, P. E. 2014. The genome and life-stage specific transcriptomes of *Globodera pallida* elucidate key aspects of plant parasitism by a cyst nematode. *Genome Biol.* 15:R43.
- Davies, J. M. 2014. Annexin-mediated calcium signalling in plants. *Plants* 3:128-140.
- Davies, L. J., Zhang, L., and Elling, A. A. 2015. The *Arabidopsis thaliana* papain-like cysteine protease RD21 interacts with a root-knot nematode effector protein. *Nematology* 17:655-666.
- de Almeida Engler, J., and Gheysen, G. 2013. Nematode-induced endoreduplication in plant host cells: Why and how? *Mol. Plant-Microbe Interact.* 26:17-24.
- de Almeida Engler, J., Rodiuc, N., Smertenko, A., and Abad, P. 2010. Plant actin cytoskeleton remodeling by plant parasitic nematodes. *Plant Signal. Behav.* 5:213-217.
- Derevnina, L., Contreras, M. P., Adachi, H., Upson, J., Vergara Cruces, A., Xie, R., Sklenar, J., Menke, F. L. H., Mugford, S. T., MacLean, D., Ma, W., Hogenhout, S. A., Govere, A., Maqbool, A., Wu, C.-H., and Kamoun, S. 2021. Plant pathogens convergently evolved to counteract redundant nodes of an NLR immune receptor network. *PLoS Biol.* 19:e3001136.
- Díaz-Granados, A., Petrescu, A.-J., Govere, A., and Smant, G. 2016. SPRYSEC effectors: A versatile protein-binding platform to disrupt plant innate immunity. *Front. Plant Sci.* 7:1575.
- Díaz-Granados, A., Sterken, M. G., Overmars, H., Ariaans, R., Holterman, M., Pokhare, S. S., Yuan, Y., Pomp, R., Finkers-Tomczak, A., Roosien, J., Sloopweg, E., Elashry, A., Grundler, F. M. W., Xiao, F., Govere, A., and Smant, G. 2020. The effector GpRbp-1 of *Globodera pallida* targets a nuclear HECT E3 ubiquitin ligase to modulate gene expression in the host. *Mol. Plant Pathol.* 21:66-82.
- Ding, X., Shields, J., Allen, R., and Hussey, R. S. 1998. A secretory cellulose-binding protein cDNA cloned from the root-knot nematode (*Meloidogyne incognita*). *Mol. Plant-Microbe Interact.* 11:952-959.
- Eves-van den Akker, S. 2021. Plant-nematode interactions. *Curr. Opin. Plant Biol.* 62:102035.
- FAO. 2019. International Year of Plant Health, 2020: Communication Guide. FAO, Rome, Italy.
- Fujimoto, T., Tomitaka, Y., Abe, H., Tsuda, S., Futai, K., and Mizukubo, T. 2011. Expression profile of jasmonic acid-induced genes and the induced resistance against the root-knot nematode (*Meloidogyne incognita*) in tomato plants (*Solanum lycopersicum*) after foliar treatment with methyl jasmonate. *J. Plant Physiol.* 168:1084-1097.
- Furniss, J. J., Grey, H., Wang, Z., Nomoto, M., Jackson, L., Tada, Y., and Spoel, S. H. 2018. Proteasome-associated HECT-type ubiquitin ligase activity is required for plant immunity. *PLoS Pathog.* 14:e1007447.
- Ganguly, A., Zhu, C., Chen, W., and Dixit, R. 2020. FRA1 kinesin modulates the lateral stability of cortical microtubules through cellulose synthase-microtubule uncoupling proteins. *Plant Cell* 32:2508-2524.
- Gao, B., Allen, R., Davis, E. L., Baum, T. J., and Hussey, R. S. 2004. Molecular characterisation and developmental expression of a cellulose-binding protein gene in the soybean cyst nematode *Heterodera glycines*. *Int. J. Parasitol.* 34:1377-1383.
- Gao, B., Allen, R., Maier, T., Davis, E. L., Baum, T. J., and Hussey, R. S. 2003. The parasitome of the phytonematode *Heterodera glycines*. *Mol. Plant-Microbe Interact.* 16:720-726.
- Gartner, U., Hein, I., Brown, L. H., Chen, X., Mantelin, S., Sharma, S. K., Dandurand, L.-M., Kuhl, J. C., Jones, J. T., Bryan, G. J., and Blok, V. C.

2021. Resisting potato cyst nematodes with resistance. *Front. Plant Sci.* 12:661194.
- Gheysen, G., and Mitchum, M. G. 2019. Phytoparasitic nematode control of plant hormone pathways. *Plant Physiol.* 179:1212-1226.
- Gleason, C., Leelarasamee, N., Meldau, D., and Feussner, I. 2016. OPDA has key role in regulating plant susceptibility to the root-knot nematode *Meloidogyne hapla* in *Arabidopsis*. *Front. Plant Sci.* 7:1565.
- Godinho Mendes, R. A., Basso, M. F., Fernandes de Araújo, J., Paes de Melo, B., Lima, R. N., Ribeiro, T. P., da Silva Mattos, V., Saliba Albuquerque, E. V., Grossi-de-Sa, M., Dessaune Tameirao, S. N., da Rocha Fragoso, R., Mattar da Silva, M. C., Vignols, F., Fernandez, D., and Grossi-de-Sa, M. F. 2021. Minc00344 and Mj-NULG1a effectors interact with GmHub10 protein to promote the soybean parasitism by *Meloidogyne incognita* and *M. javanica*. *Exp. Parasitol.* 229:108153.
- Grundler, F. M. W., Sobczak, M., and Golinowski, W. 1998. Formation of wall openings in root cells of *Arabidopsis thaliana* following infection by the plant-parasitic nematode *Heterodera schachtii*. *Eur. J. Plant Pathol.* 104:545-551.
- Guo, X., Chronis, D., De La Torre, C. M., Smeda, J., Wang, X., and Mitchum, M. G. 2015. Enhanced resistance to soybean cyst nematode *Heterodera glycines* in transgenic soybean by silencing putative CLE receptors. *Plant Biotechnol. J.* 13:801-810.
- Guo, Y., Ni, J., Denver, R., Wang, X., and Clark, S. E. 2011. Mechanisms of molecular mimicry of plant CLE peptide ligands by the parasitic nematode *Globodera rostochiensis*. *Plant Physiol.* 157:476-484.
- Hamamouch, N., Li, C., Hewezi, T., Baum, T. J., Mitchum, M. G., Hussey, R. S., Vodkin, L. O., and Davis, E. L. 2012. The interaction of the novel 30C02 cyst nematode effector protein with a plant  $\beta$ -1,3-endoglucanase may suppress host defence to promote parasitism. *J. Exp. Bot.* 63:3683-3695.
- Harris, W., Kim, S., Völz, R., and Lee, Y.-H. 2023. Nuclear effectors of plant pathogens: Distinct strategies to be one step ahead. *Mol. Plant Pathol.* 24:637-650.
- Heller, J., and Tudzynski, P. 2011. Reactive oxygen species in phytopathogenic fungi: Signaling, development, and disease. *Annu. Rev. Phytopathol.* 49:369-390.
- Hewezi, T., Howe, P., Maier, T. R., Hussey, R. S., Mitchum, M. G., Davis, E. L., and Baum, T. J. 2008. Cellulose binding protein from the parasitic nematode *Heterodera schachtii* interacts with *Arabidopsis* pectin methyltransferase: Cooperative cell wall modification during parasitism. *Plant Cell* 20:3080-3093.
- Hewezi, T., Howe, P. J., Maier, T. R., Hussey, R. S., Mitchum, M. G., Davis, E. L., and Baum, T. J. 2010. *Arabidopsis* spermidine synthase is targeted by an effector protein of the cyst nematode *Heterodera schachtii*. *Plant Physiol.* 152:968-984.
- Hewezi, T., Juvale, P. S., Piya, S., Maier, T. R., Rambani, A., Rice, J. H., Mitchum, M. G., Davis, E. L., Hussey, R. S., and Baum, T. J. 2015. The cyst nematode effector protein 10A07 targets and recruits host post-translational machinery to mediate its nuclear trafficking and to promote parasitism in *Arabidopsis*. *Plant Cell* 27:891-907.
- Holbein, J., Franke, R. B., Marhavý, P., Fujita, S., Górecka, M., Sobczak, M., Geldner, N., Schreiber, L., Grundler, F. M. W., and Siddique, S. 2019. Root endodermal barrier system contributes to defence against plant-parasitic cyst and root-knot nematodes. *Plant J.* 100:221-236.
- Huang, G., Allen, R., Davis, E. L., Baum, T. J., and Hussey, R. S. 2006. Engineering broad root-knot resistance in transgenic plants by RNAi silencing of a conserved and essential root-knot nematode parasitism gene. *Proc. Natl. Acad. Sci. U.S.A.* 103:14302-14306.
- Huang, L., Yuan, Y., Lewis, C., Kud, J., Kuhl, J. C., Caplan, A., Dandurand, L.-M., Zasada, I., and Xiao, F. 2023. NLR1 perceives a nematode ascaroside triggering immune signaling and resistance. *Curr. Biol.* 33:3992-3997.e3.
- Jauannet, M., Perfus-Barbeoch, L., Deleury, E., Magliano, M., Engler, G., Vieira, P., Danchin, E. G. J., Rocha, M. D., Coquillard, P., Abad, P., and Rosso, M.-N. 2012. A root-knot nematode-secreted protein is injected into giant cells and targeted to the nuclei. *New Phytol.* 194:924-931.
- Jones, J. T., Kumar, A., Pylypenko, L. A., Thirugnanasambandam, A., Castelli, L., Chapman, S., Cock, P. J. A., Grenier, E., Lilley, C. J., Phillips, M. S., and Blok, V. C. 2009. Identification and functional characterization of effectors in expressed sequence tags from various life cycle stages of the potato cyst nematode *Globodera pallida*. *Mol. Plant Pathol.* 10:815-828.
- Jones, J. T., and Robertson, W. M. 1997. Nematode Secretions. Pages 98-106 in: *Cellular and Molecular Aspects of Plant-Nematode Interactions*. C. Fenoll, F. M. W. Grundler, and S. A. Ohl, eds. Springer, Dordrecht, Netherlands.
- Kaloshian, I., and Teixeira, M. 2019. Advances in plant-nematode interactions with emphasis on the notorious nematode genus *Meloidogyne*. *Phytopathology* 109:1988-1996.
- Kiyohara, S., and Sawa, S. 2012. CLE signaling systems during plant development and nematode infection. *Plant Cell Physiol.* 53:1989-1999.
- Lee, C., Chronis, D., Kenning, C., Peret, B., Hewezi, T., Davis, E. L., Baum, T. J., Hussey, R., Bennett, M., and Mitchum, M. G. 2011. The novel cyst nematode effector protein 19C07 interacts with the *Arabidopsis* auxin influx transporter LAX3 to control feeding site development. *Plant Physiol.* 155:866-880.
- Lin, B., Zhuo, K., Chen, S., Hu, L., Sun, L., Wang, X., Zhang, L.-H., and Liao, J. 2016. A novel nematode effector suppresses plant immunity by activating host reactive oxygen species-scavenging system. *New Phytol.* 209:1159-1173.
- Lolle, S., Stevens, D., and Coaker, G. 2020. Plant NLR-triggered immunity: From receptor activation to downstream signaling. *Curr. Opin. Immunol.* 62:99-105.
- Lozano-Torres, J. L., Wilbers, R. H. P., Gawronski, P., Boshoven, J. C., Finkers-Tomczak, A., Cordewener, J. H. G., America, A. H. P., Overmars, H. A., Van 't Klooster, J. W., Baranowski, L., Sobczak, M., Ilyas, M., van der Hoorn, R. A. L., Schots, A., de Wit, P. J. G. M., Bakker, J., Goverse, A., and Smant, G. 2012. Dual disease resistance mediated by the immune receptor Cf-2 in tomato requires a common virulence target of a fungus and a nematode. *Proc. Natl. Acad. Sci. U.S.A.* 109:10119-10124.
- Lozano-Torres, J. L., Wilbers, R. H. P., Warmerdam, S., Finkers-Tomczak, A., Diaz-Granados, A., van Schaik, C. C., Helder, J., Bakker, J., Goverse, A., Schots, A., and Smant, G. 2014. Apoplastic venom allergen-like proteins of cyst nematodes modulate the activation of basal plant innate immunity by cell surface receptors. *PLoS Pathog.* 10:e1004569.
- Luo, S., Liu, S., Kong, L., Peng, H., Huang, W., Jian, H., and Peng, D. 2019. Two venom allergen-like proteins, HaVAP1 and HaVAP2, are involved in the parasitism of *Heterodera avenae*. *Mol. Plant Pathol.* 20:471-484.
- Manohar, M., Tenjo-Castano, F., Chen, S., Zhang, Y. K., Kumari, A., Williamson, V. M., Wang, X., Klessig, D. F., and Schroeder, F. C. 2020. Plant metabolism of nematode pheromones mediates plant-nematode interactions. *Nat. Commun.* 11:208.
- Mei, Y., Thorpe, P., Guzha, A., Haegeman, A., Blok, V. C., MacKenzie, K., Gheysen, G., Jones, J. T., and Mantelin, S. 2015. Only a small subset of the SPRY domain gene family in *Globodera pallida* is likely to encode effectors, two of which suppress host defences induced by the potato resistance gene *Gpa2*. *Nematology* 17:409-424.
- Mei, Y., Wright, K. M., Haegeman, A., Bauters, L., Diaz-Granados, A., Goverse, A., Gheysen, G., Jones, J. T., and Mantelin, S. 2018. The *Globodera pallida* SPRYSEC effector *GpSPRY-414-2* that suppresses plant defenses targets a regulatory component of the dynamic microtubule network. *Front. Plant Sci.* 9:1019.
- Mejias, J., Bazin, J., Truong, N.-M., Chen, Y., Marteu, N., Bouteiller, N., Sawa, S., Crespi, M. D., Vaucheret, H., Abad, P., Favery, B., and Quentin, M. 2021. The root-knot nematode effector MiEFF18 interacts with the plant core spliceosomal protein SmD1 required for giant cell formation. *New Phytol.* 229:3408-3423.
- Mejias, J., Truong, N. M., Abad, P., Favery, B., and Quentin, M. 2019. Plant proteins and processes targeted by parasitic nematode effectors. *Front. Plant Sci.* 10:970.
- Mendy, B., Wang'ombe, M. W., Radakovic, Z. S., Holbein, J., Ilyas, M., Chopra, D., Holton, N., Zipfel, C., Grundler, F. M. W., and Siddique, S. 2017. *Arabidopsis* leucine-rich repeat receptor-like kinase NLR1 is required for induction of innate immunity to parasitic nematodes. *PLoS Pathog.* 13:e1006284.
- Misas-Villamil, J. C., van der Hoorn, R. A. L., and Doehlemann, G. 2016. Papain-like cysteine proteases as hubs in plant immunity. *New Phytol.* 212:902-907.
- Mitchum, M. G., and Liu, X. 2022. Peptide effectors in phytonematode parasitism and beyond. *Annu. Rev. Phytopathol.* 60:97-119.
- Molloy, B., Baum, T., and Eves-van den Akker, S. 2023. Unlocking the development- and physiology-altering "effector toolbox" of plant-parasitic nematodes. *Trends Parasitol.* 39:732-738.
- Mukhtar, M. S., Carvunis, A.-R., Dreze, M., Epple, P., Steinbrenner, J., Moore, J., Tasan, M., Galli, M., Hao, T., Nishimura, M. T., Pevzner, S. J., Donovan, S. E., Ghamsari, L., Santhanam, B., Romero, V., Poulin, M. M., Gebreab, F., Gutierrez, B. J., Tam, S., Monachello, D., Boxem, M., Harbort, C. J., McDonald, N., Gai, L., Chen, H., He, Y., European Union Effectoromics Consortium, Vandenhoute, J., Roth, F. P., Hill, D. E., Ecker, J. R., Vidal, M., Beynon, J., Braun, P., Dangl, J. L. 2011. Independently evolved virulence effectors converge onto hubs in a plant immune system network. *Science* 333:596-601.

- Naalden, D., Haegeman, A., de Almeida-Engler, J., Birhane Eshetu, F., Bauters, L., and Gheysen, G. 2018. The *Meloidogyne graminicola* effector Mg16820 is secreted in the apoplast and cytoplasm to suppress plant host defense responses. *Mol. Plant Pathol.* 19:2416-2430.
- Patel, N., Hamamouch, N., Li, C., Hewezi, T., Hussey, R. S., Baum, T. J., Mitchum, M. G., and Davis, E. L. 2010. A nematode effector protein similar to annexins in host plants. *J. Exp. Bot.* 61:235-248.
- Perrine-Walker, F. 2019. Interactions of endoparasitic and ectoparasitic nematodes within the plant root system. *Funct. Plant Biol.* 46:295-303.
- Perry, R. N., Moens, M., and Starr, J. L., eds. 2009. *Root-Knot Nematodes*. CABI, Wallingford, U.K.
- Pogorelko, G., Juvale, P. S., Rutter, W. B., Hewezi, T., Hussey, R., Davis, E. L., Mitchum, M. G., and Baum, T. J. 2016. A cyst nematode effector binds to diverse plant proteins, increases nematode susceptibility and affects root morphology. *Mol. Plant Pathol.* 17:832-844.
- Pogorelko, G. V., Juvale, P. S., Rutter, W. B., Hütten, M., Maier, T. R., Hewezi, T., Paulus, J., van der Hoorn, R. A., Grundler, F. M., Siddique, S., Lionetti, V., Zabolina, O. A., and Baum, T. J. 2019. Re-targeting of a plant defense protease by a cyst nematode effector. *Plant J.* 98:1000-1014.
- Postma, W. J., Slootweg, E. J., Rehman, S., Finkers-Tomczak, A., Tytgat, T. O. G., van Gelderen, K., Lozano-Torres, J. L., Roosien, J., Pomp, R., van Schaik, C., Bakker, J., Govere, A., and Smant, G. 2012. The effector SPRYSEC-19 of *Globodera rostochiensis* suppresses CC-NB-LRR-mediated disease resistance in plants. *Plant Physiol.* 160:944-954.
- Przybylska, A., Wiczorek, P., and Obrepalska-Stepłowska, A. 2023. *Meloidogyne arenaria* candidate effector MaMsp4 interacts with maize (*Zea mays* L.) proteins involved in host defense response and cell wall modifications. *Plant Soil* 491:501-523.
- Quist, C. W., Smant, G., and Helder, J. 2015. Evolution of plant parasitism in the phylum Nematoda. *Annu. Rev. Phytopathol.* 53:289-310.
- Rehman, S., Postma, W., Tytgat, T., Prins, P., Qin, L., Overmars, H., Vossen, J. J., Spiridon, L.-N., Petrescu, A.-J., Govere, A., Bakker, J., and Smant, G. 2009. A secreted SPRY domain-containing protein (SPRYSEC) from the plant-parasitic nematode *Globodera rostochiensis* interacts with a CC-NB-LRR protein from a susceptible tomato. *Mol. Plant-Microbe Interact.* 22:330-340.
- Rosso, M.-N., Vieira, P., de Almeida-Engler, J., and Castagnone-Sereno, P. 2011. Proteins secreted by root-knot nematodes accumulate in the extracellular compartment during root infection. *Plant Signal. Behav.* 6:1232-1234.
- Ruan, Y., Halat, L. S., Khan, D., Jancowski, S., Ambrose, C., Belmonte, M. F., and Wasteneys, G. O. 2018. The microtubule-associated protein CLASP sustains cell proliferation through a brassinosteroid signaling negative feedback loop. *Curr. Biol.* 28:2718-2729.e5.
- Rutter, W. B., Franco, J., and Gleason, C. 2022. Rooting out the mechanisms of root-knot nematode-plant interactions. *Annu. Rev. Phytopathol.* 60:43-76.
- Sacco, M. A., Koropacka, K., Grenier, E., Jaubert, M. J., Blanchard, A., Govere, A., Smant, G., and Moffett, P. 2009. The cyst nematode SPRYSEC protein RBP-1 elicits Gpa2- and RanGAP2-dependent plant cell death. *PLoS Pathog.* 5:e1000564.
- Sato, K., Kadota, Y., and Shirasu, K. 2019. Plant immune responses to parasitic nematodes. *Front. Plant Sci.* 10:1165.
- Shinohara, H., and Matsubayashi, Y. 2015. Reevaluation of the CLV3-receptor interaction in the shoot apical meristem: Dissection of the CLV3 signaling pathway from a direct ligand-binding point of view. *Plant J.* 82:328-336.
- Siddique, S., Matera, C., Radakovic, Z. S., Hasan, M. S., Gutbrod, P., Rozanska, E., Sobczak, M., Torres, M. A., and Grundler, F. M. W. 2014. Parasitic worms stimulate host NADPH oxidases to produce reactive oxygen species that limit plant cell death and promote infection. *Sci. Signal.* 7:ra33.
- Song, H., Lin, B., Huang, Q., Sun, L., Chen, J., Hu, L., Zhuo, K., and Liao, J. 2021. The *Meloidogyne graminicola* effector MgMO289 targets a novel copper metallochaperone to suppress immunity in rice. *J. Exp. Bot.* 72:5638-5655.
- Stojilković, B., Xiang, H., Chen, Y., Bauters, L., Van de Put, H., Steppe, K., Liao, J., de Almeida Engler, J., and Gheysen, G. 2022. The nematode effector Mj-NEROS interacts with ISP influencing plastid ROS production to suppress plant immunity. *bioRxiv* 513376.
- Thorpe, P., Mantelin, S., Cock, P. J., Blok, V. C., Coke, M. C., Eves-van den Akker, S., Guzeva, E., Lilley, C. J., Smant, G., Reid, A. J., Wright, K. M., Urwin, P. E., and Jones, J. T. 2014. Genomic characterisation of the effector complement of the potato cyst nematode *Globodera pallida*. *BMC Genom.* 15:923.
- Truong, N. M., Chen, Y., Mejias, J., Soulé, S., Mulet, K., Jaouannet, M., Jaubert-Possamai, S., Sawa, S., Abad, P., Favery, B., and Quentin, M. 2021. The *Meloidogyne incognita* nuclear effector MiEFF1 interacts with *Arabidopsis* cytosolic glyceraldehyde-3-phosphate dehydrogenases to promote parasitism. *Front. Plant Sci.* 12:641480.
- Tytgat, T., Vanholme, B., De Meutter, J., Claeys, M., Couvreur, M., Vanhoutte, I., Gheysen, G., Van Criekinge, W., Borgonie, G., Coomans, A., and Gheysen, G. 2004. A new class of ubiquitin extension proteins secreted by the dorsal pharyngeal gland in plant parasitic cyst nematodes. *Mol. Plant-Microbe Interact.* 17:846-852.
- Van Damme, M., Huibers, R. P., Elberse, J., and Van den Ackerveken, G. 2008. *Arabidopsis* *DMR6* encodes a putative 2OG-Fe(II) oxygenase that is defense-associated but required for susceptibility to downy mildew. *Plant J.* 54:785-793.
- Van Der Vossen, E. A. G., Van Der Voort, J. N. A. M. R., Kanyuka, K., Bendahmane, A., Sandbrink, H., Baulcombe, D. C., Bakker, J., Stiekema, W. J., and Klein-Lankhorst, R. M. 2000. Homologues of a single resistance-gene cluster in potato confer resistance to distinct pathogens: A virus and a nematode. *Plant J.* 23:567-576.
- Verhoeven, A., Finkers-Tomczak, A., Prins, P., Valkenburg-van Raaij, D. R., van Schaik, C. C., Overmars, H., van Steenbrugge, J. J. M., Tacke, W., Varossieau, K., Slootweg, E. J., Kappers, I. F., Quentin, M., Govere, A., Sterken, M. G., and Smant, G. 2023. The root-knot nematode effector MiMSP32 targets host 12-oxophytodienoate reductase 2 to regulate plant susceptibility. *New Phytol.* 237:2360-2374.
- Verma, A., Lee, C., Morriss, S., Odu, F., Kenning, C., Rizzo, N., Spollen, W. G., Lin, M., McRae, A. G., Givan, S. A., Hewezi, T., Hussey, R., Davis, E. L., Baum, T. J., and Mitchum, M. G. 2018. The novel cyst nematode effector protein 30D08 targets host nuclear functions to alter gene expression in feeding sites. *New Phytol.* 219:697-713.
- Verma, A., Lin, M., Smith, D., Walker, J. C., Hewezi, T., Davis, E. L., Hussey, R. S., Baum, T. J., and Mitchum, M. G. 2022. A novel sugar beet cyst nematode effector 2D01 targets the *Arabidopsis* HAESA receptor-like kinase. *Mol. Plant Pathol.* 23:1765-1782.
- Vieira, P., and Gleason, C. 2019. Plant-parasitic nematode effectors - insights into their diversity and new tools for their identification. *Curr. Opin. Plant Biol.* 50:37-43.
- Vijayapalani, P., Hewezi, T., Pontvianne, F., and Baum, T. J. 2018. An effector from the cyst nematode *Heterodera schachtii* derepresses host rRNA genes by altering histone acetylation. *Plant Cell* 30:2795-2812.
- Williamson, V. M., and Kumar, A. 2006. Nematode resistance in plants: The battle underground. *Trends Genet.* 22:396-403.
- Wu, C.-H., Abd-El-Halim, A., Bozkurt, T. O., Belhaj, K., Terauchi, R., Vossen, J. H., and Kamoun, S. 2017. NLR network mediates immunity to diverse plant pathogens. *Proc. Natl. Acad. Sci. U.S.A.* 114:8113-8118.
- Xue, B., Hamamouch, N., Li, C., Huang, G., Hussey, R. S., Baum, T. J., and Davis, E. L. 2013. The *8D05* parasitism gene of *Meloidogyne incognita* is required for successful infection of host roots. *Phytopathology* 103:175-181.
- Yang, N., Yu, Q., Li, W., Chen, D., Jian, J., Zhang, H., Wang, G., Liu, S., Huang, W., Peng, H., Peng, D., Yang, J., Guo, X., and Kong, L. 2022. Functional characterization of the soybean cyst nematode effector SCN-27D09 using the model plant pathogenic fungus *Magnaporthe oryzae*-mediated delivery system. *Phytopathol. Res.* 4:33.
- Yang, S., Dai, Y., Chen, Y., Yang, J., Yang, D., Liu, Q., and Jian, H. 2019a. A novel G16B09-like effector from *Heterodera avenae* suppresses plant defenses and promotes parasitism. *Front. Plant Sci.* 10:66.
- Yang, S., Pan, L., Chen, Y., Yang, D., Liu, Q., and Jian, H. 2019b. *Heterodera avenae* GLAND5 effector interacts with pyruvate dehydrogenase subunit of plant to promote nematode parasitism. *Front. Microbiol.* 10:1241.
- Yeh, Y.-C. 2021. Functional characterization of 29D09 effectors from the potato cyst nematode *Globodera rostochiensis*. Master's thesis. Cornell University, Ithaca, NY.
- Yun, H. S., Sul, W. J., Chung, H. S., Lee, J.-H., and Kwon, C. 2023. Secretory membrane traffic in plant-microbe interactions. *New Phytol.* 237:53-59.
- Zhang, L., Davies, L. J., and Elling, A. A. 2015. A *Meloidogyne incognita* effector is imported into the nucleus and exhibits transcriptional activation activity in planta. *Mol. Plant Pathol.* 16:48-60.
- Zhang, R., Shi, P.-T., Zhou, M., Liu, H.-Z., Xu, X.-J., Liu, W.-T., and Chen, K.-M. 2023. Rapid alkalization factor: Function, regulation, and potential applications in agriculture. *Stress Biol.* 3:16.
- Zhang, X., Peng, H., Zhu, S., Xing, J., Li, X., Zhu, Z., Zheng, J., Wang, L., Wang, B., Chen, J., Ming, Z., Yao, K., Jian, J., Luan, S., Coleman-Derr, D., Liao, H., Peng, Y., Peng, D., and Yu, F. 2020. Nematode-encoded RALF peptide mimics facilitate parasitism of plants through the FERONIA receptor kinase. *Mol. Plant* 13:1434-1454.

- Zhao, J., Duan, Y., Kong, L., Huang, W., Peng, D., and Liu, S. 2022. Opposite beet cyst nematode infection phenotypes of transgenic *Arabidopsis* between overexpressing *GmSNAP18* and *AtSNAP2* and between overexpressing *GmSHMT08* and *AtSHMT4*. *Phytopathology* 112: 2383-2390.
- Zhao, J., Li, L., Liu, Q., Liu, P., Li, S., Yang, D., Chen, Y., Pagnotta, S., Favery, B., Abad, P., and Jian, H. 2019. A MIF-like effector suppresses plant immunity and facilitates nematode parasitism by interacting with plant annexins. *J. Exp. Bot.* 70:5943-5958.
- Zhao, J., and Liu, S. 2023. Beet cyst nematode HsSNARE1 interacts with both *AtSNAP2* and *AtPR1* and promotes disease in *Arabidopsis*. *J. Adv. Res.* 47:27-40.
- Zhao, J., Mejias, J., Quentin, M., Chen, Y., de Almeida-Engler, J., Mao, Z., Sun, Q., Liu, Q., Xie, B., Abad, P., Favery, B., and Jian, H. 2020. The root-knot nematode effector MiPDI1 targets a stress-associated protein (SAP) to establish disease in Solanaceae and *Arabidopsis*. *New Phytol.* 228:1417-1430.
- Zhao, J., Sun, Q., Quentin, M., Ling, J., Abad, P., Zhang, X., Li, Y., Yang, Y., Favery, B., Mao, Z., and Xie, B. 2021. A *Meloidogyne incognita* C-type lectin effector targets plant catalases to promote parasitism. *New Phytol.* 232:2124-2137.
- Zheng, Q., Putker, V., and Goverse, A. 2021. Molecular and cellular mechanisms involved in host-specific resistance to cyst nematodes in crops. *Front. Plant Sci.* 12:641582.



## Review

## Structural mechanisms of mitochondrial uncoupling protein 1 regulation in thermogenesis

Scott A. Jones <sup>1</sup>, Jonathan J. Ruprecht <sup>1</sup>, Paul G. Crichton <sup>2</sup>, and Edmund R.S. Kunji <sup>1,\*</sup>

In mitochondria, the oxidation of nutrients is coupled to ATP synthesis by the generation of a protonmotive force across the mitochondrial inner membrane. In mammalian brown adipose tissue (BAT), uncoupling protein 1 (UCP1, SLC25A7), a member of the SLC25 mitochondrial carrier family, dissipates the protonmotive force by facilitating the return of protons to the mitochondrial matrix. This process short-circuits the mitochondrion, generating heat for non-shivering thermogenesis. Recent cryo-electron microscopy (cryo-EM) structures of human UCP1 have provided new molecular insights into the inhibition and activation of thermogenesis. Here, we discuss these structures, describing how purine nucleotides lock UCP1 in a proton-impermeable conformation and rationalizing potential conformational changes of this carrier in response to fatty acid activators that enable proton leak for thermogenesis.

## Thermogenesis and UCP1

To support thermoregulation, many mammals, particularly newborns, use specialized fat deposits, known as **BAT** (see [Glossary](#)), to carry out **non-shivering thermogenesis** to protect against cold temperatures [1–3]. BAT is distinct from white adipose tissue due to its higher density of mitochondria, larger number of blood vessels, and smaller fat deposits. Heat generation by BAT occurs due to the presence of mitochondrial **UCP1**, also called thermogenin. The oxidation of breakdown products from fats and sugars is used by the complexes of the mitochondrial electron transport chain to generate a **protonmotive force**, which is used for ATP synthesis ([Figure 1A](#)). However, UCP1 short-circuits the mitochondrion by allowing protons to leak back from the intermembrane space to the mitochondrial matrix, dissipating the protonmotive force, which generates heat rather than chemical energy in the form of ATP ([Figure 1A](#)).

The activity of UCP1 is tightly controlled by regulatory ligands. Cytosolic **purine nucleotides** bind with high affinity to UCP1 to inhibit proton leak activity [4–6]. However, a cold environment induces sympathetic nerve activity and the adrenergic stimulation of BAT, leading to UCP1 activation. Intracellular signaling within brown adipocytes results in the phosphorylation of hormone-sensitive lipase and perilipin, lipolysis of triglyceride stores, and the release of free fatty acids [7]. These **fatty acid activators** directly interact with UCP1 [8], stimulating proton leak across the mitochondrial inner membrane by a mechanism that is as yet unresolved [9].

UCP1 activation is of therapeutic interest as it can help increase calorie expenditure and the clearance of triglycerides and glucose from the bloodstream [10,11]. Human brown fat correlates inversely with age-related obesity [12–14] and, when activated, improves insulin sensitivity and glucose homeostasis [15]. Recently, several atomic structures of UCP1 were determined by

## Highlights

Uncoupling protein 1 (UCP1, SLC25A7) is a mitochondrial carrier that facilitates thermogenesis in brown adipose tissue (BAT) by dissipating the protonmotive force and is a potential target for therapeutics against metabolic disease.

Recent cryo-electron microscopy (cryo-EM) structures of human UCP1 have provided molecular insights into pH-dependent purine nucleotide inhibition of thermogenesis in BAT.

UCP1 has the structural fold of members of the SLC25 mitochondrial carrier family, suggesting a related mechanism involving conformational changes, required to initiate proton conductance.

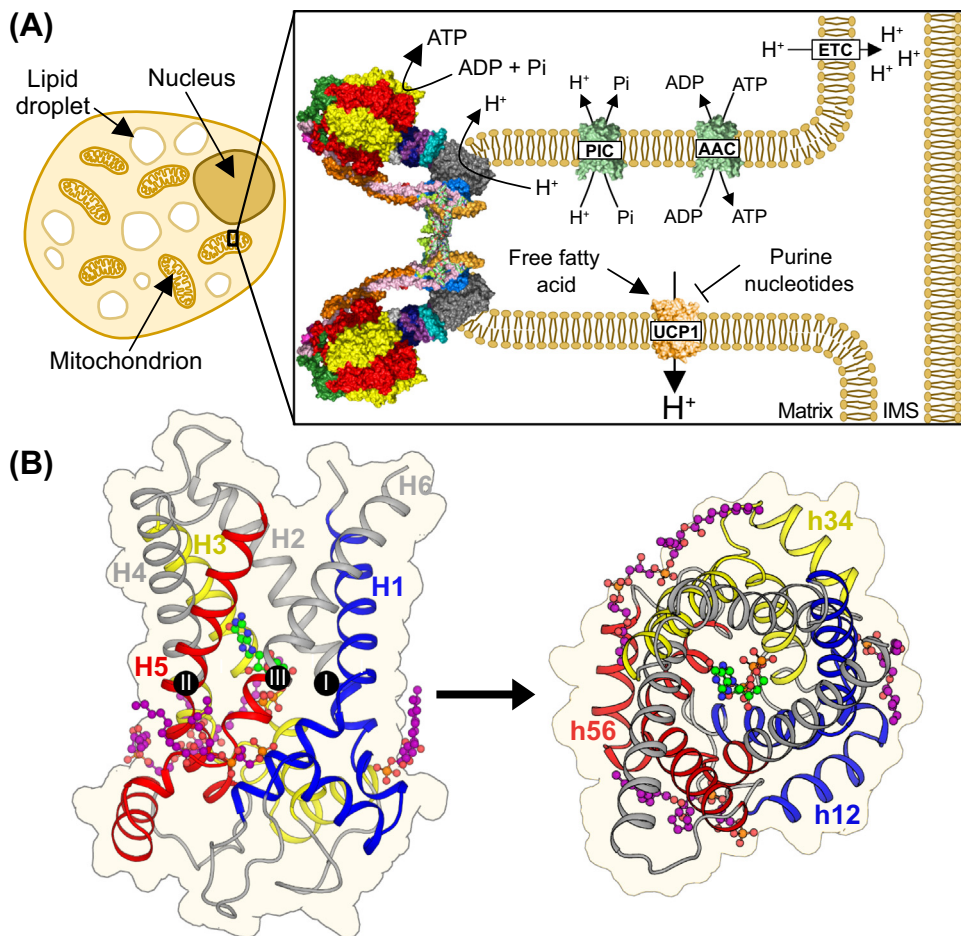
The chemical uncoupler 2,4-dinitrophenol (DNP), which binds weakly to UCP1, does not induce conformational changes and retains a proton-impermeable state.

Analysis of the structural features reveals potential molecular components of an activation mechanism of UCP1 in thermogenesis.

<sup>1</sup>MRC Mitochondrial Biology Unit, University of Cambridge, Cambridge Biomedical Campus, Keith Peters Building, Cambridge, CB2 0XY, UK

<sup>2</sup>Biomedical Research Centre, Norwich Medical School, University of East Anglia, Norwich, NR4 7TJ, UK

\*Correspondence: [ek@mrc-mbu.cam.ac.uk](mailto:ek@mrc-mbu.cam.ac.uk) (E.R.S. Kunji).



**Figure 1. Physiological role and structure of uncoupling protein 1 (UCP1).** (A) Brown adipose tissue cell (left). Mitochondrial membranes (right) showing a crista with ATP synthase at the tip with the phosphate carrier (PIC), ADP/ATP carrier (AAC), UCP1, and electron transport chain (ETC) within the mitochondrial inner membrane. UCP1 dissipates the protonmotive force across the inner mitochondrial membrane, which is activated by free fatty acids and inhibited by purine nucleotides. (B) Structure of UCP1 (PDB ID: 8G8W chain A) [16]. Lateral (left) and cytoplasmic (right) view with bound GTP (green) and three cardiolipin molecules (purple). Substrate binding in mitochondrial carriers occurs at the contact point residues (black circles), which are situated between the core (blue, yellow, and red helices) and gate elements (gray helices). The core and gate elements undergo substantial conformational changes during c- to m-state interconversion.

Trends in Biochemical Sciences

cryo-EM, revealing the molecular details of the protein for the first time [16,17]. In this review, we discuss the new molecular insights into the inhibition and activation mechanisms of UCP1. Understanding these is vital for the advancement of therapeutic strategies to stimulate thermogenesis for the treatment of obesity and associated metabolic diseases.

### UCP1 is a mitochondrial carrier

UCP1 (SLC25A7) is a member of the SLC25 **mitochondrial carrier** family, which is the largest transporter family in humans. Most mitochondrial carriers are involved in metabolite transport across the mitochondrial inner membrane [18]. SLC25 members are defined by three homologous domains, each comprising two transmembrane  $\alpha$ -helices separated by a loop and a matrix helix (Figure 1B). Studies of the mitochondrial ADP/ATP carrier, another member of the SLC25

### Glossary

**Brown adipose tissue (BAT):**

specialized fat tissue found in mammals that contains mitochondria with UCP1 expressed at high levels. When exposed to cold temperatures, mitochondria in BAT uncouple oxidative phosphorylation from ATP synthesis to generate heat and maintain body temperature.

**Chemical uncouplers:**

compounds that carry protons across biological membranes, dissipating the protonmotive force generated by the electron transport chain, which removes the driving force for ATP synthase.

**Contact points:**

amino acid residues on the even-numbered transmembrane helices of mitochondrial carriers, including UCP1, involved in substrate binding and coupling to conformational changes.

**Cytoplasmic network:**

salt bridge network on the cytoplasmic side of a carrier protein, which closes the central cavity and substrate-binding site to the intermembrane space when formed, holding the carrier in a m-state conformation. The network is highly conserved in mitochondrial carriers.

**Cytoplasmic state (c-state):**

conformation of mitochondrial carriers where the central substrate-binding site is accessible from the cytoplasm.

**2,4-Dinitrophenol (DNP):**

protonophore that disrupts the protonmotive force across the mitochondrial inner membrane by transporting protons.

**Fatty acid activators:**

free fatty acids generated from lipolysis of stored triglycerides in response to cold exposure, which are known activators of proton conductance by UCP1 and, therefore, thermogenesis in BAT.

**Matrix network:**

salt bridge network formed on the matrix side of UCP1, observed in all UCP1 structures, which holds the protein in a c-state closed to the mitochondrial matrix.

**Matrix state (m-state):**

conformation of mitochondrial carriers where the central substrate-binding site is open to the mitochondrial matrix.

**Mitochondrial carrier:**

group of transporters of the SLC25 family that transport metabolites and ions across the mitochondrial inner membrane.

**Nanobodies or sybodies:**

single-domain antibodies raised to bind folded protein epitopes specifically, which aide structural determination by cryo-EM or X-ray crystallography.

family, revealed the fundamental transport mechanism of mitochondrial carriers [19–21]. These carriers cycle between two conformations: one open to the intermembrane space, which is confluent with the cytoplasm (**cytoplasmic state; c-state**) [22,23] and another open to the mitochondrial matrix (**matrix state; m-state**) [24]. The conformational changes are coupled to the formation and disruption of two networks on either side of the carrier, which are part of the access gates to a central substrate-binding site [25]. The binding site comprises three **contact points** on the even-numbered transmembrane helices [26,27]. In the c-state, the residues of the matrix salt bridge network [22] and glutamine braces [23] interact, closing the central cavity to the matrix, whereas the residues of the cytoplasmic salt bridge network and tyrosine braces are not [24], opening the carrier to the intermembrane space. In the m-state, the configurations of the two networks are switched, opening the carrier to the mitochondrial matrix [24]. The conformational changes involve all three domains, each comprising a core and a gate element [24]. When the carrier changes from the c-state to the m-state, the three core elements rotate outward, opening the matrix side, whereas the three gate elements rotate inward, closing the cytoplasmic side of the carrier [24]. The same elements operate in reverse when the carrier moves from the m-state to the c-state.

### Cryo-EM structures of human UCP1

Determining protein structures by cryo-EM is challenging for proteins <100 kDa [28], but UCP1 is particularly difficult due to its small size (33 kDa), the presence of three pseudo-symmetrical domains, and its dynamic nature. Therefore, to facilitate its structure determination, fiducial markers were used to increase both the size and asymmetry of the protein. **Nanobodies** or synthetic nanobodies (**sybodies**) are attractive fiducials, because they are single-chain antibody fragments that can bind conformational epitopes and stabilize states [29,30]. Since nanobodies are small (15 kDa), they need to be enlarged, either by creating a fusion protein, such as a Pro-macrobody [31], or by creating protein complexes, such as legobodies [32]. Both strategies enabled the processing of particle images, taken by cryo-EM, to generate a density map of sufficient resolution for the building of atomic models of UCP1.

Structures of human UCP1 were solved in the GTP-bound state (PDB ID: 8G8W), ATP-bound state (PDB ID: 8HBW), and unliganded state (PDB ID: 8HBV), as well as in a **2,4-dinitrophenol (DNP; a proposed activator [33])**-bound state (PDB ID: 8J1N) [16,17]. The structures show that UCP1 has a typical fold for a SLC25 family member [20] and that the protein is monomeric with three bound cardiolipin molecules [16] (Figure 1B), as determined previously [34]. Notably, all solved structures used nanobodies or sybodies selected against the nucleotide-bound state of UCP1. While the amino acid numbering differs between the deposited structures, 8G8W follows the Uniprot sequence (P25874), which is also used here. These structures revealed the molecular properties of UCP1 for the first time, enabling analysis of the molecular mechanisms that control non-shivering thermogenesis in humans.

### Purine nucleotides lock UCP1 in a proton-impermeable configuration

The structures of UCP1 with bound GTP [16] and ATP [17] have revealed how purine nucleotides inhibit UCP1, preventing proton leak. The independently solved nucleotide-bound UCP1 structures are very similar [**root mean squared deviation (RMSD): 0.86 Å**] despite using distinct lipidic mimetic systems (detergent and nanodiscs) and fiducial markers that bind to different epitopes [16,17,29]. Purine nucleotides bind in the central cavity of UCP1 and lock the protein in a conformation that is open to the cytoplasm (c-state). Given the similarities between the chemical structures of ATP and GTP, most interactions are conserved. The binding arrangement accounts for their comparable affinities, although GTP has a slightly lower  $K_D$  in radiolabeled nucleotide-binding experiments compared with ATP [35]. The triphosphate and ribose groups

#### Non-shivering thermogenesis:

physiological process in which heat is generated through the dissipation of the protonmotive force produced by oxidative phosphorylation in BAT, rather than through shivering muscle contractions.

**Protonmotive force:** membrane potential and pH difference across inner mitochondrial membrane generated by the proton pumps of the electron transport chain, which is used for ATP synthesis and other cellular processes, such as transport.

**Purine nucleotides:** phosphorylated ribose bonded to purine bases, such as guanine and adenine, which comprise fused pyrimidine and imidazole rings.

#### Root mean squared deviation

**(RMSD):** measure of the structural similarity of two proteins, quantifying the average deviation of atom positions.

#### Uncoupling protein 1 (UCP1):

mitochondrial carrier (SLC25A7) responsible for dissipating the protonmotive force generated across the mitochondrial inner membrane during oxidative phosphorylation, short-circuiting the mitochondrion and causing thermogenesis.

of both purine nucleotides have similar interactions with UCP1 (Figure 2A,B). The triphosphate group forms electrostatic interactions with the positively charged central cavity comprising contact point residues R84, R277, and R183 (the ‘arginine triplet’) [26,27] and **matrix network** residues K38 and K138. In addition, Q85 forms a hydrogen bond with the  $\alpha$ -phosphate and R183 forms a hydrogen bond with the ribose group. R92 forms a cation- $\pi$  interaction with both adenine and guanine heterocyclic rings. The interactions of the nucleotides with UCP1 differ at the base: adenine forms hydrogen bonds with two asparagine residues, connecting N<sub>6</sub> with N188 and N<sub>3</sub> with N282 (Figure 2A); by contrast, the N<sub>2</sub> of guanine interacts electrostatically with E191, whereas N<sub>3</sub> and N<sub>2</sub> form hydrogen bonds to N188 (Figure 2B).

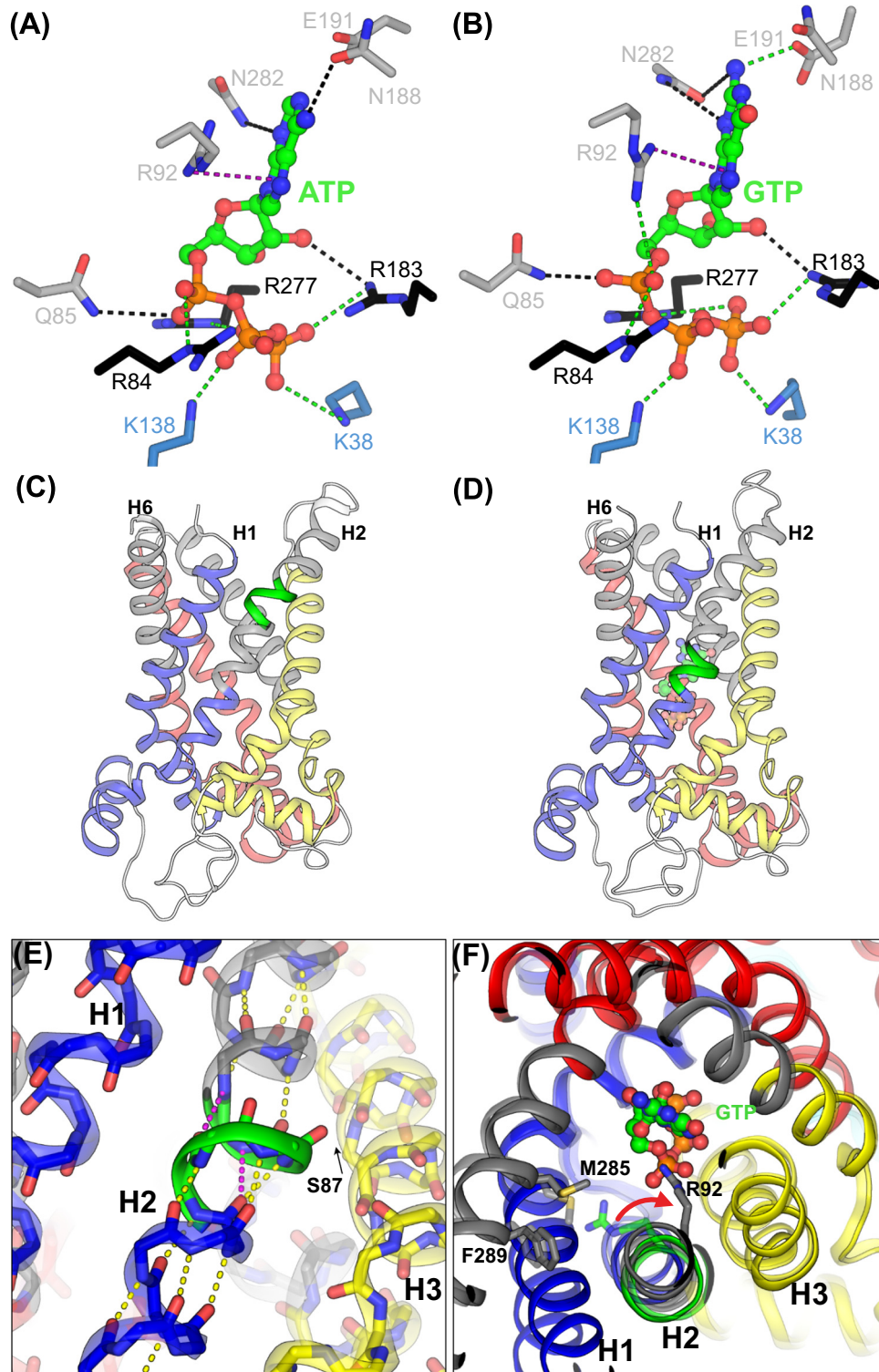
Notably, nucleotide binding to UCP1 [16,17] differs fundamentally from ATP binding to the mitochondrial ADP/ATP carrier (AAC) [36] in agreement with ATP being an inhibitor of UCP1 and a transport substrate of AAC [4,37]. In mitochondrial carriers, substrate binding primarily occurs at the contact points [26,27], which also act as hinges between the core and gate elements, which undergo large conformational changes between the c- and m-states [24,26,36]. In the mitochondrial ADP/ATP carrier, both the triphosphate and adenine groups interact with the contact point residues to disrupt one salt bridge network and form another, leading to substrate translocation [36]. In UCP1, purine nucleotides interact not only with the contact points, but also with the matrix network and gate elements, as well as with residues toward the cytoplasmic side of the cavity, explaining how they act as inhibitors, preventing conformational changes [16].

The UCP1 structures reveal an unusual state-dependent distortion on transmembrane helix H2. In the unliganded UCP1 structure, residues 92–96 form a  $\pi$ -bulge (Figure 2C), where the standard main-chain  $i, i+4$  hydrogen bonding is changed to  $i, i+5$ . By contrast, in the GTP- and ATP-inhibited structures, residues 92–96 form a standard  $\alpha$ -helix, but residues 85–89 adopt a  $\pi$ -bulge (Figure 2D). As a result, the main-chain carbonyl oxygen of S87 misses a hydrogen bond donor and rotates toward the membrane (Figure 2E). The switch from a  $\pi$ - to  $\alpha$ -helical arrangement in the region of R92 induces localized rotation of the main chain (Figure 2F). Consequently, R92 moves from interacting with M285 and F289 (both on helix H6) in the unliganded state to interacting directly with the  $\alpha$ -phosphate and purine base of the inhibitor in the bound state (Figure 2F). Consistent with this notion,  $\pi$ -bulges in helices are frequently located in functionally important sites, and transformations between  $\alpha$ - and  $\pi$ -helical arrangements have been proposed to help active sites accommodate ligands of different shapes and sizes [38]. Superimposition of the three domains of UCP1 for the nucleotide-bound and unliganded state shows movement of the gate element of transmembrane H2. A similar movement occurs during c- and m-state conversion of the ADP/ATP carrier and suggests that UCP1 is in an intermediate state when inhibited by nucleotides [16,24]. Otherwise, the unliganded state is very similar to the nucleotide-bound state and is also likely to be proton impermeable, having a closed matrix gate preventing the leak of protons.

### A plausible molecular explanation for the pH-dependence of purine nucleotide binding

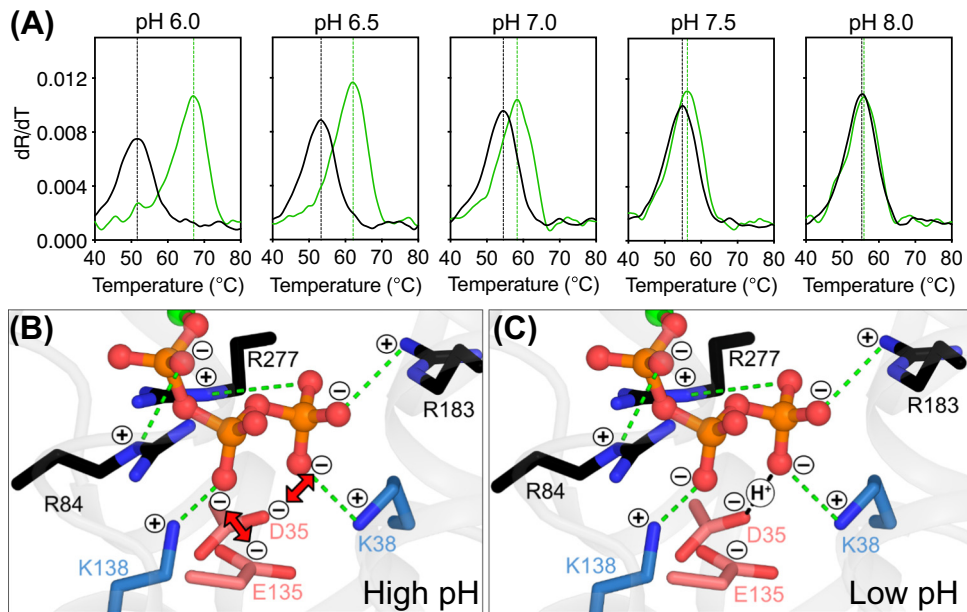
The binding affinity of purine nucleotides to UCP1 is pH dependent (Figure 3A) and decreases above pH 6.5 [16,35,37,39]. Previously, residue E191 was proposed to be responsible for the pH dependency of nucleotide binding [40–42], and was suggested to form an ionic bond with R92 [43]. Above pH 6.5, this interaction would block nucleotide binding [44], whereas, below pH 6.5, the ionic bond would be broken by protonation, opening the site for nucleotide binding [6,39,45]. This interaction has been observed in molecular dynamics simulations of a c-state model of UCP1, based on the ADP/ATP carrier [46]. However, E191 is not in bonding distance with R92 in the unliganded state structure (11.3 Å) [17], even when all alternative conformers are





Trends in Biochemical Sciences

(See figure legend at the bottom of the next page.)



Trends in Biochemical Sciences

**Figure 3. The pH dependency of purine nucleotide inhibition of uncoupling protein 1 (UCP1).** (A) The pH-dependent binding of GTP to purified human UCP1, as measured by a thermal stability shift assay [16]. The first derivatives of the unfolding curves are plotted in the absence (black) and presence of 1 mM GTP (green). The peak of the first derivative is the apparent melting temperature (T<sub>m</sub>). The change in apparent melting temperature (ΔT<sub>m</sub>) caused by GTP binding as a function of pH is indicated in green. (B) Nucleotide binding at high pH, showing the negatively charged residues of the matrix gate (in red) repelled (red arrows) by the negatively charged phosphate groups of the nucleotide (in orange and red). (C) Nucleotide binding at low pH, showing the negatively charged residues of the matrix gate (in red) interacting via a proton-mediated bond to the terminal phosphate group of the nucleotide (in orange and red).

considered. Indeed, in simulations, substantial distortion of the helical arrangement is required for this bond to occur [46]. Furthermore, E191 would not be protonated unless the local pH is <4, which does not match the pH-dependent profile of nucleotide binding. Importantly, E191 does interact with GTP [16], but not with ATP (Figure 2A) [17], even though the pH effect is observed for both nucleotides [6]. In agreement, the E191Q mutation weakens binding, but does not abolish the pH dependency of GTP binding [41], indicating that there must be an alternative explanation.

Both nucleotide-bound structures show that the negatively charged D35 and E135 of the matrix gate are positioned within bonding distance of the negatively charged triphosphate moiety of the purine nucleotides, which, at high pH, could cause repulsion (Figure 3B), but, at low pH, could form proton-mediated bonds (Figure 3C) [16]. The nucleotide-binding site has net negative overall

**Figure 2. Purine nucleotide binding to uncoupling protein 1 (UCP1).** (A) ATP (green) binding site (PDB ID: 8HBW chain A). The arginine triplet residues are shown in black; matrix network residues are shown in blue, and other residues involved in nucleotide binding are shown in gray. Salt bridge, hydrogen bond, and cation-π interactions are shown in green, black, and purple broken lines, respectively. (B) GTP (green) binding site (PDB ID: 8G8W chain A). Residues and interactions are depicted as in (A). (C) Unliganded UCP1 (PDB ID: 8HBV chain A) shown in cartoon representation, as in Figure 1B in the main text, but with the π-helix on H2 represented in green. (D) GTP-inhibited UCP1 (PDB ID: 8G8W chain A) shown as in (C). (E) Close-up view of the π-helix shown in (C), with mainchain atoms shown in stick representation. Conventional *i, i+4* hydrogen-bonds are shown as yellow broken lines, and π-helix *i, i+5* hydrogen-bonds are shown as magenta broken lines. (F) Comparison of uninhibited UCP1 (semitransparent cartoon) and GTP-inhibited UCP1 (opaque cartoon) showing the localized main-chain rotation (red arrow) and shift in the position of R92, which allows the residue to interact with the guanine base and α-phosphate of GTP.

charge and, thus, neutralizing this charge with protons will lead to stronger binding of the nucleotide. One proton could mediate bonding between the terminal phosphate and D35, in agreement with the pKa of 6.5 for purine nucleotides in solution (Figure 3C). However, in this conjugated system, another proton could also be drawn in to provide an additional proton-mediated interaction with E135. This bonding arrangement provides a plausible molecular mechanism for the pH dependency of purine nucleotide binding to UCP1.

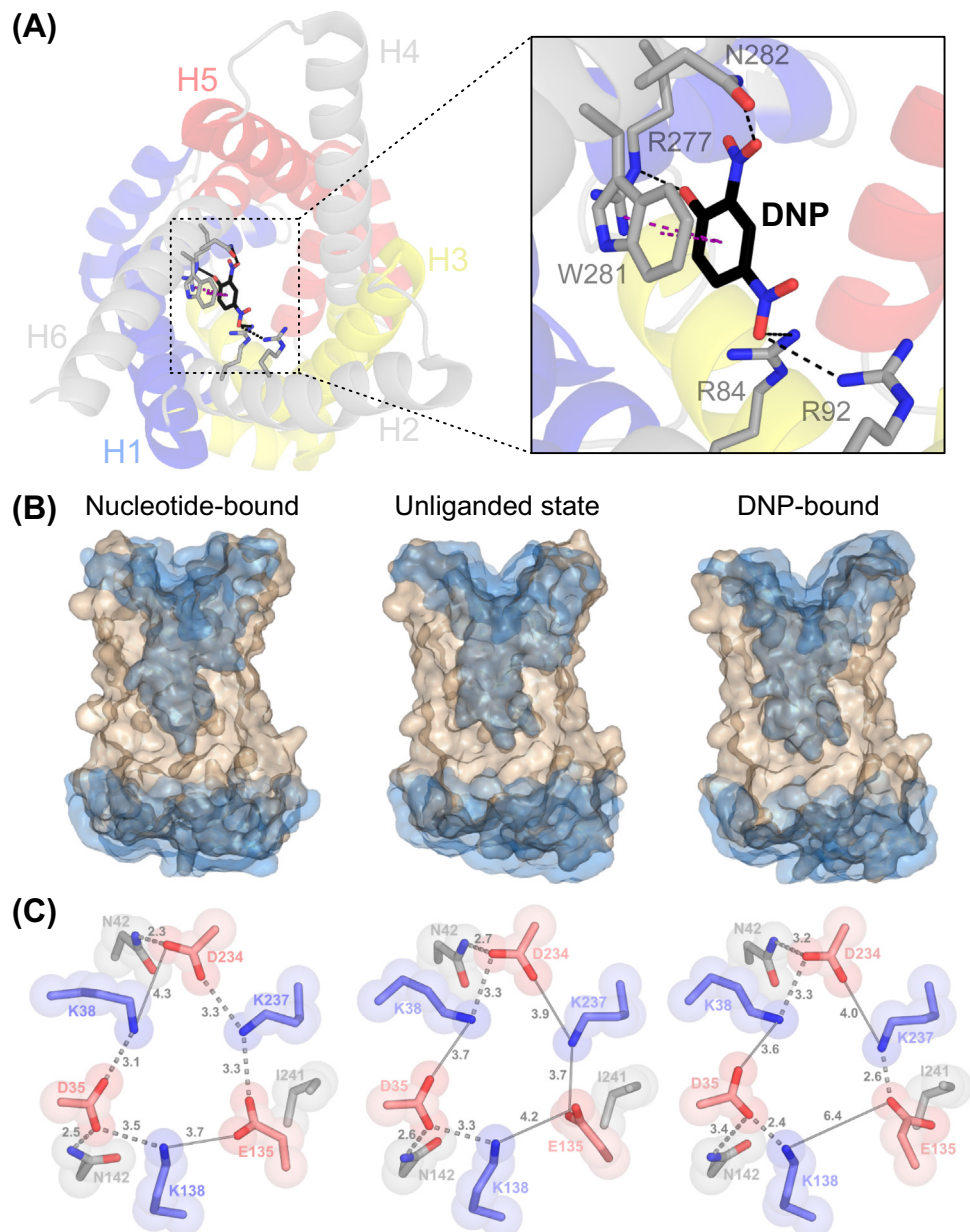
### The dinitrophenol-bound structure does not resolve the activation mechanism of UCP1

Weak acids can transport protons across biological membranes [47] and can function in mitochondria as **chemical uncouplers** by acting as protonophores, dissipating the protonmotive force. DNP is a well-known chemical uncoupler, which was used as weight-loss medication during the 1930s because of its ability to uncouple mitochondria [48]. However, the compound is highly toxic due to its unpredictable accumulation and distribution, which can lead to dangerous levels of oxidative phosphorylation inhibition [49]. Even though banned, DNP is still used as a weight loss agent, causing numerous deaths [49,50]. It was recently reported that UCP1 mediates DNP-induced proton leak across the mitochondrial inner membrane, which was abolished in mitochondria from UCP1-knockout animals [33].

Kang and Chen solved a UCP1 structure with bound DNP, used as an ‘activator’, by using the same sybody that was selected against the nucleotide-bound inhibited state [17]. The authors reported that DNP binds with extremely low potency, with an  $IC_{50}$  of 3.86 mM being measured for DNP-induced destabilization of UCP1; thus, a concentration of 10 mM DNP was utilized in the study. Despite using a high DNP concentration, only ~42% of 3D classes showed density for DNP during cryo-EM reconstruction, meaning that the  $K_D$  must be >10 mM [17]. In the structure, DNP binds to UCP1 via hydrogen bonds between R92 and R84 and  $N_4$ , a hydrogen bond between N282 and  $N_2$ , an ionic bond between R277 and the hydroxyl group, and a  $\pi$ - $\pi$  stacking interaction between W281 and the benzene ring of DNP (Figure 4A). Importantly, DNP does not induce any significant structural alterations, as the RMSD between unliganded and DNP-bound UCP1 is 0.25 Å. In addition, the central cavity is closed to the mitochondrial matrix with a large proton-impermeable layer similar to the unliganded and nucleotide-bound states (Figure 4B). As such, there is no structural evidence to support the hypothesis that DNP binding generates a proton-conducting state. Therefore, Kang and Chen propose that UCP1 is activated when bound to DNP by dynamic thermal motions, which transiently open the matrix gate to allow protons to pass through the central cavity [17]. However, there are several problems with this proposal.

First, DNP is chemically distinct and not a clear surrogate for fatty acids, the physiological activators, and may utilize a different mechanism. In addition, at the high concentrations used, DNP may bind to a UCP1 site that is unrelated entirely to the activation mechanism by either class of compounds. Moreover, the sybody was raised against the nucleotide-inhibited state and may restrict alternative activator-accessible conformations. Notably, the destabilization by DNP reported by the authors [17] only occurs at millimolar concentrations in a manner compatible with nonspecific denaturation, compared with the distinct ordered destabilization induced by micromolar concentrations of fatty acids [51]. As such, the DNP-binding site may not be relevant to fatty acid binding or the reported DNP activity [33].

Second, the reported structure of the DNP-bound UCP1 [17] is in a c-state through the formation of the matrix network and glutamine braces, comprising three ionic and two hydrogen bonds (Figure 4C) [16]. This matrix network arrangement is similar to those of the unliganded and purine nucleotide-bound states (Figure 4C), which do not conduct protons [3]. Therefore, it is unlikely



Trends in Biochemical Sciences

**Figure 4. The dinitrophenol-bound structure does not reveal an activated state.** (A) Cytoplasmic view of uncoupling protein 1 (UCP1) bound to 2,4-dinitrophenol (DNP). Core elements are colored by domain in blue, yellow, and red, and gate elements are in gray. Insert shows the DNP-binding site, black dashes represent hydrogen bonds and purple dashes represent  $\pi$ - $\pi$  stacking. (B) Cross-section of the surface view of nucleotide-bound (PDB ID: 8G8W chain A), unliganded (PDB ID: 8HBV chain A), and DNP-bound (PDB ID: 8J1N chain A) UCP1 states. The blue semitransparent surfaces reveal the internal cavities. (C) Matrix gate of the nucleotide-bound (left), unliganded (center), and DNP-bound (right) UCP1 states. Positively charged residues are in blue, negatively charged residues are in red, and neutral residues are in gray. Broken lines show polar interactions and thin gray lines represent distances between non-interacting residues.

that thermal energy is sufficient to break this bonding arrangement. In the related mitochondrial ADP/ATP carrier, which has a weaker interaction network, conformational changes only occur upon substrate binding [52], explaining the equimolar adenine nucleotide exchange [53]. Here,



the structure clearly shows that DNP binding does not provide sufficient energy input to break the networks. Any shift to a proton permeable state without breaking the network in an unrelated mechanism would likely require significant novel features in UCP1 distinct from other mitochondrial carriers, although none are apparent [54].

Third, many of the residues interacting with DNP in UCP1, for example, key residues W281, R92, and N282, are not conserved in the mitochondrial ADP/ATP carrier, which has also been proposed as a facilitator for DNP-induced uncoupling [33]. The UCP1 site is situated between H2 and H6 (Figure 4A), whereas the binding site of DNP proposed for the ADP/ATP carrier based on molecular dynamics simulations is between H4 and H5 [33].

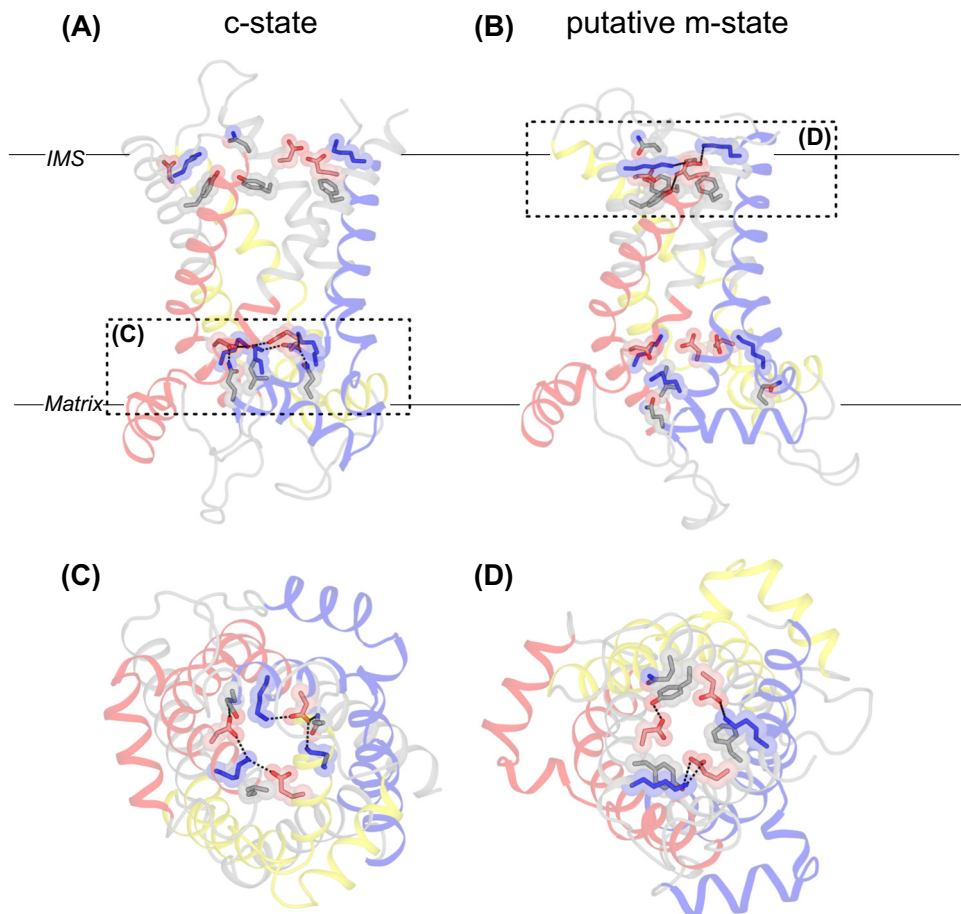
In conclusion, DNP binding does not provide evidence for any activation mechanism. The binding is extremely weak, does not induce conformational changes, has insufficient binding energy to release the matrix network and glutamine brace interactions, and is most likely unrelated to fatty acid binding.

### UCP1 has all the functional elements for a carrier-like mechanism

As described earlier, all structures of UCP1 are likely to be in a proton-impermeable c-state (Figure 4B), which could be due to the matrix network being stronger compared with the **cytoplasmic network**, increasing the probability of its formation [20]. Indeed, the matrix network is stronger than the cytoplasmic network due to the absence of a salt bridge between H2 and H4 (Figure 5D) [20]. Importantly, for all determined structures, nanobodies or sybodies were used that were selected against the nucleotide-bound state of UCP1 and, thus, could be responsible for maintaining a single conformation.

While some assert that the c-state is the only functional state of UCP1 [17], we believe that the observed changes in proteolysis sensitivity in response to fatty acids have provided evidence that conformational changes are part of the activation mechanism [55,56]. Furthermore, thermostability studies show that UCP1 shifts to a less stable population in response to activators, indicating that some conformational changes accompany activation [37,51,54]. In addition, UCP1 has both salt bridge networks and braces, which are key conserved features in the mechanism of mitochondrial carriers (Figure 5). Moreover, the interaction energy of the cytoplasmic network is even stronger than in many other mitochondrial carriers with an established transport function [20], indicating that formation of an m-state is a conserved property of UCP1 [20].

Vital to the conformational cycle of mitochondrial carriers is the rotation of the even-numbered helices across the surface of odd-numbered helices [23,24]. In addition, the transmembrane helices on the cytoplasmic side need to come close together in the m-state for the formation of the cytoplasmic network [16,24] (Figure 5B,D). Although the m-state structure has not been solved, all small residues in the helical interfaces that allow these movements to occur in other mitochondrial carriers are also present in UCP1, indicating that the ability to cycle between states is likely to be conserved [16]. Finally, the nucleotide-bound structures are in an intermediary rather than a pure c-state [16]. More specifically, the gate element of H2 is rotated into a position closer to that of an m-state, suggesting that UCP1 has retained its ability to convert to the m-state. Notably, the related UCP2 and UCP3 proteins share these functional features and have four-carbon metabolite transport activity [57–59], suggesting that a transport conformational cycle is retained in these related proteins to support their respective functions. Thus, all the sequence and structural data are consistent with the notion that UCP1 is a mitochondrial carrier capable of undergoing conformational changes, similar to those of a transport cycle.

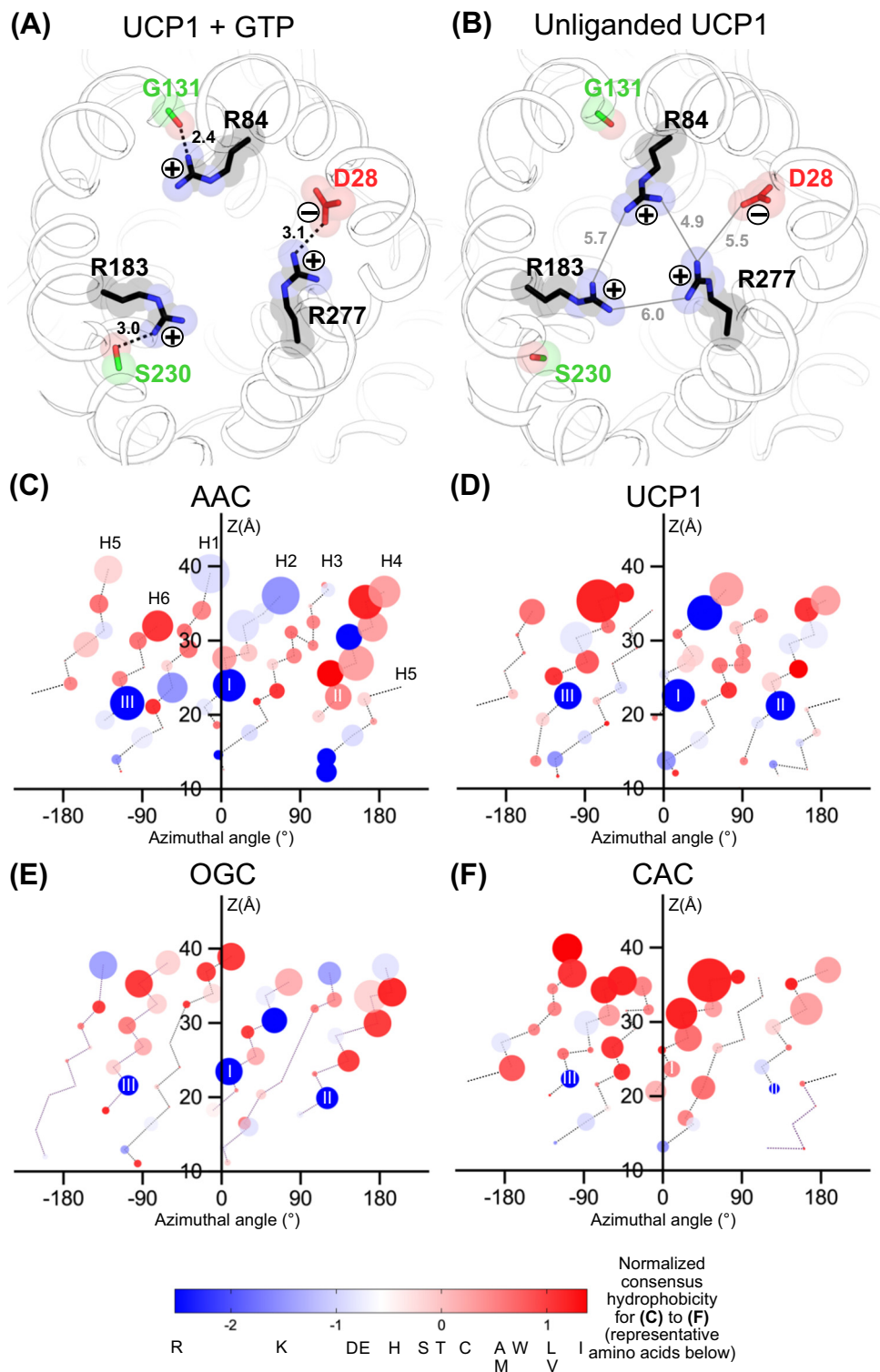


Trends in Biochemical Sciences

**Figure 5. Proposed conformational changes of uncoupling protein 1 (UCP1).** (A) Structure of UCP1 in cytoplasmic (c)-state (PDB ID: 8G8W chain A) and (B) a model of UCP1 in the matrix (m)-state based on the m-state of the mitochondrial ADP/ATP carrier (PDB ID: 6GCI, chain A) [24] (C) Matrix-side view of c-state structure of UCP1, showing the matrix gate residues, with interactions between residues shown in black broken lines. Positively charged residues are in blue, negatively charged residues are in red and neutral residues are in gray (D) Cytoplasmic view of the m-state model of UCP1, showing the cytoplasmic gate residues colored as in (C).

### Physiological activation of UCP1

The mechanism by which free fatty acids activate UCP1 remains undefined, but four main biochemical models have been proposed [54]. The competition model suggests that fatty acids act merely to remove purine nucleotides from an otherwise intrinsically active protein [60]. The cofactor model proposes that fatty acids become part of the proton conductance pathway, acting as a protonatable group [61]. Two alternative models propose that fatty acids are transport substrates of UCP1. In the shuttling model, fatty acids are transported by UCP1 but remain bound to the protein, chaperoning the proton across the membrane [62]. In the cycling model, fatty acid anions are exported by UCP1, after which the protonated fatty acids flip back across the membrane, independently from the protein, causing a net proton flux across the membrane [63,64]. However, the recently published structural studies have not revealed an activated state [16,17] and, thus, are unable to discriminate between the current proposed models of activation.



Trends in Biochemical Sciences

(See figure legend at the bottom of the next page.)

UCP1 likely evolved from the closely related dicarboxylate (DIC) and oxoglutarate (OGC) carriers, which transport their substrates without proton coupling [25]. A direct comparison with these proteins provides insights into how UCP1 might have evolved a proton-translocating function. All three proteins share a similar central substrate-binding site, containing the symmetric arginine triplet R84, R183, R277 at the substrate contact points and Q85 [25]. UCP1-specific adaptations in this site include the negatively charged D28, which has been shown to be important for proton conductance by UCP1 [25,65]. In the nucleotide-bound state, the arginine triplet interacts with other residues: R84 with the carbonyl backbone of G131; R183 with S230; and R277 with D28 (Figure 6A). However, in the nucleotide-free state, these interactions are broken and all three arginine residues point into the central cavity (Figure 6B). Importantly, D28, which was in an ionic bond with R277, is released and, thus, could become available for proton binding. The proposed substrates of UCP1, free fatty acids, are highly hydrophobic, especially when compared with the substrates of dicarboxylate and oxoglutarate carriers. The only mitochondrial carrier that transports substrates with a fatty acyl group is the carnitine/acylcarnitine carrier (CAC) [66]. Comparison of the central cavity residues with respect to hydrophobicity reveals that UCP1 is overall less hydrophobic compared with the carnitine/acylcarnitine carrier and more like the ADP/ATP carrier and dicarboxylate carrier (Figure 6). We believe that this observation suggests that the central cavity of UCP1 has not evolved extensively to enable fatty acid binding as a transport substrate, but does not exclude it. It is likely that UCP1 evolved from dicarboxylate carriers along with the other related UCP2 and UCP3 for a particular anion transport function, such as for small four-carbon metabolites [57–59]. A limited number of subsequent adaptations occurred for its role in thermogenesis [25], as a late evolutionary development, consistent with phylogenetic analysis [67,68]. Further studies of UCP1 activation, both structural and functional, will be required to determine how UCP1 activates thermogenesis.

### Concluding remarks

The structures of human UCP1 in the nucleotide-bound and free state represent a major advance in understanding the molecular mechanism of UCP1 inhibition in non-shivering thermogenesis. Comparison of both states shows how purine nucleotides inhibit UCP1 in a pH-dependent manner. While the molecular mechanism of activation of UCP1 remains unresolved, several potential elements have been uncovered, such as the release of the arginine triplet and D28, the presence of networks and braces, the conservation of small interhelical residues, and the partial conformational changes, indicative of state interconversion as part of the activation mechanism. However, many questions remain (see Outstanding questions), in particular how UCP1 is activated *in vivo* by free fatty acids. A molecular understanding of how fatty acids initiate proton leak across the mitochondrial inner membrane by UCP1 will provide an improved mechanistic understanding of non-shivering thermogenesis in humans and offer potential therapeutic strategies for addressing metabolic disorders.

**Figure 6. Substrate-binding site area of mitochondrial carriers.** (A) Contact point residues of uncoupling protein 1 (UCP1) inhibited with GTP (PDB ID: 8G8W chain A). Black broken lines represent hydrogen bonds with distance in angstrom labeled. (B) Contact point residues of UCP1 free state (PDB ID: 8HBV chain A). Gray lines show distance in angstrom between arginine residues. (C) C $\alpha$  atoms of residues lining the cavity of bovine ADP/ATP carrier (AAC) (PDB ID: 1OKC) [22], (D) human UCP1 (PDB ID: 8HBV chain A) [17], (E) *Tetrahymena thermophila* oxoglutarate carrier (OGC) (PDB ID: 7W5Z chain M2) [69], and (F) human carnitine/acylcarnitine carrier (CAC) AlphaFold homology model (PDB ID: AF-O43772-F1-model\_v4.pdb). (C–F) 2D projections of the cavity-lining residues, viewed from the center of the cavity. Each C $\alpha$  atom is represented as a sphere at its appropriate azimuthal angle and Z-height above a plane perpendicular to the longitudinal axis. The sphere area is proportional to the solvent-accessible surface area of the residue (calculated by AREAIMOL [70]), and the color represents the normalized consensus hydrophobicity of the residue [71], where red is the most hydrophobic and blue the most hydrophilic. Cavity-lining residues lie on transmembrane helices, indicated by the dotted lines, and labeled in (C). Contact point residues are labeled I, II, and III.

### Outstanding questions

Where in the structure of UCP1 do fatty acid activators bind and how do they activate? Do fatty acids act as transport substrates or cofactors?

What conformational changes are required to activate UCP1 proton conductance. Is there a specific activated state, or does UCP1 use a dynamic mechanism?

How do fatty acids overcome purine nucleotide inhibition during activation, when purine nucleotides are present at millimolar concentrations in the cytoplasm?

How do various purine nucleotide species in the cell compete to influence regulation of UCP1? Are there UCP1 isoform-specific adaptations?

Does UCP1 cycle between a c-state and m-state for proton leak activity, as observed for the metabolite transport cycle of other SLC25 family members, or are these states used for a similar, but so far unknown, metabolite transport function of UCP1?

What specific molecular adaptations has UCP1 acquired to facilitate proton leak and thermogenesis? Do these occur in other carriers and bestow a related proton conductance activity?

What biochemical and physiological function do other uncoupling proteins have (UCP2–5)? Understanding whether these related proteins also cause mitochondrial uncoupling may help to reveal the mechanism of activation.

Do purine nucleotides inhibit other mitochondrial carriers? Multiple conserved residues in the central cavity between UCP1 and related proteins suggest similar regulatory mechanisms.

What is the physiological activation mechanism of UCP1? Do other endogenous molecules, in addition to fatty acids, activate *in vivo*, related to other physiological stimuli, such as in response to eating?

Can we develop novel pharmacological agents that specifically target UCP1 *in vivo* and enhance its thermogenic



### Acknowledgments

We would like to acknowledge financial support from the UKRI Medical Research Council (MC\_UU\_00028/2) and Biotechnology and Biological Sciences Research Council (BB/S00940X/1).

### Declaration of interests

None declared by authors.

### References

- Lidell, M.E. (2019) Brown adipose tissue in human infants. *Handb. Exp. Pharmacol.* 251, 107–123
- Harms, M. and Seale, P. (2013) Brown and beige fat: development, function and therapeutic potential. *Nat. Med.* 19, 1252–1263
- Nicholls, D.G. (2021) Mitochondrial proton leaks and uncoupling proteins. *Biochim. Biophys. Acta Bioenerg.* 1862, 148428
- Rial, E. *et al.* (1983) Brown-adipose-tissue mitochondria: the regulation of the 32 000-Mr uncoupling protein by fatty acids and purine nucleotides. *Eur. J. Biochem.* 137, 197–203
- Heaton, G.M. *et al.* (1978) Brown-adipose-tissue mitochondria: photoaffinity labelling of the regulatory site of energy dissipation. *Eur. J. Biochem.* 82, 515–521
- Klingenberg, M. (1988) Nucleotide binding to uncoupling protein. Mechanism of control by protonation. *Biochemistry* 27, 781–791
- Shih, M.F. and Taberner, P.V. (1995) Selective activation of brown adipocyte hormone-sensitive lipase and cAMP production in the mouse by  $\beta$ 3-adrenoceptor agonists. *Biochem. Pharmacol.* 50, 601–608
- Locke, R.M. *et al.* (1982) Fatty acids as acute regulators of the proton conductance of hamster brown-fat mitochondria. *Eur. J. Biochem.* 129, 373–380
- Nicholls, D.G. (2017) The hunt for the molecular mechanism of brown fat thermogenesis. *Biochimie* 134, 9–18
- Stanford, K.I. *et al.* (2013) Brown adipose tissue regulates glucose homeostasis and insulin sensitivity. *J. Clin. Invest.* 123, 215–223
- Bartelt, A. *et al.* (2011) Brown adipose tissue activity controls triglyceride clearance. *Nat. Med.* 17, 200–206
- Nedergaard, J. *et al.* (2007) Unexpected evidence for active brown adipose tissue in adult humans. *Am. J. Physiol. Endocrinol. Metab.* 293, E444–E452
- Cypess, A.M. *et al.* (2009) Identification and importance of brown adipose tissue in adult humans. *N. Engl. J. Med.* 360, 1509
- Saito, M. *et al.* (2009) High incidence of metabolically active brown adipose tissue in healthy adult humans: effects of cold exposure and adiposity. *Diabetes* 58, 1526
- Chondronikola, M. *et al.* (2014) Brown adipose tissue improves whole-body glucose homeostasis and insulin sensitivity in humans. *Diabetes* 63, 4089–4099
- Jones, S.A. *et al.* (2023) Structural basis of purine nucleotide inhibition of human uncoupling protein 1. *Sci. Adv.* 9, ead425
- Kang, Y. and Chen, L. (2023) Structural basis for the binding of DNP and purine nucleotides onto UCP1. *Nature* 620, 226–231
- Kunjii, E.R.S. *et al.* (2020) The SLC25 carrier family: Important transport proteins in mitochondrial physiology and pathology. *Physiology* 35, 302–327
- Ruprecht, J.J. and Kunjii, E.R. (2019) Structural changes in the transport cycle of the mitochondrial ADP/ATP carrier. *Curr. Opin. Struct. Biol.* 57, 135–144
- Ruprecht, J.J. and Kunjii, E.R.S. (2020) The SLC25 mitochondrial carrier family: structure and mechanism. *Trends Biochem. Sci.* 45, 244–258
- Ruprecht, J.J. and Kunjii, E.R.S. (2021) Structural mechanism of transport of mitochondrial carriers. *Annu. Rev. Biochem.* 20, 535–558
- Peabay-Peyroula, E. *et al.* (2003) Structure of mitochondrial ADP/ATP carrier in complex with carboxyatractylolide. *Nature* 426, 39–44
- Ruprecht, J.J. *et al.* (2014) Structures of yeast mitochondrial ADP/ATP carriers support a domain-based alternating-access transport mechanism. *Proc. Natl. Acad. Sci. USA* 111, E426–E434
- Ruprecht, J.J. *et al.* (2019) The molecular mechanism of transport by the mitochondrial ADP/ATP carrier. *Cell* 176, 435–447
- Robinson, A.J. *et al.* (2008) The mechanism of transport by mitochondrial carriers based on analysis of symmetry. *Proc. Natl. Acad. Sci. USA* 105, 17766–17771
- Kunjii, E.R.S. and Robinson, A.J. (2006) The conserved substrate binding site of mitochondrial carriers. *Biochim. Biophys. Acta Bioenerg.* 1757, 1237–1248
- Robinson, A.J. and Kunjii, E.R.S. (2006) Mitochondrial carriers in the cytoplasmic state have a common substrate binding site. *Proc. Natl. Acad. Sci. USA* 103, 2617–2622
- Wentnick, K. *et al.* (2022) Putting on molecular weight: enabling cryo-EM structure determination of sub-100-kDa proteins. *Curr. Res. Struct. Biol.* 4, 332
- Uchański, T. *et al.* (2020) Nanobodies to study protein conformational states. *Curr. Opin. Struct. Biol.* 60, 117–123
- Pardon, E. *et al.* (2014) A general protocol for the generation of nanobodies for structural biology. *Nat. Protoc.* 9, 674–693
- Botte, M. *et al.* (2022) Cryo-EM structures of a LptDE transporter in complex with Pro-macrobodies offer insight into lipopolysaccharide translocation. *Nat. Commun.* 13, 1826
- Wu, X. and Rapoport, T.A. (2021) Cryo-EM structure determination of small proteins by nanobody-binding scaffolds (Legobodies). *Proc. Natl. Acad. Sci. USA* 118, e2115001118
- Bertholet, A.M. *et al.* (2022) Mitochondrial uncouplers induce proton leak by activating AAC and UCP1. *Nature* 606, 180–187
- Lee, Y. *et al.* (2015) Uncoupling protein 1 binds one nucleotide per monomer and is stabilized by tightly bound cardiolipin. *Proc. Natl. Acad. Sci. USA* 112, 6973–6978
- Lin, C.S. and Klingenberg, M. (1982) Characteristics of the isolated purine nucleotide binding protein from brown fat mitochondria. *Biochemistry* 21, 2950–2956
- Mavidou, V. *et al.* (2022) Substrate binding in the mitochondrial ADP/ATP carrier is a step-wise process guiding the structural changes in the transport cycle. *Nat. Commun.* 13, 3585
- Crichton, P.G. *et al.* (2015) Trends in thermostability provide information on the nature of substrate, inhibitor, and lipid interactions with mitochondrial carriers. *J. Biol. Chem.* 290, 8206–8217
- Kumar, P. *et al.* (2015) Dissecting p-helices: sequence, structure and function. *FEBS J.* 282, 4415–4432
- Echtay, K.S. *et al.* (2018) Uncoupling proteins: Martin Klingenberg's contributions for 40 years. *Arch. Biochem. Biophys.* 657, 41–55
- Winkler, E. *et al.* (1997) Identification of the pH sensor for nucleotide binding in the uncoupling protein from brown adipose tissue. *Biochemistry* 36, 148–155
- Echtay, K.S. *et al.* (1997) Mutagenesis of the uncoupling protein of brown adipose tissue. Neutralization of E190 largely abolishes pH control of nucleotide binding. *Biochemistry* 36, 8253–8260
- Shuvalov, V.A. *et al.* (1988) Nucleotide binding to uncoupling protein. Mechanism of control by protonation? *Proc. Natl. Acad. Sci. USA* 27, 3830–3832
- Echtay, K.S. *et al.* (2001) Role of intrahelical arginine residues in functional properties of uncoupling protein (UCP1). *Biochemistry* 40, 5243–5248
- Klingenberg, M. (2017) UCP1 - a sophisticated energy valve. *Biochimie* 134, 19–27
- Klingenberg, M. and Huang, S.-G. (1999) Structure and function of the uncoupling protein from brown adipose tissue. *Biochim. Biophys. Acta* 1415, 271–296
- Gagelin, A. *et al.* (2023) Molecular determinants of inhibition of UCP1-mediated respiratory uncoupling. *Nat. Commun.* 14, 2594
- McLaughlin, S.G.A. and Dilger, J.P. (1980) Transport of protons across membranes by weak acids. *Physiol. Rev.* 60, 825–863

activity? Are there molecules that can activate both brown adipocyte proliferation and UCP1 directly in the absence of physiological stimuli?

48. Colman, E. (2007) Dinitrophenol and obesity: an early twentieth-century regulatory dilemma. *Regul. Toxicol. Pharmacol.* 48, 115–117
49. Grundlingh, J. *et al.* (2011) 2,4-Dinitrophenol (DNP): a weight loss agent with significant acute toxicity and risk of death. *J. Med. Toxicol.* 7, 205
50. Siegmüller, C. and Narasimhaiah, R. (2010) Fatal 2,4-dinitrophenol poisoning... coming to a hospital near you. *Emerg. Med. J.* 27, 639–640
51. Cavaleri, R. *et al.* (2022) Activating ligands of Uncoupling protein 1 identified by rapid membrane protein thermostability shift analysis. *Mol. Metab.* 62, 101526
52. Springett, R. *et al.* (2017) Modelling the free energy profile of the mitochondrial ADP/ATP carrier. *Biochim. Biophys. Acta Bioenerg.* 1858, 906–914
53. Pfaff, E. *et al.* (1965) Unspecific permeation and specific exchange of adenine nucleotides in liver mitochondria. *Biochim. Biophys. Acta Gen. Subj.* 104, 312–315
54. Crichton, P.G. *et al.* (2017) The molecular features of uncoupling protein 1 support a conventional mitochondrial carrier-like mechanism. *Biochimie* 134, 35–50
55. Huang, S.G. (2003) Limited proteolysis reveals conformational changes in uncoupling protein-1 from brown adipose tissue mitochondria. *Arch. Biochem. Biophys.* 420, 40–45
56. Divakaruni, A.S. *et al.* (2012) Fatty acids change the conformation of uncoupling protein 1 (UCP1). *J. Biol. Chem.* 287, 36845–36853
57. Vozza, A. *et al.* (2014) UCP2 transports C4 metabolites out of mitochondria, regulating glucose and glutamine oxidation. *Proc. Natl. Acad. Sci. USA* 111, 960–965
58. Kreiter, J. *et al.* (2023) Uncoupling protein 3 catalyzes the exchange of C4 metabolites similar to UCP2. *Biomolecules* 14, 21
59. De Leonardis, F. *et al.* (2024) Human mitochondrial uncoupling protein 3 functions as a metabolite transporter. *FEBS Lett.* 598, 338–346
60. Shabalina, I.G. *et al.* (2004) Native UCP1 displays simple competitive kinetics between the regulators purine nucleotides and fatty acids. *J. Biol. Chem.* 279, 38236–38248
61. Winkler, E. and Klingenberg, M. (1994) Effect of fatty acids on H<sup>+</sup> transport activity of the reconstituted uncoupling protein. *J. Biol. Chem.* 269, 2508–2515
62. Fedorenko, A. *et al.* (2012) Mechanism of fatty-acid-dependent UCP1 uncoupling in brown fat mitochondria. *Cell* 151, 400–413
63. Garlid, K.D. *et al.* (1998) The mechanism of proton transport mediated by mitochondrial uncoupling proteins. *FEBS Lett.* 438, 10–14
64. Garlid, K.D. *et al.* (1996) On the mechanism of fatty acid-induced proton transport by mitochondrial uncoupling protein. *J. Biol. Chem.* 271, 2615–2620
65. Echtay, K.S. *et al.* (2000) Site-directed mutagenesis identifies residues in uncoupling protein (UCP1) involved in three different functions. *Biochemistry* 39, 3311–3317
66. Tonazzi, A. *et al.* (2021) The mitochondrial carnitine acyl-carnitine carrier (SLC25A20): molecular mechanisms of transport, role in redox sensing and interaction with drugs. *Biomolecules* 11, 521
67. Hughes, D.A. *et al.* (2009) Molecular evolution of UCP1 and the evolutionary history of mammalian non-shivering thermogenesis. *BMC Evol. Biol.* 9, 4
68. Gaudry, M.J. *et al.* (2019) Molecular evolution of thermogenic uncoupling protein 1 and implications for medical intervention of human disease. *Mol. Asp. Med.* 68, 6–17
69. Zhou, L. *et al.* (2022) Structures of Tetrahymena's respiratory chain reveal the diversity of eukaryotic core metabolism. *Science* 376, 831–839
70. Winn, M.D. *et al.* (2011) Biological crystallography overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* 67, 235–242
71. Eisenberg, D. *et al.* (1984) Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J. Mol. Biol.* 179, 125–142