C9orf72-mediated ALS and FTD: multiple pathways to disease

Rubika Balendra[1,2] and Adrian M. Isaacs[1,3*]

[1]Department of Neurodegenerative Disease, UCL Institute of Neurology, London, UK.

[2]Department of Genetics, Evolution and Environment, Institute of Healthy Ageing, University College London, London, UK.

[3]UK Dementia Research Institute at UCL, UCL Institute of Neurology, London, UK.

*email: a.isaacs@ucl.ac.uk

Abstract | The discovery that repeat expansions in the *C9orf72* gene are a frequent cause of amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) has revolutionized our understanding of these diseases. Substantial headway has been made in characterizing *C9orf72*-mediated disease and unravelling its underlying aetiopathogenesis. Three main disease mechanisms have been proposed: loss of function of the C9orf72 protein, toxic gain of function from *C9orf72* repeat RNA or from dipeptide repeat proteins produced by repeat-associated non-ATG translation. Several downstream processes across a range of cellular functions have also been implicated. In this article, we review the pathological and mechanistic features of *C9orf72*-associated FTD and ALS (collectively termed C9FTD/ALS), the model systems used to study these conditions, and the probable initiators of downstream disease mechanisms. We suggest that a combination of upstream mechanisms involving both loss and gain of function, and downstream cellular pathways involving both cell-autonomous and non-cell-autonomous effects, contribute to disease progression.

Introduction

Amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) are devastating and fatal neurodegenerative diseases. In common with other neurodegenerative diseases, progress towards finding disease-modifying therapies in ALS and FTD has been slow, in large part owing to an

incomplete understanding of disease aetiopathogenesis. In 2011, ground-breaking progress was made with the discovery that a hexanucleotide GGGGCC repeat expansion in the *C9orf72* gene (Fig. 1) is the most frequent genetic cause of both diseases in Europe and North America[1,2]. Interestingly, C9FTD/ALS — the collective term for *C9orf72*-associated diseases with clinical features of FTD, ALS or both —is extremely rare in Asia and the Middle East[3,4], indicating a different genetic architecture underlying FTD and ALS in these populations. The age of onset of C9FTD/ALS ranges from 27–83 years of age[3,4], and the disease duration ranges from 1–22 years[3,4]. *C9orf72* repeat expansions have also been identified as a rare cause of other neurodegenerative diseases[4], including Parkinson disease, progressive supranuclear palsy, ataxia, corticobasal syndrome, Huntington disease-like syndrome, Creutzfeldt–Jakob disease and Alzheimer disease.

The vast majority (>95%) of neurologically healthy individuals have ≤11 hexanucleotide repeats in the *C9orf72* gene[5]. The pathological repeat-length threshold has not been clearly defined; an arbitrary cut-off of 30 repeats is used in most studies, but larger expansions ranging from hundreds to thousands of repeats are most commonly observed in patients with C9FTD/ALS[5-10]. Importantly, the discovery of C9FTD/ALS has heightened the realization that ALS and FTD are intimately linked on a clinical, genetic, pathological and mechanistic spectrum.

The expanded GGGGCC repeats are bidirectionally transcribed into repetitive RNA, which forms sense and antisense RNA foci[11-15]. Remarkably, despite being within a non-coding region of *C9orf72*, these repetitive RNAs can be translated in every reading frame to form five different dipeptide repeat proteins (DPRs) — poly-GA, poly-GP poly-GR, poly-PA and poly-PR — via a non-canonical mechanism known as repeat-associated non-ATG (RAN) translation (Fig. 2)[14-18]. Although *C9orf72* mutations are a relatively recent discovery, progress in understanding their pathogenenic effects has been rapid. Three competing but non-exclusive mechanisms have arisen: loss of function of C9orf72 protein, and toxic gain of function from sense and antisense *C9orf72* repeat RNA or from DPRs. These mechanisms are all likely to contribute to disease to some extent, but it is crucial to determine their relative importance at various disease stages, so as to inform therapeutic strategies.

The differential involvement of these mechanisms might also explain clinical, pathological and prognostic heterogeneity that is observed in patients with C9FTD/ALS.

In this article, we review the current understanding of the mechanisms underlying C9FTD/ALS, and the questions that remain unanswered. The clinical and genetic aspects of C9FTD/ALS have been extensively reviewed elsewhere[19,20] and will not be discussed in detail here.

Repeat size and somatic mosaicism

Defining the minimum number of hexanucleotide repeats in *C9orf72* that cause disease would be invaluable for genetic counselling and to guide disease modelling, but this issue is currently unresolved. Neurologically healthy individuals and patients with ALS or FTD can all have 20–30 repeats, so whether repeat lengths within this range can drive disease is unclear. For example, a screen of control post-mortem brains identified an individual with 30 repeats, sparse RNA foci and DPR inclusions, but no neurological symptoms up until death at 84 years of age[21].

One confounding factor for accurate repeat sizing is somatic instability of the mutation. Individuals have been identified with large expansions within the CNS but an intermediate repeat length in DNA extracted from blood[22-24]. These studies show that >50 repeats in blood-derived DNA can be associated with large CNS expansions, suggesting that 50 repeats is a useful cut-off when analysing blood DNA. However, the current data do not support a single precise cut-off; for instance, 70 repeats were insufficient to cause disease in an 89-year-old whose children inherited much larger expansions and went on to develop C9FTD/ALS[25]. Furthermore, defining a cut-off for blood-derived DNA does not help answer the key question of the minimum repeat size required in the CNS to cause disease.

A further confounder is that repeat size varies between different brain regions[6,10]. This so-called somatic mosaicism within the CNS might explain some of the clinical heterogeneity that is observed between patients, but is currently not well understood.

Pathological features of C9FTD/ALS

TDP-43 inclusions

The overwhelming majority of ALS cases and approximately 50% of FTD cases are characterized by inclusions consisting of the RNA-binding protein TDP-43 (TAR DNA-binding protein 43) in neurons and glia[26,27]. The fact that *C9orf72* mutations can lead to TDP-43 inclusions (Fig. 3) in both ALS and FTD implies a final common pathway in these diseases[28]. This pathology is evident in various brain regions, including the frontal, temporal and primary motor cortices, hippocampus, basal ganglia, amygdala, thalamus and midbrain[29-33]. Patients with *C9orf72* mutations who have a predominant ALS syndrome can still exhibit extramotor pathological features consistent with FTD and, conversely, those with a predominant FTD syndrome can show pathology in the motor system[31].

RNA foci

Sense and antisense RNA foci comprising *C9orf72* repeat RNA are widely distributed across the CNS in patients with C9FTD/ALS (Fig. 3). These foci are found predominantly within neuronal nuclei in the frontal and motor cortices, hippocampus, cerebellum and in the spinal cord, in motor neurons and occasionally in interneurons, and sporadically in the cytoplasm[11,13-15,34]. Less frequently, foci are detected in glia (astrocytes, microglia, astrocytes and oligodendrocytes [11,13,15]. Sense RNA foci occur in ~37% and antisense foci in ~26% of neurons in the frontal cortex, respectively[11,34], and they co-occur in ~14% of frontal cortex neurons[11].

DPR pathology

DPR inclusions are p62-positive and TDP-43-negative and can consist of more than one DPR (Fig. 3)[14,15,17,18,35]. Several studies have examined DPRs in patients with *C9orf72* repeat expansions, using immunohistochemistry[14-18,29,35-42], immunoblotting[14,16,17,38,41] or immunoassays[43], and have reached similar conclusions. DPRs most commonly form neuronal cytoplasmic inclusions[15-17,29], but can also exist as neuritic inclusions[29,44] or as 'pre-inclusions'[15,29], which appear as diffuse cytoplasmic staining.

Neuronal intranuclear inclusions are sometimes observed[15-17,29], and are occasionally paranucleolar[36]. Sense-derived poly-GA is the most frequent form of DPR [16,18,38-41], followed by poly-GP then poly-GR. The antisense-derived DPRs poly-PA and poly-PR are the least frequent forms[15,18,35,36,38-40]. Although sense-derived DPRs seem to be more prevalent than antisense-derived DPRs in patients with *C9orf72* repeat expansions, both sense-derived and antisense-derived poly-GP are detected in hippocampal neurons in these individuals, with strand-specific antibodies suggesting a preponderance of antisense-derived poly-GP[14]. Staining of cortical tissue with antibodies specific to the carboxy-terminal region of the translated DPRs revealed that translation occurs beyond the 3′ end of the repeats in the cortex in *C9orf72* repeat expansion carriers[14,18].

DPR pathology is most prominent in the cerebellum, hippocampus and neocortex, is less frequent in subcortical regions, and is rarely observed in the brainstem and spinal cord. In addition, all DPRs can be detected in insoluble fractions from patient frontal cortex or cerebellum, as high-molecular-weight species, indicating they are aggregation prone [14,16,17,38]. A Meso Scale Discovery immunoassay has quantitatively detected poly-GP and poly-GA — but not, to date, any of the other DPRs — in both CNS tissue[43] and cerebrospinal fluid (CSF)[45]. CSF poly-GP levels might have utility as a biomarker for both diagnosis and pharmacodynamic response[46-48].

Pathogenesis of C9FTD/ALS

The complementary use of human tissue and *in vitro* and *in vivo* models, including illuminating mouse models examining both loss-of-function and gain-of-function mechanisms, has informed our current understanding of the contribution of C9orf72 loss of function, *C9orf72* repeat RNA and DPRs to pathogenesis. These studies also have implicated several downstream mechanisms resulting from *C9orf72* expansions (Fig. 4).

Loss-of-function mechanisms

C9orf72 transcription and splicing. In *C9orf72* transcript variant 2, the repeats are located within the promoter region (Fig. 1), so are not incorporated into variant 2 pre-mRNA but have the potential to affect the expression of this variant. By contrast, in variants 1 and 3, the repeats are within intron 1, so are included in the respective pre-mRNAs. Variant 2 is expressed at higher levels than variants 1 and 3 in CNS tissue[138,139].

Studies have demonstrated reduced levels of one or more of the *C9orf72* transcript variants in blood lymphocytes[1,132,140], induced pluripotent stem cell (iPSC)-derived neurons[12,90,94,138], frontal cortex[1,132,141-145], cerebellum[12,16,143-145], motor cortex[12] and cervical spinal cord[12] from *C9orf72* expansion carriers compared with controls. The findings are particularly robust for variants 1 and 2[145]. C9orf72 protein levels might be correspondingly reduced in the frontal cortex[110,144]. Transcripts upstream of the repeat are increased relative to downstream transcripts in blood lymphocytes and brain and spinal cord tissue[66] from *C9orf72* expansion carriers, possibly owing to abortive transcription in the presence of the repeat expansion[66]. Raised levels of variant 1 in the frontal cortex and cerebellum are associated with increased survival[145] — an important consideration when developing therapies that affect transcript levels.

The intronic location of the repeats in variants 1 and 3 means they should be spliced out of the transcript, but the fact that they are translated into DPRs implies that they are either retained in the transcript or that the spliced intron is translated. Although levels of mature spliced *C9orf72* mRNA are reduced in the brains of individuals with C9FTD/ALS, levels of sense and antisense transcripts containing intron 1, where the repeats are located, are increased[14,16], suggesting stabilization of repeat RNAs. Mature *C9orf72* transcripts with correct splicing or with retention of intron 1 are both detected in C9FTD/ALS lymphoblasts and brain tissue[138,146], indicating that both species contribute to RAN translation.

The expression of *C9orf72* is also modified by epigenetic effects (Box 1).

***Autophagy and lysosomal function*.** Bioinformatic analysis shows that C9orf72 is structurally related to the differentially expressed in normal and neoplastic cells (DENN) guanine nucleotide exchange factor (GEF) proteins, which activate Rab proteins. Rabs are crucial for a wide range of vesicular trafficking events, and multiple lines of evidence from several independent groups point to a role for C9orf72 in autophagy and endolysosomal trafficking and function.

Knockdown of *C9orf72* in human cell lines and primary neurons specifically inhibits autophagy induction, but not later stages of the autophagy pathway[91,96], leading to accumulation of p62[91,96] and cytoplasmic aggregation of TDP-43[96]. Consistent with these findings, accumulation of autophagy substrates, including p62, is observed in the spleens of *C9orf72*-knockout mice[54,56]. Conversely, overexpression of C9orf72 can activate autophagy, leading to an increase in autophagosomes in cell lines[91]. The role in autophagy seems to be mediated by the long C9orf72 protein isoform (C9orf72-L) rather than the short isoform (C9orf72-S)[96]. The mechanism underlying these changes involves Rab proteins, although no consensus has been reached on which are the most important, with Rab1a, Rab8a and Rab39b all being implicated[56,78,91,96,154]. C9orf72 interacts with guanine nucleotide exchange protein SMCR8 and WD repeat-containing protein 41 (WDR41)[56,58,96,101,122,154], and the effect on autophagy is generally agreed to be mediated through an interaction with the serine/threonine-protein kinase ULK1 complex[56,91,96,100,154], a key initiator of autophagy. Potentially, this interaction can also occur via TBK1[96], another serine/threonine-protein kinase implicated in autophagy. This finding is particularly intriguing, as loss-of-function mutations in *TBK1* cause ALS and FTD[155].

Another link to known ALS-associated genes is through ataxin-2 (*ATXN2*), in which intermediate expansions of polyglutamine increase the risk of ALS[156]. *C9orf72* knockdown specifically increases the aggregation and toxicity of ataxin-2 protein with intermediate polyglutamine repeats[96].

The relevance of the role of C9orf72 role in autophagy for disease pathogenesis is unclear, but neurons from patients with C9FTD/ALS have impaired basal autophagy[91,93] and increased sensitivity to autophagy inhibition[90], suggesting that reductions in C9orf72 levels contribute to

cellular distress. Furthermore, a Src–c-Abl pathway inhibitor, which increases autophagic processes, rescues survival defects in neurons derived from patients with ALS[92].

In addition to effects on autophagy, reduced endocytosis was reported in *C9orf72* knockdown cell lines[78], and impaired endosomal and lysosomal trafficking were observed in bone marrow-derived macrophages and microglia from homozygous *C9orf72* knockout mice[54], as well as in patient-derived fibroblasts and neurons[93]. Moreover, C9orf72 has been shown to reside on lysosomes and can directly affect lysosomal function, which might also explain the effects of *C9orf72* loss of function on both endolysosomal trafficking and autophagy[101]. In iPSC-derived motor neurons, C9orf72 primarily localizes to early endosomes, and iPSC-derived motor neurons from patients with C9FTD/ALS have fewer lysosomes[94]. Both patient neurons and CRISPR–Cas9 *C9orf72*-knockout iPSC-derived neurons have reduced vesicular trafficking, which can be rescued by C9orf72 overexpression[94]. These cells also have elevated glutamate receptor levels and increased sensitivity to excitotoxicity[94,127]. Consistent with this observation, increased glutamate receptor levels were found in spinal cord tissue in a *C9orf72*-knockout mice[94], and in spinal cord[94,127] and cortical tissue[94] from patients with C9FTD/ALS.

Taken together, these data suggest that C9orf72 is involved in multiple cellular trafficking events, and that loss of C9orf72 in both microglia and neurons can sensitize cells to other insults, thereby contributing to neurodegeneration in C9FTD/ALS.

Further insights from loss-of-function mouse models. The mouse *C9orf72* orthologue shares 98% homology with human *C9orf72* and is expressed in embryonic and early postnatal neurons, various regions of the adult brain and spinal cord, glia, and non-neuronal tissues, including muscle, spleen, kidney and testes[50-54,157]. Several *C9orf72* knockout or knockdown models have now been reported (Table 1)[51-58].

Transient reduction of *C9orf72* expression in the CNS by antisense oligonucleotides (ASOs)[13] and conditional homozygous knockouts of *C9orf72* in neurons and glia[51] do not lead to motor or

behavioural phenotypes. By contrast, ubiquitous knockouts of *C9orf72*[52-55,57,58] or CRISPR–Cas9-mediated knockouts of C9orf72 isoforms[56,57] throughout development led to immune system dysregulation in homozygous mice. The phenotypes included changes in myeloid and/or lymphoid cell populations in the spleen and lymph nodes, increased levels of inflammatory cytokines, and cervical or systemic lymphadenopathy and splenomegaly, sometimes with reduced body weight[52,53,55,57], neoplasia[55] or increased autoimmune antibody titres[53,57]. In comparison, haploinsufficiency of *C9orf72* does not lead to severe phenotypes.

Although some studies reported mild motor or cognitive phenotypes[52,53] or reduced lifespan[52,53,55,57,58] in homozygous *C9orf72*-knockout mice, none reported neuronal loss. Transcriptomic analysis confirmed changes in immune pathways[53,54], similar to those observed in CNS tissue from patients with *C9orf72* repeat expansions[54]. Transcriptomic analysis in human tissue has shown that *C9orf72* transcripts are particularly prevalent in $CD14^+$ myeloid cells, which are involved in innate and adaptive immunity[139]. Overall, in line with cellular studies of C9orf72, these findings suggest an important role for C9orf72 in immune regulation, possibly through its effects on autophagosome and lysosome function and/or microglial activity, or through alteration of autoimmune responses. However, in contrast to *C9orf72* loss of function, knockout of other autophagy-related genes, including *ATG7*[158] and *ATG5*[159], in neurons in mice leads to neurodegeneration. These findings suggest that C9orf72 is not an essential component of the autophagy pathway that mediates neuronal survival.

Crucially, none of the mouse *C9orf72* knockouts recapitulate ALS or FTD, suggesting that *C9orf72* loss of function is insufficient to precipitate disease. However, given the role of C9orf72 in pathways previously implicated in FTD and ALS[160], haploinsufficiency might contribute to the disease process in combination with gain-of-function mechanisms, and an interesting approach will be to breed loss-of-function mouse models with gain-of-function models.

Gain-of-function mechanisms

The question of whether *C9orf72* repeat RNA or DPRs produced by RAN translation are the toxic species in aetiopathogenesis is hotly debated in the field. Various approaches, each of which has limitations, have been used to address this issue. Post-mortem studies generally do not capture the earliest pathogenic events. *In vitro* and *in vivo* models often do not feature the long repeats that are found in patients, owing to methodological difficulties in cloning GC-rich repeats of this length. Therefore, these models might not fully recapitulate the disease mechanisms. Overexpression models do not necessarily reflect endogenous expression levels in patients. Many models express expanded repeats, but as these repeats can go on to produce both *C9orf72* repeat RNA and DPRs, attribution of downstream mechanisms to either entity is challenging. Despite these uncertainties, however, clear mechanistic pathways have emerged (Table 3).

C9orf72 repeat RNA. *In vitro*, GGGGCC repeat RNA forms secondary structures, including hairpins[45,66] and highly stable G-quadruplexes[45,66,161-163]. Other secondary structures, including DNA–RNA heteroduplexes, RNA duplexes and i-motifs[164-167], might arise from the sense and antisense repeat RNA and DNA sequences. *In vivo*, such secondary structures are likely to mediate the sequestration — and, as a consequence, depletion — of RNA-binding proteins (RBPs)[66,161] (Box 2), thus providing a clear potential route to RNA toxicity (reviewed extensively elsewhere[164,168]).

DPRs: insights from post-mortem studies. The results of post-mortem studies have raised suspicions that DPR inclusions are not the primary culprit in C9FTD/ALS pathogenesis[29,35-40,43,181-183]. TDP-43 pathology and neurodegeneration co-occur in affected regions of the CNS in ALS and FTD[29,35,37-39,43]. By contrast, DPR pathology does not coincide neuroanatomically with TDP-43 pathology[43] and is generally not found in the same neurons as TDP-43 inclusions[16,29,35,40]. Furthermore, DPR inclusions do not differ in neuroanatomical distribution between FTD and ALS cases[29,37,38] and are rare in the spinal cord in *C9orf72*-associated ALS, whereas TDP-43 pathology is common[14,17,18,29,36-40]. DPR inclusions are frequent across several brain regions[17,35-37,39], including structures that are thought to

be minimally affected in ALS and FTD, such as the cerebellum and occipital and parietal lobes. However, one study showed that poly-GR inclusions — but not the other DPRs — correlated with areas of neurodegeneration in *C9orf72*-associated ALS, and also, of all the DPRs, uniquely colocalized with TDP-43 pathology in a small sample of brains that were obtained shortly after death[44]. These data suggest that further investigation in larger, deep-phenotyped post-mortem cohorts will provide important insights.

Despite these observations, strong counterarguments to support a pathogenic role for DPRs in C9FTD/ALS have been put forward. Post-mortem studies tend to represent the final stages of the disease process and might not reflect the early pathogenicity of DPRs. Aggregates observed at post-mortem could represent protective species, and correlations with inclusions might be misleading if soluble species mediate neurotoxicity. CNS regions that have extensive DPR pathology but are unaffected by neurodegeneration might contain protective factors; indeed, selective vulnerability is a frequent observation in neurodegenerative diseases[184].

TDP-43 pathology is likely to be downstream of DPR pathology, probably explaining why it correlates more closely with neurodegeneration. This idea is consistent with downstream effects on TDP-43 in some experimental models expressing expanded repeats[109,185] or pure DPRs[106,185], suggesting that one or both of these gain-of-function mechanisms are linked to TDP-43. Affected individuals with DPR pathology but relatively mild or absent TDP-43 pathology have been reported[16,21,43,141,182,186,187]. These individuals include a young patient with evidence of intellectual disability[186], a patient with *C9orf72*-associated ALS who had little extramotor TDP-43 pathology but showed evidence of cognitive impairment and high cerebellar DPR levels[43], and patients with pathological or clinical diagnoses of FTLD or FTD[16,141,182,186,187], some of whom died prematurely from other causes[182]. Therefore, DPR pathology without substantial TDP-43 pathology seems to be sufficient for disease to develop in some cases. Furthermore, as discussed below, considerable evidence from model systems indicates that DPRs can cause neurodegeneration.

Insights from gain-of-function models

Evidence for C9orf72 repeat RNA toxicity. The effects of *C9orf72* repeat RNA were modelled in primary cortical and motor neurons transfected with expanded GGGGCC repeats within an artificial intronic region of the green fluorescent protein gene, reflecting the intronic human genomic context of *C9orf72* expansions[42]. These neurons demonstrated nuclear RNA foci and reduced survival. Dot blots and immunocytochemistry revealed no DPRs in these cells, suggesting that the reduced survival was attributable to repeat RNA. Interestingly, however, co-expression of poly-PR and the intronic expanded GGGGCC repeats had a synergistic detrimental effect on neuronal survival.

RNA toxicity has also been implicated in eye and motor neuron degeneration in a *Drosophila* model that expresses 30 GGGGCC repeats with a 6 bp (CTCGAG) interruption in the middle of the repeats [109,172]. DPRs were not detected in the eyes or neurons, and were only detected when the repeats were strongly induced in all tissues[109].

One caveat for the interpretation of both studies is that the inability to detect DPRs is not sufficient to exclude a role for these proteins. In our experience, poly-GR can be difficult to detect even in flies that express this protein at high levels and show overt toxicity. Therefore, more sensitive detection assays for DPRs will be required to unpick the relative contributions of RNA and DPRs in these models.

A study in developing zebrafish found that both sense and antisense RNA repeats could mediate toxicity, leading to a motor axonopathy phenotype[134]. No DPRs were detected in this model. In a *Drosophila* model, however, neither sense nor antisense repeats with similar lengths to those found in patients with C9FTD/ALS led to degeneration of adult neurons[189]. One potential explanation for this difference is that developing neurons are more susceptible to RNA toxicity than adult neurons. Another possibility is that RNA foci in humans sequester RBPs that are not present in *Drosophila*, thereby limiting the utility of *Drosophila* as a model for RNA toxicity.

Evidence for DPR toxicity. Two studies in *Drosophila* models have provided evidence that neurotoxicity is attributable to DPRs rather than repeat RNA[138,188]. In the first study, overexpression of expanded GGGGCC repeats in *Drosophila* eyes or adult neurons led to neurodegeneration[188]. This effect was inhibited when the repeats were interrupted by stop codons in each reading frame that prevented translation of the repeats into DPRs. The second study involved ubiquitous overexpression of 160 GGGGCC repeats in an intronic context, flanked by human *C9orf72* sequence[138]. The intron containing the repeat was spliced out and formed large numbers of sense RNA foci in neurons and glia, without production of DPRs. This model showed no evidence of neurodegeneration, reduced survival or widespread transcriptomic changes. Increasing transgene expression in this model led to DPR production and lifespan reduction, although the frequency of sense foci remained the same[138], supporting the idea that DPRs rather than sense foci mediate neurodegeneration.

It should be noted that although the RNA sequence interrupted by stop codons in the first study forms the same G-quadruplex secondary structure as uninterrupted *C9orf72* repeat RNA[188], the tertiary and quaternary structures are not necessarily identical, which might affect the dynamics of RBP sequestration. However, the intronic model did not have an interrupted repeat sequence and exhibited multiple sense foci, yet still showed no evidence of toxicity[138].

Toxicity of individual DPRs. In numerous studies in cell models[14,42,69,87,104-106,190], *Drosophila* [42,72,108,115,123], zebrafish[133,134,191] and mice[63,64], the repetitive GGGGCC sequence was altered to generate coding sequences that expressed each DPR in isolation, which was often sufficient to produce neurotoxicity and implicate several downstream mechanisms. The main limitation of these models is that DPR overexpression might not reflect the endogenous mechanisms that are seen in patients.

Studies in cultured neuronal or non-neuronal cell lines or primary neuronal cultures indicate that poly-GR and poly-PR are the most toxic of the DPRs[42,69,87,106]. These arginine-rich DPRs — in

particular, poly-PR — are toxic to yeast.[89] Poly-GA and poly-PA are also toxic, but to a lesser extent[89]. Synthetic poly-PR and poly-GR are highly toxic when exogenously applied to cultured human astrocytes. Poly-PR has a longer half-life than poly-GR, and is especially potent in this context[68]. Synthetic poly-GR and poly-GA are also toxic to primary neurons[192]. Furthermore, overexpression of poly-PR was found to be toxic to control iPSC-derived neurons[42].

Of all the DPRs, poly-GR had the greatest detrimental effect on development, locomotor activity and survival in a zebrafish model[133]. In *Drosophila* models, poly-GR and poly-PR were neurotoxic when expressed in the eyes[42,72,108,115,123,188]. In addition, flies expressing these proteins exhibited reduced survival[42,108,123,188] and locomotor phenotypes[42,123,135]. Most of these studies found that poly-GA[42,115,123], poly-PA[42,188] and poly-GP[108] were not toxic in *Drosophila*, although one study reported a mild reduced survival phenotype when poly-GA was expressed in adult neurons[188]. Consistent with this finding, poly-GA overexpression in cultured cells[104,105], primary neurons[104], zebrafish[133,191] or mouse brain[63] leads to toxicity, and synthetic poly-GA exogenously applied to human cells[190] or primary neurons[192] is also toxic. In addition, cryo-electron tomography revealed that poly-GA forms twisted ribbon structures that sequester the 26S proteasome[107].

Overall, these studies show that poly-GR and poly-PR are potently neurotoxic and poly-GA also exerts toxicity. The remaining DPRs, poly-PA and poly-GP, are unlikely to be toxic species. As poly-GR, poly-PR and poly-GA can all be damaging when overexpressed, a key aim is to establish the levels of each of these species in patient tissue, particularly at early disease stages. New techniques to extract and measure DPRs from patient tissue and assess their toxicity will be essential for this endeavour.

Mouse gain-of-function models. Several mouse models of *C9orf72* gain of function have been characterized. Adeno-associated virus-mediated CNS overexpression of 66 GGGGCC repeats leads to motor and behavioural phenotypes by 6 months, with RNA foci, DPRs, phospho-TDP-43 inclusions and neuronal loss being observed[59]. The same approach was used to generate mice specifically overexpressing poly-GA in the CNS[63]. These mice developed neuronal cytoplasmic inclusions of fibrillar poly-GA, as well as neurodegeneration and motor, cognitive and behavioural phenotypes. However, these effects were not observed when the poly-GA sequence was mutated to a sequence that was unable to aggregate. Phospho-TDP-43 inclusions were rarely found; therefore, this model does not fully recapitulate the $(GGGGCC)_{66}$ repeat mouse phenotype, indicating that species other than poly-GA are the main drivers of phospho-TDP-43 accumulation. A further poly-GA model, with expression levels more comparable to those in the cortex in patients with C9FTD/ALS, demonstrated motor deficits, but overall a less severe phenotype than the viral poly-GA model [64].

Bacterial artificial chromosome transgenic mouse models. Bacterial artificial chromosome (BAC) transgenic mice have the advantages of expressing the human *C9orf72* gene, with surrounding regulatory regions and flanking sequences, at more physiological levels. Three BAC transgenic models have produced similar findings. Two of these models used BAC constructs[52,60], containing exons 1–5 of the gene and 300–500 repeats (Table 2). The third BAC model expressed the full gene and protein with 100–1000 repeats[61]. All three models developed RNA foci and DPRs in the CNS, but none demonstrated TDP-43 inclusion pathology, neuronal loss or reduced survival. One model developed memory impairment and loss of hippocampal neurons, with an increase in levels of phosphorylated TDP-43, but was not fully reflective of ALS or FTD[52]. In this model, a single injection of an ASO targeting *C9orf72* RNA led to sustained reductions in RNA foci and DPR pathology, and improved cognition. Repeat length had a strong influence on the formation of RNA foci, with 100-repeat mice developing no foci and 400-repeat mice developing many foci, despite considerably higher transgene expression in the 100-repeat mice[52].

Only one study has reported BAC transgenic mice with a striking neurodegenerative phenotype[62]. The mice expressed the full gene and different repeat lengths in different lines. Three independent lines, two with ~500 repeats and one with high expression levels of 36 repeats, developed RNA foci, DPR aggregates, TDP-43 pathology and neurodegeneration. A subset of female mice developed an acute motor phenotype, with paralysis, weight loss and decreased survival, with other female and male mice showing slower progressive neurodegeneration. Antisense transcripts and foci were observed at especially high levels in some CNS regions that showed neurodegeneration, whereas sense foci were more evenly distributed. These data show that high expression of short repeats can be toxic, indicating that large repeats are not the critical factor for toxicity in this mouse model. The reason why female mice are particularly susceptible and only a subset develops acute disease is currently unclear, although factors such as methylation might be important.

In-depth molecular and phenotypic comparisons between all the BAC models should further our understanding of *C9orf72*-associated pathogenic mechanisms. One model that expresses the full-length human gene with the repeat expansion recapitulates disease[62] whereas the other does not[61]. These mice feature different genetic backgrounds and flanking *C9orf72* sequences (Table 2), which might be contributory factors. Backcrossing the models that did not show a phenotype into different genetic backgrounds, so as to determine whether certain backgrounds facilitate disease, would provide new insight.

Downstream mechanisms

A range of downstream mechanisms in C9FTD/ALS have been validated across multiple human and non-human model systems and different laboratories (Table 3). Dysfunctional nucleocytoplasmic transport (NCT) is a prominent mechanism that has been identified in genome-wide screens in yeast[89] and *Drosophila*[108,109,115], and in CRISPR–Cas9 screens in human cells and primary neurons[88] (reviewed elsewhere[168,193]).

Poly-PR and poly-GR have been shown to interact with proteins that contain low complexity domains (LCDs), which include many RBPs[72,194]. LCD proteins can undergo liquid–liquid phase separation (LLPS) to form droplets. Through this process, the proteins become compartmentalized in the cell, forming membrane-less organelles such as nucleoli and stress granules. These organelles facilitate the assembly of RNA and RBPs into ribonucleoproteins, and also aid subsequent RNA metabolism. Furthermore, these LCDs can form reversible hydrogels, which have the propensity to fibrillize into irreversible hydrogel-like structures. Mutations frequently occur in the LCD domains in several RBPs that are involved in ALS, including FUS and TDP-43[195]. These mutations can alter the LLPS dynamics of RBPs, leading to their fibrillization and aggregation[196-198]. Poly-GR and poly-PR interact with LCD proteins in nucleoli and stress granules, thereby impairing LLPS, disrupting the dynamics of assembly of these organelles, and affecting mRNA translation and NCT[72,73,82]. Aliphatic alcohols, which disrupt phase separation and hydrogel formation, can reduce poly-PR's protein interactions[194] and disrupt the nucleolar localization of poly-PR and poly-GR[72]. Knockdown of several of these LCD proteins modifies the eye degeneration phenotype in a *Drosophila* poly-GR model[72]. Poly-PR also interacts with LCDs in intermediate filament proteins[194]. Interestingly, arginine-rich DPRs at high concentrations can undergo LLPS[73]. In addition, GGGGCC repeat RNA can undergo gel transition[81,199] and induce phase transition of RNA granule proteins[81] in the absence of other LCD proteins. These data further implicate altered LLPS in C9FTD/ALS pathogenesis.

Effects on membrane-less organelles have been observed in several other studies. Primary cortical neurons that overexpress poly-PR, leading to the formation of nuclear poly-PR aggregates, show a reduction in cytoplasmic P-bodies and an increase in stress granule formation[42]. In cultured cells treated with arsenite, a stress granule inducer, overexpression of poly-PR and poly-GR reduced stress granule formation, whereas poly-GA, poly-GP and poly-PA induced stress granule formation[69].

Poly-PR interacts with translation initiation and elongation factors, and overexpression of poly-GR and poly-PR, but not poly-GA, inhibited translation *in vitro* and in cell lines[87]. Translational inhibition was attributed to direct binding of poly-GR and poly-PR to mRNA, thereby blocking access

to the translation machinery[87]. A second study also observed translational inhibition when $(GGGGCC)_{31}$ was expressed in cell lines[79]. This inhibition was accompanied by an increase in stress granule formation (a marker of translational arrest) and nuclear accumulation of poly-A mRNAs and PABPc, a protein that facilitates mRNA translation in the cytoplasm[79]. GGGGCC repeat-induced stress granule formation was also observed in subsequent studies[80,81]. The translational inhibition might be explained by sequestration of PABPc by GGGGCC repeat RNA, but could also be due to the DPR-mediated mechanisms described above. Interestingly, two studies that identified translation inhibition suggested that this process was independent of eukaryotic translation initiation factor 2 subunit 1 (eIF2α) phosphorylation, an important master regulator of translation that has been implicated in other neurodegenerative diseases[200]. However, stress granule induction secondary to poly-PR has been shown to require eIF2α phosphorylation[73].

A study published in 2018 showed that poly-GR and poly-PR induced stress granule assembly and localization of NCT factors into these stress granules, thereby mediating NCT dysfunction[82]. Inhibition of stress granule assembly abrogated NCT dysfunction and neurodegeneration in patient-derived neurons and *in vivo*[82]. Stress granules and translation inhibition have been implicated in FTD and ALS associated with TDP-43 and FUS inclusions[201], and they provide an interesting potential common theme across the disease spectrum.

In three elegant studies, GGGGCC repeat overexpression constructs were used to investigate the mechanisms underlying RAN translation. Two of these studies showed that the integrated stress response, via eIF2α phosphorylation, selectively increased RAN translation[80,83]. Therefore, repeat-induced cellular stress could lead to both impaired translation and enhanced RAN translation, causing a negative feedback loop. In addition, two of the studies showed that RAN translation occurred on unspliced, capped mRNA (in which the repeat-containing intron is retained), and was initiated by an upstream CUG acting as a start codon[80,202]. By contrast, the third study found that RAN translation occurred on the spliced intronic RNA and was cap-independent, although cap-

dependent RAN translation was more efficient[83]. These mechanisms now need to be investigated in the context of endogenous RAN translation.

When applied exogenously to astrocyte cultures, poly-GR and poly-PR 20mers accumulate in the nucleolus, leading to splicing changes and impaired ribosomal RNA (rRNA) maturation[68]. This finding led to the suggestion that these DPRs bind to and impair nuclear RBP complexes that are involved in mRNA splicing, which could in turn affect ribosome biogenesis and other important genes[68]. Interestingly, given the effect of poly-PR and poly-GR on NCT, splicing of RanGAP, an essential regulator of NCT, was altered[68]. Poly-PR and poly-GR also colocalize with nucleoli in cultured cells, primary neurons, iPSC-derived neurons and *Drosophila*, leading to abnormal nucleolar morphology[36,42,67,69,71]. Importantly, the brains of individuals with *C9orf79*-associated FTLD exhibit bidirectional nucleolar volume changes, with smaller neuronal nucleoli overall but enlarged nucleoli in neurons containing poly-GR inclusions[67]. Other interactome studies suggest that poly-PR and poly-GR associate with the spliceosome component U2 snRNP in cell nuclear extracts, block spliceosome assembly, and disrupt splicing when applied to cell nuclear extracts[85]. These proteins also interact with mitochondrial ribosomal proteins and other ribosomal proteins[117], and evidence of impaired mitochondrial function has been observed in iPSC-derived neurons from patients with C9FTD/ALS[75,117]. Ribosomal proteins, hnRNPs, nucleolar proteins and RBPs associated with RNA granules have been shown to interact with poly-GR, and overexpression of poly-GR and poly-PR led to decreased levels of rRNA in human cell lines[69]. In agreement with these data, poly-PR was shown to interact with proteins involved in mRNA splicing and ribosome assembly, and with ribosomal proteins[87]. Three studies have shown that in addition to effects on RNA, *C9orf72* repeats can induce DNA damage, probably mediated by DPRs[117-119].

Overall, many cellular pathways have been implicated in gain-of-function toxicity, with the majority of studies focused on poly-GR and poly-PR. Given the highly toxic nature of these DPRs, they are likely to influence multiple pathways. Links between the downstream mechanisms and specific DPRs or repeat RNA have not yet been established, and sensitive techniques to measure

both repeat RNA and DPRs in models expressing pure GGGGCC repeats will be required to enable the effects of individual molecular species to be distinguished.

Another important issue that remains to be addressed is whether different DPRs or repeat RNA and DPRs act synergistically, potentially in conjunction with loss of function of *C9orf72*, to elicit downstream effects. Crossing of models that express different DPR species should help to address this question. A further priority is to develop physiologically relevant models to reflect endogenous levels of *C9orf72* repeat RNA and DPRs, and human iPSC-derived neuronal models will be a key tool in this regard. Furthermore, C9orf72 protein depletion, repeat RNA and DPRs might have non-cell-autonomous effects, and iPSC models of neuronal, glial and muscle co-culture should increase our understanding of these complex interactions.

Therapeutic strategies

Targeting C9orf72 RNA or DNA

ASOs that target *C9orf72* RNA can rescue *C9orf72*-specific pathologies[12,13,46,126], downstream gene expression changes[12,126], NCT defects[109] and TDP-43 mislocalization[109] in *C9orf72* fibroblasts or iPSC-derived neurons. These ASOs also reduce NCT defects and neurodegeneration in *Drosophila*[109] and diminish sense RNA foci and DPRs in mice[46,52]. ASO trials in humans are planned, and the feasibility of this strategy has a precedent in studies of superoxide dismutase 1 (*SOD1*)-targeting ASOs in patients with ALS[203].

An alternative strategy is to use compounds that target the secondary structure of *C9orf72* repeat RNA[45,109,204-206]. These compounds could prevent sequestration of RBPs and/or interfere with the RNA structure to prevent RAN translation. Such molecules have been shown to affect the secondary structure of *C9orf72* repeat RNA *in vitro*[45,204,206], to reduce the production of RNA foci and DPRs when applied to patient neurons[45,205] and *Drosophila*[205], and to rescue nuclear import defects and neurodegeneration in a *Drosophila* model[109].

Targeting of the transcription elongation factor SPT4 reduces levels of sense and antisense *C9orf72* repeat transcripts and ameliorates disease phenotypes *in vitro* and *in vivo*, including in *C9orf72* iPSC-derived neurons[207]. Single-stranded small inhibitory RNAs have also been proposed as a strategy to silence the *C9orf72* repeat RNA[208]. In addition, two studies have shown that use of the CRISPR–Cas9 system to target either GGGGCC repeat DNA[209] — thus reducing repeat transcription — or repeat RNA[210] can reduce RNA foci and DPR levels in cell lines. On the basis of these studies, targeting of *C9orf72* RNA or DNA are promising strategies, and ASOs are currently the most advanced in terms of clinical development.

Targeting DPRs and TDP-43

In other neurodegenerative diseases, active and passive immunological approaches have been used to target toxic proteins such as amyloid-β[211], tau[212,213] and α-synuclein[214,215], leading to improved pathology and phenotypes in model systems. With the ongoing development of specific DPR antibodies, passive immunization to the DPRs could present a novel therapeutic approach for C9FTD/ALS, and poly-GA-specific antibodies have been shown to reduce intracellular poly-GA aggregation and seeding activity of C9FTD/ALS brain extracts[216]. In addition, given that C9FTD/ALS pathogenesis converges on TDP-43, anti-TDP-43 immunotherapy would be a compelling strategy.

Current strategies in other neurodegenerative diseases are based on targeting of the extracellular pool of protein, which is presumably involved in cell-to-cell transmission. The success of such an approach in the case of DPRs and TDP-43 would depend on whether there is a disease-relevant extracellular pool to target. If not, methods to target antibodies intracellularly would be required. Other important factors to consider are the ideal timing of these treatments, which DPRs and conformations to target, the specificity and safety of the treatment, and the risks of precipitating an autoimmune response. As short repetitive sequences similar to DPRs are present across the proteome, specificity might present a barrier. Increasing clearance of DPRs by other

mechanisms could also be effective; for example, the small heat shock protein HSPB8 was shown to reduce DPR levels, probably via the autophagy pathway[98].

Targeting downstream mechanisms

Targeting of downstream mechanisms might represent a useful therapeutic strategy for C9FTD/ALS. Reducing nuclear export by targeting the nuclear export factors SRSF1 or exportin 1 ameliorates toxicity in *C9orf72*-repeat *Drosophila*[84,109]; this was suggested to be due to either by reducing the levels of cytoplasmic repeat RNA (and, thus, DPRs) or by a more general mechanism to alterations in NCT. Importantly, in addition to the proof of concept genetic knockdown approaches that were used to reduce SRSF1 and exportin 1, small molecule exportin 1 inhibitors were also used, indicating a potential route to the clinic. It will now be important to address whether there is a large enough therapeutic window when targeting nuclear export, which is a fundamental cellular process. Inhibition of stress granule formation using either ASOs targeting Ataxin-2 or small molecules, also prevents NCT defects in C9-ALS iPSC-neurons and neurodegeneration in a *C9orf72* repeat fly model[82]. Given the success of Ataxin-2-targeting ASOs in extending lifespan in an ALS mouse model overexpressing TDP-43, this strategy and the mechanisms underlying it, are of clear interest. A Src–c-Abl pathway inhibitor that augments autophagy, improved the survival of iPSC-derived neurons from patients with ALS, including those with C9FTD/ALS, suggesting this approach could be beneficial for several forms of FTD/ALS[92]. Knockdown of TMX2, an endoplasmic reticulum protein that was identified in CRISPR–Cas9 screens in human cells and primary neurons, modulated endoplasmic reticulum stress in primary neurons overexpressing DPRs and increased survival of neurons derived from patients with *C9orf72*-associated ALS [88]. Whether TMX2 is druggable awaits further study. High-throughput drug screens — for example, using iPSC-derived neuronal models[92,94,217] — will be of great importance in establishing treatments to reduce levels of repeat RNA and DPRs, repair dysfunctional cellular mechanisms and rescue disease phenotypes[88].

Conclusions

Our knowledge of C9FTD/ALS has increased exponentially within a relatively short time period. Dissection of the disease mechanisms has not been straightforward, and the emerging picture is one of a combination of a diverse range of factors that lead to neurodegeneration. Given the relative importance of *C9orf72* repeat expansions as a causative factor for ALS and FTD, rapid translation of the accumulated knowledge into therapeutic strategies would have a substantial impact on patients with these devastating neurodegenerative diseases.

1. DeJesus-Hernandez, M. et al. Expanded GGGGCC hexanucleotide repeat in noncoding region of *C9ORF72* causes chromosome 9p-linked FTD and ALS. *Neuron* **72**, 245-256 (2011).
2. Renton, A. E. et al. A hexanucleotide repeat expansion in *C9ORF72* is the cause of chromosome 9p21-linked ALS–FTD. *Neuron* **72**, 257-268 (2011).
3. Majounie, E. et al. Frequency of the *C9orf72* hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. *Lancet Neurol.* **11**, 323-330 (2012).
4. Woollacott, I. O. & Mead, S. The *C9ORF72* expansion mutation: gene structure, phenotypic and diagnostic issues. *Acta Neuropathol.* **127**, 319-332 (2014).
5. Rutherford, N. J. et al. Length of normal alleles of *C9ORF72* GGGGCC repeat do not influence disease phenotype. *Neurobiol. Aging* **33**, 2950.e5-2950.e7 (2012).
6. Beck, J. et al. Large *C9orf72* hexanucleotide repeat expansions are seen in multiple neurodegenerative syndromes and are more frequent than expected in the UK population. *Am. J. Hum. Genet.* **92**, 345-353 (2013).
7. Simón-Sánchez, J. et al. The clinical and pathological phenotype of *C9ORF72* hexanucleotide repeat expansions. *Brain* **135**, 723-735 (2012).
8. Harms, M. B. et al. Lack of *C9ORF72* coding mutations supports a gain of function for repeat expansions in amyotrophic lateral sclerosis. *Neurobiol. Aging* **34**, 2234.e13-2234.e19 (2013).

9. van der Zee, J. et al. A pan-European study of the *C9orf72* repeat associated with FTLD: geographic prevalence, genomic instability, and intermediate repeats. *Hum. Mutat.* **34**, 363-373 (2013).

10. van Blitterswijk, M. et al. Association between repeat sizes and clinical and pathological characteristics in carriers of *C9ORF72* repeat expansions (Xpansize-72): a cross-sectional cohort study. *Lancet Neurol.* **12**, 978-988 (2013).

11. Mizielinska, S. et al. *C9orf72* frontotemporal lobar degeneration is characterised by frequent neuronal sense and antisense RNA foci. *Acta Neuropathol.* **126**, 845-857 (2013).

12. Donnelly, C. J. et al. RNA toxicity from the ALS/FTD *C9ORF72* expansion is mitigated by antisense intervention. *Neuron* **80**, 415-428 (2013).

13. Lagier-Tourenne, C. et al. Targeted degradation of sense and antisense *C9orf72* RNA foci as therapy for ALS and frontotemporal degeneration. *Proc. Natl Acad. Sci. USA* **110**, E4530-E4539 (2013).

14. Zu, T. et al. RAN proteins and RNA foci from antisense transcripts in *C9ORF72* ALS and frontotemporal dementia. *Proc. Natl Acad. Sci. USA* **110**, E4968-E4977 (2013).

15. Gendron, T. F. et al. Antisense transcripts of the expanded *C9ORF72* hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS. *Acta Neuropathol.* **126**, 829-844 (2013).

16. Mori, K. et al. The *C9orf72* GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. *Science* **339**, 1335-1338 (2013).

17. Ash, P. E. et al. Unconventional translation of *C9ORF72* GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron* **77**, 639-646 (2013).

18. Mori, K. et al. Bidirectional transcripts of the expanded *C9orf72* hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. *Acta Neuropathol.* **126**, 881-893 (2013).

19. Rohrer, J. D. et al. *C9orf72* expansions in frontotemporal dementia and amyotrophic lateral sclerosis. *Lancet Neurol.* **14**, 291-301 (2015).

20. Renton, A. E., Chio, A. & Traynor, B. J. State of play in amyotrophic lateral sclerosis genetics. *Nat. Neurosci.* **17**, 17-23 (2014).

21. Gami, P. et al. A 30-unit hexanucleotide repeat expansion in *C9orf72* induces pathological lesions with dipeptide-repeat proteins and RNA foci, but not TDP-43 inclusions and clinical disease. *Acta Neuropathol.* **130**, 599-601 (2015).

22. Gijselinck, I. et al. The *C9orf72* repeat size correlates with onset age of disease, DNA methylation and transcriptional downregulation of the promoter. *Mol. Psychiatry* **21**, 1112-1124 (2016).

23. Nordin, A. et al. Extensive size variability of the GGGGCC expansion in *C9orf72* in both neuronal and non-neuronal tissues in 18 patients with ALS or FTD. *Hum. Mol. Genet.* **24**, 3133-3142 (2015).

24. Fratta, P. et al. Screening a UK amyotrophic lateral sclerosis cohort provides evidence of multiple origins of the *C9orf72* expansion. *Neurobiol. Aging* **36**, 546.e1-546.e7 (2015).

25. Xi, Z. et al. Jump from pre-mutation to pathologic expansion in *C9orf72*. *Am. J. Hum. Genet.* **96**, 962-970 (2015).

26. Neumann, M. et al. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* **314**, 130-133 (2006).

27. Arai, T. et al. TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Biochem. Biophys. Res. Commun.* **351**, 602-611 (2006).

28. Ling, S. C., Polymenidou, M. & Cleveland, D. W. Converging mechanisms in ALS and FTD: disrupted RNA and protein homeostasis. *Neuron* **79**, 416-438 (2013).

29. Mackenzie, I. R. et al. Dipeptide repeat protein pathology in *C9ORF72* mutation cases: clinico-pathological correlations. *Acta Neuropathol.* **126**, 859-879 (2013).

30. Murray, M. E. et al. Clinical and neuropathologic heterogeneity of c9FTD/ALS associated with hexanucleotide repeat expansion in *C9ORF72*. *Acta Neuropathol.* **122**, 673-690 (2011).

31. Hsiung, G. Y. et al. Clinical and pathological features of familial frontotemporal dementia caused by *C9ORF72* mutation on chromosome 9p. *Brain* **135**, 709-722 (2012).

32. Mahoney, C. J. et al. Frontotemporal dementia with the *C9ORF72* hexanucleotide repeat expansion: clinical, neuroanatomical and neuropathological features. *Brain* **135**, 736-750 (2012).

33. Irwin, D. J. et al. Cognitive decline and reduced survival in *C9orf72* expansion frontotemporal degeneration and amyotrophic lateral sclerosis. *J. Neurol. Neurosurg. Psychiatry* **84**, 163-169 (2013).

34. DeJesus-Hernandez, M. et al. In-depth clinico-pathological examination of RNA foci in a large cohort of *C9ORF72* expansion carriers. *Acta Neuropathol.* **134**, 255-269 (2017).

35. Mann, D. M. et al. Dipeptide repeat proteins are present in the p62 positive inclusions in patients with frontotemporal lobar degeneration and motor neurone disease associated with expansions in *C9ORF72*. *Acta Neuropathol. Commun.* **1**, 68 (2013).

36. Schludi, M. H. et al. Distribution of dipeptide repeat proteins in cellular models and *C9orf72* mutation cases suggests link to transcriptional silencing. *Acta Neuropathol.* **130**, 537-555 (2015).

37. Davidson, Y. S. et al. Brain distribution of dipeptide repeat proteins in frontotemporal lobar degeneration and motor neurone disease associated with expansions in *C9ORF72*. *Acta Neuropathol. Commun.* **2**, 70 (2014).

38. Mackenzie, I. R. et al. Quantitative analysis and clinico-pathological correlations of different dipeptide repeat protein pathologies in *C9ORF72* mutation carriers. *Acta Neuropathol.* **130**, 845-861 (2015).

39. Davidson, Y. et al. Neurodegeneration in frontotemporal lobar degeneration and motor neurone disease associated with expansions in *C9orf72* is linked to TDP-43 pathology and

not associated with aggregated forms of dipeptide repeat proteins. *Neuropathol. Appl. Neurobiol.* **42**, 242-254 (2016).

40. Gomez-Deza, J. et al. Dipeptide repeat protein inclusions are rare in the spinal cord and almost absent from motor neurons in *C9ORF72* mutant amyotrophic lateral sclerosis and are unlikely to cause their degeneration. *Acta Neuropathol. Commun.* **3**, 38 (2015).

41. Liu, E. Y. et al. *C9orf72* hypermethylation protects against repeat expansion-associated pathology in ALS/FTD. *Acta Neuropathol.* **128**, 525-541 (2014).

42. Wen, X. et al. Antisense proline-arginine RAN dipeptides linked to *C9ORF72*-ALS/FTD form toxic nuclear aggregates that initiate in vitro and in vivo neuronal death. *Neuron* **84**, 1213-1225 (2014).

43. Gendron, T.F. et al. Cerebellar c9RAN proteins associate with clinical and neuropathological characteristics of *C9ORF72* repeat expansion carriers. *Acta Neuropathol.* **130**, 559-573 (2015).

44. Saberi, S. et al. Sense-encoded poly-GR dipeptide repeat proteins correlate to neurodegeneration and uniquely co-localize with TDP-43 in dendrites of repeat-expanded C9orf72 amyotrophic lateral sclerosis. *Acta Neuropathol.* 135, 459-474 (2018).

45. Su, Z. et al. Discovery of a biomarker and lead small molecules to target r(GGGGCC)-associated defects in c9FTD/ALS. *Neuron* **83**, 1043-1050 (2014).

46. Gendron, T. F. et al. Poly(GP) proteins are a useful pharmacodynamic marker for *C9ORF72*-associated amyotrophic lateral sclerosis. *Sci. Transl. Med.* **9**, eaai7866 (2017).

47. Lehmer, C. et al. Poly-GP in cerebrospinal fluid links *C9orf72*-associated dipeptide repeat expression to the asymptomatic phase of ALS/FTD. *EMBO Mol. Med.* **9**, 859-868 (2017).

48. Balendra, R., Moens, T.G. & Isaacs, A. M. Specific biomarkers for *C9orf72* FTD/ALS could expedite the journey towards effective therapies. *EMBO Mol. Med.* **9**, 853-855 (2017).

49. Panda, S. K. et al. Highly efficient targeted mutagenesis in mice using TALENs. *Genetics* **195**, 703-713 (2013).

50. Suzuki, N. et al. The mouse *C9ORF72* ortholog is enriched in neurons known to degenerate in ALS and FTD. *Nat. Neurosci.* **16**, 1725-1727 (2013).

51. Koppers, M. et al. *C9orf72* ablation in mice does not cause motor neuron degeneration or motor deficits. *Ann. Neurol.* **78**, 426-438 (2015).

52. Jiang, J. et al. Gain of toxicity from ALS/FTD-linked repeat expansions in *C9ORF72* Is alleviated by antisense oligonucleotides targeting GGGGCC-containing RNAs. *Neuron* **90**, 535-550 (2016).

53. Atanasio, A. et al. *C9orf72* ablation causes immune dysregulation characterized by leukocyte expansion, autoantibody production, and glomerulonephropathy in mice. *Sci. Rep.* **6**, 23204 (2016).

54. O'Rourke, J. G. et al. *C9orf72* is required for proper macrophage and microglial function in mice. *Science* **351**, 1324-1329 (2016).

55. Sudria-Lopez, E. et al. Full ablation of C9orf72 in mice causes immune system-related pathology and neoplastic events but no motor neuron defects. *Acta Neuropathol.* **132**, 145-147 (2016).

56. Sullivan, P.M. et al. The ALS/FTLD associated protein C9orf72 associates with SMCR8 and WDR41 to regulate the autophagy-lysosome pathway. *Acta Neuropathol. Commun.* **4**, 51 (2016).

57. Burberry, A. et al. Loss-of-function mutations in the *C9ORF72* mouse ortholog cause fatal autoimmune disease. *Sci. Transl. Med.* **8**, 347ra93 (2016).

58. Ugolino, J. et al. Loss of C9orf72 enhances autophagic activity via deregulated mTOR and TFEB signaling. *PLoS Genet.* **12**, e1006443 (2016).

59. Chew, J. et al. *C9ORF72* repeat expansions in mice cause TDP-43 pathology, neuronal loss, and behavioral deficits. *Science* **348**, 1151-1154 (2015).

60. Peters, O. M. et al. Human *C9ORF72* hexanucleotide expansion reproduces RNA foci and dipeptide repeat proteins but not neurodegeneration in BAC transgenic mice. *Neuron* **88**, 902-909 (2015).

61. O'Rourke, J. G. et al. *C9orf72* BAC transgenic mice display typical pathologic features of ALS/FTD. *Neuron* **88**, 892-901 (2015).

62. Liu, Y. et al. *C9orf72* BAC mouse model with motor deficits and neurodegenerative features of ALS/FTD. *Neuron* **90**, 521-534 (2016).

63. Zhang, Y. J. et al. *C9ORF72* poly(GA) aggregates sequester and impair HR23 and nucleocytoplasmic transport proteins. *Nat. Neurosci.* **19**, 668-677 (2016).

64. Schludi, M. H. et al. Spinal poly-GA inclusions in a *C9orf72* mouse model trigger motor deficits and inflammation without neuron loss. *Acta Neuropathol.* **134**, 241-254 (2017).

65. Herranz-Martin, S. et al. Viral delivery of *C9orf72* hexanucleotide repeat expansions in mice leads to repeat-length-dependent neuropathology and behavioural deficits. *Dis. Model. Mech.* **10**, 859-868 (2017).

66. Haeusler, A. R. et al. *C9orf72* nucleotide repeat structures initiate molecular cascades of disease. *Nature* **507**, 195-200 (2014).

67. Mizielinska, S. et al. Bidirectional nucleolar dysfunction in *C9orf72* frontotemporal lobar degeneration. *Acta Neuropathol. Commun.* **5**, 29 (2017).

68. Kwon, I. et al. Poly-dipeptides encoded by the *C9orf72* repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science* **345**, 1139-1145 (2014).

69. Tao, Z. et al. Nucleolar stress and impaired stress granule formation contribute to *C9orf72* RAN translation-induced cytotoxicity. *Hum. Mol. Genet.* **24**, 2426-2441 (2015).

70. Stopford, M. J. et al. *C9ORF72* hexanucleotide repeat exerts toxicity in a stable, inducible motor neuronal cell model, which is rescued by partial depletion of Pten. *Hum. Mol. Genet.* **26**, 1133-1145 (2017).

71. Bennion Callister, J., Ryan, S., Sim, J., Rollinson, S. & Pickering-Brown, S. M. Modelling C9orf72 dipeptide repeat proteins of a physiologically relevant size. *Hum. Mol. Genet.* **25**, 5069-5082 (2016).

72. Lee, K. H. et al. C9orf72 dipeptide repeats impair the assembly, dynamics, and function of membrane-less organelles. *Cell* **167**, 774-788.e17 (2016).

73. Boeynaems, S. et al. Phase separation of *C9orf72* dipeptide repeats perturbs stress granule dynamics. *Mol. Cell* **65**, 1044-1055.e5 (2017).

74. Maharjan, N., Kunzli, C., Buthey, K. & Saxena, S. C9ORF72 regulates stress granule formation and its deficiency impairs stress granule assembly, hypersensitizing cells to stress. *Mol. Neurobiol.* **54**, 3062-3077 (2017).

75. Dafinca, R. et al. *C9orf72* hexanucleotide expansions are associated with altered ER calcium homeostasis and stress granule formation in iPSC-derived neurons from patients with amyotrophic lateral sclerosis and frontotemporal dementia. *Stem Cells* **34**, 2063-2078 (2016).

76. McGurk, L. et al. Poly-A binding protein-1 localization to a subset of TDP-43 inclusions in amyotrophic lateral sclerosis occurs more frequently in patients harboring an expansion in *C9orf72*. *J. Neuropathol. Exp. Neurol.* **73**, 837-845 (2014).

77. Daigle, J. G. et al. Pur-alpha regulates cytoplasmic stress granule dynamics and ameliorates FUS toxicity. *Acta Neuropathol.* **131**, 605-620 (2016).

78. Farg, M. A. et al. C9ORF72, implicated in amytrophic lateral sclerosis and frontotemporal dementia, regulates endosomal trafficking. *Hum. Mol. Genet.* **23**, 3579-3595 (2014).

79. Rossi, S. et al. Nuclear accumulation of mRNAs underlies G4C2-repeat-induced translational repression in a cellular model of *C9orf72* ALS. *J. Cell Sci.* **128**, 1787-1799 (2015).

80. Green, K. M. et al. RAN translation at *C9orf72*-associated repeat expansions is selectively enhanced by the integrated stress response. *Nat. Commun.* **8**, 2005 (2017).

81. Fay, M. M., Anderson, P. J. & Ivanov, P. ALS/FTD-associated *C9ORF72* repeat RNA promotes phase transitions in vitro and in cells. *Cell Rep.* **21**, 3573-3584 (2017).

82. Zhang, K. et al. Stress granule assembly disrupts nucleocytoplasmic transport. *Cell* **173**, 958-971.e17 (2018).

83. Cheng, W. et al. C9ORF72 GGGGCC repeat-associated non-AUG translation is upregulated by stress through eIF2α phosphorylation. *Nat. Commun.* **9**, 51 (2018).

84. Hautbergue, G. M. et al. SRSF1-dependent nuclear export inhibition of *C9ORF72* repeat transcripts prevents neurodegeneration and associated motor deficits. *Nat. Commun.* **8**, 16063 (2017).

85. Yin, S. et al. Evidence that C9ORF72 dipeptide repeat proteins associate with U2 snRNP to cause mis-splicing in ALS/FTD patients. *Cell Rep.* **19**, 2244-2256 (2017).

86. Prudencio, M. et al. Distinct brain transcriptome profiles in *C9orf72*-associated and sporadic ALS. *Nat. Neurosci.* **18**, 1175-1182 (2015).

87. Kanekura, K. et al. Poly-dipeptides encoded by the C9ORF72 repeats block global protein translation. *Hum. Mol. Genet.* **25**, 1803-1813 (2016).

88. Kramer, N. J. et al. CRISPR–Cas9 screens in human cells and primary neurons identify modifiers of C9ORF72 dipeptide-repeat-protein toxicity. *Nat. Genet.* **50**, 603-612 (2018).

89. Jovicic, A. et al. Modifiers of *C9orf72* dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS. *Nat. Neurosci.* **18**, 1226-1229 (2015).

90. Almeida, S. et al. Modeling key pathological features of frontotemporal dementia with *C9ORF72* repeat expansion in iPSC-derived human neurons. *Acta Neuropathol.* **126**, 385-399 (2013).

91. Webster, C. P. et al. The C9orf72 protein interacts with Rab1a and the ULK1 complex to regulate initiation of autophagy. *EMBO J.* **35**, 1656-1676 (2016).

92. Imamura, K. et al. The Src/c–Abl pathway is a potential therapeutic target in amyotrophic lateral sclerosis. *Sci. Transl. Med.* **9**, eaaf3962 (2017).

93. Aoki, Y. et al. *C9orf72* and *RAB7L1* regulate vesicle trafficking in amyotrophic lateral sclerosis and frontotemporal dementia. *Brain* **140**, 887-897 (2017).

94. Shi, Y. et al. Haploinsufficiency leads to neurodegeneration in *C9ORF72* ALS/FTD human induced motor neurons. *Nat. Med.* **24**, 313-325 (2018).

95. Onesto, E. et al. Gene-specific mitochondria dysfunctions in human *TARDBP* and *C9ORF72* fibroblasts. *Acta Neuropathol. Commun.* **4**, 47 (2016).

96. Sellier, C. et al. Loss of C9ORF72 impairs autophagy and synergizes with polyQ Ataxin-2 to induce motor neuron dysfunction and cell death. *EMBO J* (2016).

97. Busch, J.I. et al. Increased expression of the frontotemporal dementia risk factor TMEM106B causes C9orf72-dependent alterations in lysosomes. *Hum. Mol. Genet.* **35**, 1276-1297 (2016).

98. Cristofani, R. et al. The small heat shock protein B8 (HSPB8) efficiently removes aggregating species of dipeptides produced in C9ORF72-related neurodegenerative diseases. *Cell Stress Chaperones* **23**, 1-12 (2018).

99. Cristofani, R. et al. Inhibition of retrograde transport modulates misfolded protein accumulation and clearance in motoneuron diseases. *Autophagy* **13**, 1280-1303 (2017).

100. Jung, J. et al. Multiplex image-based autophagy RNAi screening identifies SMCR8 as ULK1 kinase activity and gene expression regulator. *Elife* **6**, e23063 (2017).

101. Amick, J., Roczniak-Ferguson, A. & Ferguson, S. M. C9orf72 binds SMCR8, localizes to lysosomes, and regulates mTORC1 signaling. *Mol. Biol. Cell* **27**, 3040-3051 (2016).

102. Gupta, R. et al. The proline/arginine dipeptide from hexanucleotide repeat expanded *C9ORF72* inhibits the proteasome. *eNeuro* **4**, ENEURO.0249-16.2017 (2017).

103. Lee, S. et al. Activation of HIPK2 promotes ER stress-mediated neurodegeneration in amyotrophic lateral sclerosis. *Neuron* **91**, 41-55 (2016).

104. Zhang, Y. J. et al. Aggregation-prone c9FTD/ALS poly(GA) RAN-translated proteins cause neurotoxicity by inducing ER stress. *Acta Neuropathol.* **128**, 505-524 (2014).

105. May, S. et al. C9orf72 FTLD/ALS-associated Gly–Ala dipeptide repeat proteins cause neuronal toxicity and Unc119 sequestration. *Acta Neuropathol.* **128**, 485-503 (2014).

106. Yamakawa, M. et al. Characterization of the dipeptide repeat protein in the molecular pathogenesis of c9FTD/ALS. *Hum. Mol. Genet.* **24**, 1630-1645 (2015).

107. Guo, Q. et al. In situ structure of neuronal C9orf72 poly-GA aggregates reveals proteasome recruitment. *Cell* **172**, 696-705.e12 (2018).

108. Freibaum, B. D. et al. GGGGCC repeat expansion in *C9orf72* compromises nucleocytoplasmic transport. *Nature* **525**, 129-133 (2015).

109. Zhang, K. et al. The *C9orf72* repeat expansion disrupts nucleocytoplasmic transport. *Nature* **525**, 56-61 (2015).

110. Xiao, S. et al. Isoform-specific antibodies reveal distinct subcellular localizations of C9orf72 in amyotrophic lateral sclerosis. *Ann. Neurol.* **78**, 568-583 (2015).

111. Chou, C. C. et al. TDP-43 pathology disrupts nuclear pore complexes and nucleocytoplasmic transport in ALS/FTD. *Nat. Neurosci.* **21**, 228-239 (2018).

112. Shani, T., Levy, M. & Israelson, A. Assay to measure nucleocytoplasmic transport in real time within motor neuron-like NSC-34 cells. *J. Vis. Exp.* https://doi.org/10.3791/55676 (2017).

113. Shi, K. Y. et al. Toxic PRn poly-dipeptides encoded by the *C9orf72* repeat expansion block nuclear import and export. *Proc. Natl Acad. Sci. USA* **114**, E1111-E1117 (2017).

114. Khosravi, B. et al. Cytoplasmic poly-GA aggregates impair nuclear import of TDP-43 in *C9orf72* ALS/FTLD. *Hum. Mol. Genet.* **26**, 790-800 (2017).

115. Boeynaems, S. et al. *Drosophila* screen connects nuclear transport genes to DPR pathology in c9ALS/FTD. *Sci. Rep.* **6**, 20877 (2016).

116. Coyne, A. N. et al. Post-transcriptional inhibition of Hsc70-4/HSPA8 expression leads to synaptic vesicle cycling defects in multiple models of ALS. *Cell Rep.* **21**, 110-125 (2017).

117. Lopez-Gonzalez, R. et al. Poly(GR) in *C9ORF72*-related ALS/FTD compromises mitochondrial function and increases oxidative stress and DNA damage in iPSC-derived motor neurons. *Neuron* **92**, 383-391 (2016).

118. Walker, C. et al. *C9orf72* expansion disrupts ATM-mediated chromosomal break repair. *Nat. Neurosci.* **20**, 1225-1235 (2017).

119. Farg, M. A., Konopka, A., Ying Soo, K., Ito, D. & Atkin, J. D. The DNA damage response (DDR) is induced by the C9orf72 repeat expansion in amyotrophic lateral sclerosis. *Hum. Mol. Genet.* **26**, 2882-2896 (2017).

120. Schweizer Burguete, A. et al. GGGGCC microsatellite RNA is neuritically localized, induces branching defects, and perturbs transport granule function. *Elife* **4**, e08881 (2015).

121. Ishiguro, A., Kimura, N., Watanabe, Y., Watanabe, S. & Ishihama, A. TDP-43 binds and transports G-quadruplex-containing mRNAs into neurites for local translation. *Genes Cells* **21**, 466-481 (2016).

122. Blokhuis, A. M. et al. Comparative interactomics analysis of different ALS-associated proteins identifies converging molecular pathways. *Acta Neuropathol.* **132**, 175-196 (2016).

123. Yang, D. et al. FTD/ALS-associated poly(GR) protein impairs the Notch pathway and is recruited by poly(GA) into cytoplasmic inclusions. *Acta Neuropathol.* **130**, 525-535 (2015).

124. Devlin, A. C. et al. Human iPSC-derived motoneurons harbouring *TARDBP* or *C9ORF72* ALS mutations are dysfunctional despite maintaining viability. *Nat. Commun.* **6**, 5999 (2015).

125. Wainger, B. J. et al. Intrinsic membrane hyperexcitability of amyotrophic lateral sclerosis patient-derived motor neurons. *Cell Rep.* **7**, 1-11 (2014).

126. Sareen, D. et al. Targeting RNA foci in iPSC-derived motor neurons from ALS patients with a *C9ORF72* repeat expansion. *Sci. Transl. Med.* **5**, 208ra149 (2013).

127. Selvaraj, B. T. et al. *C9ORF72* repeat expansion causes vulnerability of motor neurons to Ca^{2+}-permeable AMPA receptor-mediated excitotoxicity. *Nat. Commun.* **9**, 347 (2018).

128. Schanz, O. et al. Cortical hyperexcitability in patients with *C9orf72* mutations: relationship to phenotype. *Muscle Nerve* **54**, 264-269 (2016).

129. Geevasinga, N. et al. Cortical function in asymptomatic carriers and patients with *C9orf72* amyotrophic lateral sclerosis. *JAMA Neurol.* **72**, 1268-1274 (2015).

130. Williams, K. L. et al. Pathophysiological insights into ALS with *C9ORF72* expansions. *J. Neurol. Neurosurg. Psychiatry* **84**, 931-935 (2013).

131. Benussi, A. et al. Impaired long-term potentiation-like cortical plasticity in presymptomatic genetic frontotemporal dementia. *Ann. Neurol.* **80**, 472-476 (2016).

132. Ciura, S. et al. Loss of function of C9orf72 causes motor deficits in a zebrafish model of amyotrophic lateral sclerosis. *Ann. Neurol.* **74**, 180-187 (2013).

133. Swaminathan, A. et al. Expression of *C9orf72*-related dipeptides impairs motor function in a vertebrate model. *Hum. Mol. Genet.* **27**, 1754-1762 (2018).

134. Swinnen, B. et al. A zebrafish model for C9orf72 ALS reveals RNA toxicity as a pathogenic mechanism. *Acta Neuropathol.* **135**, 427-443 (2018).

135. Baldwin, K. R., Godena, V. K., Hewitt, V. L. & Whitworth, A. J. Axonal transport defects are a common phenotype in *Drosophila* models of ALS. *Hum. Mol. Genet.* **25**, 2378-2392 (2016).

136. Sivadasan, R. et al. C9ORF72 interaction with cofilin modulates actin dynamics in motor neurons. *Nat. Neurosci.* **19**, 1610-1618 (2016).

137. Ferraiuolo, L. et al. Oligodendrocytes contribute to motor neuron death in ALS via SOD1-dependent mechanism. *Proc. Natl Acad. Sci. USA* **113**, E6496-E6505 (2016).

138. Tran, H. et al. Differential toxicity of nuclear RNA foci versus dipeptide repeat proteins in a *Drosophila* model of *C9ORF72* FTD/ALS. *Neuron* **87**, 1207-1214 (2015).

139. Rizzu, P. et al. *C9orf72* is differentially expressed in the central nervous system and myeloid cells and consistently reduced in *C9orf72*, *MAPT* and *GRN* mutation carriers. *Acta Neuropathol. Commun.* **4**, 37 (2016).

140. Xi, Z. et al. Hypermethylation of the CpG island near the G_4C_2 repeat in ALS with a *C9orf72* expansion. *Am. J. Hum. Genet.* **92**, 981-989 (2013).

141. Gijselinck, I. et al. A *C9orf72* promoter repeat expansion in a Flanders-Belgian cohort with disorders of the frontotemporal lobar degeneration–amyotrophic lateral sclerosis spectrum: a gene identification study. *Lancet Neurol.* **11**, 54-65 (2012).

142. Fratta, P. et al. Homozygosity for the *C9orf72* GGGGCC repeat expansion in frontotemporal dementia. *Acta Neuropathol.* **126**, 401-409 (2013).

143. Belzil, V. V. et al. Reduced *C9orf72* gene expression in c9FTD/ALS is caused by histone trimethylation, an epigenetic event detectable in blood. *Acta Neuropathol.* **126**, 895-905 (2013).

144. Waite, A. J. et al. Reduced C9orf72 protein levels in frontal cortex of amyotrophic lateral sclerosis and frontotemporal degeneration brain with the *C9ORF72* hexanucleotide repeat expansion. *Neurobiol. Aging* **35**, 1779.e5-1779.e13 (2014).

145. van Blitterswijk, M. et al. Novel clinical associations with specific *C9ORF72* transcripts in patients with repeat expansions in *C9ORF72*. *Acta Neuropathol.* **130**, 863-876 (2015).

146. Niblock, M. et al. Retention of hexanucleotide repeat-containing intron in *C9orf72* mRNA: implications for the pathogenesis of ALS/FTD. *Acta Neuropathol. Commun.* **4**, 18 (2016).

147. Belzil, V. V. et al. Characterization of DNA hypermethylation in the cerebellum of c9FTD/ALS patients. *Brain Res.* **1584**, 15-21 (2014).

148. Xi, Z. et al. The *C9orf72* repeat expansion itself is methylated in ALS and FTLD patients. *Acta Neuropathol* **129**, 715-727 (2015).

149. Esanov, R. et al. A *C9ORF72* BAC mouse model recapitulates key epigenetic perturbations of ALS/FTD. *Mol. Neurodegener.* **12**, 46 (2017).

150. Zeier, Z. et al. Bromodomain inhibitors regulate the *C9ORF72* locus in ALS. *Exp. Neurol.* **271**, 241-250 (2015).

151. Bauer, P. O. Methylation of *C9orf72* expansion reduces RNA foci formation and dipeptide-repeat proteins expression in cells. *Neurosci. Lett.* **612**, 204-209 (2016).

152. McMillan, C. T. et al. *C9orf72* promoter hypermethylation is neuroprotective: neuroimaging and neuropathologic evidence. *Neurology* **84**, 1622-1630 (2015).

153. Russ, J. et al. Hypermethylation of repeat expanded *C9orf72* is a clinical and molecular disease modifier. *Acta Neuropathol.* **129**, 39-52 (2015).

154. Yang, M. et al. A C9ORF72/SMCR8-containing complex regulates ULK1 and plays a dual role in autophagy. *Sci. Adv.* **2**, e1601167 (2016).

155. Freischmidt, A. et al. Haploinsufficiency of *TBK1* causes familial ALS and fronto-temporal dementia. *Nat. Neurosci.* **18**, 631-636 (2015).

156. Elden, A. C. et al. Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. *Nature* **466**, 1069-1075 (2010).

157. Ferguson, R., Serafeimidou-Pouliou, E. & Subramanian, V. Dynamic expression of the mouse orthologue of the human amyotropic lateral sclerosis associated gene *C9orf72* during central nervous system development and neuronal differentiation. *J. Anat.* **229**, 871-891 (2016).

158. Komatsu, M. et al. Loss of autophagy in the central nervous system causes neurodegeneration in mice. *Nature* **441**, 880-884 (2006).

159. Hara, T. et al. Suppression of basal autophagy in neural cells causes neurodegenerative disease in mice. *Nature* **441**, 885-889 (2006).

160. Robberecht, W. & Philips, T. The changing scene of amyotrophic lateral sclerosis. *Nat. Rev. Neurosci.* **14**, 248-264 (2013).

161. Fratta, P. et al. *C9orf72* hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes. *Sci. Rep.* **2**, 1016 (2012).

162. Reddy, K., Zamiri, B., Stanley, S. Y., Macgregor, R. B. Jr. & Pearson, C. E. The disease-associated r(GGGGCC)n repeat from the *C9orf72* gene forms tract length-dependent uni- and multimolecular RNA G-quadruplex structures. *J. Biol. Chem.* **288**, 9860-9866 (2013).

163. Zhou, B. et al. Characterizations of distinct parallel and antiparallel G-quadruplexes formed by two-repeat ALS and FTD related GGGGCC sequence. *Sci. Rep.* **8**, 2366 (2018).

164. Vatovec, S., Kovanda, A. & Rogelj, B. Unconventional features of *C9ORF72* expanded repeat in amyotrophic lateral sclerosis and frontotemporal lobar degeneration. *Neurobiol. Aging* **35**, 2421.e1-2421.e12 (2014).

165. Kovanda, A., Zalar, M., Sket, P., Plavec, J. & Rogelj, B. Anti-sense DNA d(GGCCCC)n expansions in *C9ORF72* form i-motifs and protonated hairpins. *Sci. Rep.* **5**, 17944 (2015).

166. Zhang, Y., Roland, C. & Sagui, C. Structure and dynamics of DNA and RNA double helices obtained from the GGGGCC and CCCCGG hexanucleotide repeats that are the hallmark of C9FTD/ALS diseases. *ACS Chem. Neurosci.* **8**, 578-591 (2017).

167. Zamiri, B. et al. Stress-induced acidification may contribute to formation of unusual structures in *C9orf72*-repeats. *Biochim. Biophys. Acta* **1862**, 1482-1491 (2018).

168. Haeusler, A. R., Donnelly, C. J. & Rothstein, J. D. The expanding biology of the *C9orf72* nucleotide repeat expansion in neurodegenerative disease. *Nat. Rev. Neurosci.* **17**, 383-395 (2016).

169. Lee, Y. B. et al. Hexanucleotide repeats in ALS/FTD form length-dependent RNA foci, sequester RNA binding proteins, and are neurotoxic. *Cell Rep.* **5**, 1178-1186 (2013).

170. Cooper-Knock, J. et al. Sequestration of multiple RNA recognition motif-containing proteins by *C9orf72* repeat expansions. *Brain* **137**, 2040-2051 (2014).

171. Cooper-Knock, J. et al. Antisense RNA foci in the motor neurons of *C9ORF72*-ALS patients are associated with TDP-43 proteinopathy. *Acta Neuropathol.* **130**, 63-75(2015).

172. Xu, Z. et al. Expanded GGGGCC repeat RNA associated with amyotrophic lateral sclerosis and frontotemporal dementia causes neurodegeneration. *Proc. Natl Acad. Sci. USA* **110**, 7778-7783 (2013).

173. Davidson, Y. S. et al. Heterogeneous ribonuclear protein A3 (hnRNP A3) is present in dipeptide repeat protein containing inclusions in frontotemporal lobar degeneration and

motor neurone disease associated with expansions in *C9orf72* gene. *Acta Neuropathol. Commun.* **5**, 31 (2017).

174. Conlon, E. G. et al. The *C9ORF72* GGGGCC expansion forms RNA G-quadruplex inclusions and sequesters hnRNP H to disrupt splicing in ALS brains. *Elife* **5**, e17820 (2016).

175. Mori, K. et al. hnRNP A3 binds to GGGGCC repeats and is a constituent of p62-positive/TDP43-negative inclusions in the hippocampus of patients with *C9orf72* mutations. *Acta Neuropathol.* **125**, 413-423 (2013).

176. Haas, S. et al. A 39-kD DNA-binding protein from mouse brain stimulates transcription of myelin basic protein gene in oligodendrocytic cells. *J. Cell Biol.* **130**, 1171-1179 (1995).

177. Gallia, G. L., Johnson, E. M. & Khalili, K. Puralpha: a multifunctional single-stranded DNA- and RNA-binding protein. *Nucleic Acids Res.* **28**, 3197-3205 (2000).

178. Ohashi, S. et al. The single-stranded DNA- and RNA-binding proteins pur alpha and pur beta link BC1 RNA to microtubules through binding to the dendrite-targeting RNA motifs. *J. Neurochem.* **75**, 1781-1790 (2000).

179. Kanai, Y., Dohmae, N. & Hirokawa, N. Kinesin transports RNA: isolation and characterization of an RNA-transporting granule. *Neuron* **43**, 513-525 (2004).

180. Celona, B. et al. Suppression of *C9orf72* RNA repeat-induced neurotoxicity by the ALS-associated RNA-binding protein Zfp106. *Elife* **6**, e19032 (2017).

181. Mackenzie, I. R. The role of dipeptide-repeat protein pathology in *C9orf72* mutation cases. *Neuropathol. Appl. Neurobiol.* **42**, 217-219 (2016).

182. Baborie, A. et al. Accumulation of dipeptide repeat proteins predates that of TDP-43 in frontotemporal lobar degeneration associated with hexanucleotide repeat expansions in *C9ORF72* gene. *Neuropathol. Appl. Neurobiol.* **41**, 601-612 (2015).

183. Mann, D. M. Dipeptide repeat protein toxicity in frontotemporal lobar degeneration and in motor neurone disease associated with expansions in *C9ORF72* — a cautionary note. *Neurobiol. Aging* **36**, 1224-1226 (2015).

184. Mattsson, N., Schott, J. M., Hardy, J., Turner, M. R. & Zetterberg, H. Selective vulnerability in neurodegeneration: insights from clinical variants of Alzheimer's disease. *J. Neurol. Neurosurg. Psychiatry* **87**, 1000-1004 (2016).

185. Mori, K. et al. Reduced hnRNPA3 increases *C9orf72* repeat RNA levels and dipeptide-repeat protein deposition. *EMBO Rep.* **17**, 1314-1325 (2016).

186. Proudfoot, M. et al. Early dipeptide repeat pathology in a frontotemporal dementia kindred with *C9ORF72* mutation and intellectual disability. *Acta Neuropathol.* **127**, 451-458 (2014).

187. Vatsavayai, S. C. et al. Timing and significance of pathological features in *C9orf72* expansion-associated frontotemporal dementia. *Brain* **139**, 3202-3216 (2016).

188. Mizielinska, S. et al. *C9orf72* repeat expansions cause neurodegeneration in *Drosophila* through arginine-rich proteins. *Science* **345**, 1192-1194 (2014).

189. Moens, T. G. et al. Sense and antisense RNA are not toxic in *Drosophila* models of *C9orf72*-associated ALS/FTD. *Acta Neuropathol.* **135**, 445-457 (2018).

190. Chang, Y. J., Jeng, U. S., Chiang, Y. L., Hwang, I. S. & Chen, Y. R. The glycine–alanine dipeptide repeat from *C9orf72* hexanucleotide expansions forms toxic amyloids possessing cell-to-cell transmission properties. *J. Biol. Chem.* **291**, 4903-4911 (2016).

191. Ohki, Y. et al. Glycine–alanine dipeptide repeat protein contributes to toxicity in a zebrafish model of *C9orf72* associated neurodegeneration. *Mol. Neurodegener.* **12**, 6 (2017).

192. Flores, B. N. et al. Distinct *C9orf72*-associated dipeptide repeat structures correlate with neuronal toxicity. *PLoS ONE* **11**, e0165084 (2016).

193. Prpar Mihevc, S. et al. Nuclear trafficking in amyotrophic lateral sclerosis and frontotemporal lobar degeneration. *Brain* **140**, 13-26 (2017).

194. Lin, Y. et al. Toxic PR poly-dipeptides encoded by the *C9orf72* repeat expansion target LC domain polymers. *Cell* **167**, 789-802.e12 (2016).

195. Lagier-Tourenne, C., Polymenidou, M. & Cleveland, D. W. TDP-43 and FUS/TLS: emerging roles in RNA processing and neurodegeneration. *Hum. Mol. Genet.* **19**, R46-R64 (2010).

196. Molliex, A. et al. Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* **163**, 123-133 (2015).

197. Patel, A. et al. A liquid-to-solid phase transition of the ALS protein FUS accelerated by disease mutation. *Cell* **162**, 1066-1077 (2015).

198. Murakami, T. et al. ALS/FTD mutation-induced phase transition of FUS liquid droplets and reversible hydrogels into irreversible hydrogels impairs RNP granule function. *Neuron* **88**, 678-690 (2015).

199. Jain, A. & Vale, R.D. RNA phase transitions in repeat expansion disorders. *Nature* **546**, 243-247 (2017).

200. Halliday, M. & Mallucci, G.R. Review: Modulating the unfolded protein response to prevent neurodegeneration and enhance memory. *Neuropathol. Appl. Neurobiol.* **41**, 414-427 (2015).

201. Bowden, H. A. & Dormann, D. Altered mRNP granule dynamics in FTLD pathogenesis. *J. Neurochem.* **138** (Suppl. 1), 112-133 (2016).

202. Tabet, R. et al. CUG initiation and frameshifting enable production of dipeptide repeat proteins from ALS/FTD *C9ORF72* transcripts. *Nat. Commun.* **9**, 152 (2018).

203. Miller, T. M. et al. An antisense oligonucleotide against *SOD1* delivered intrathecally for patients with *SOD1* familial amyotrophic lateral sclerosis: a phase 1, randomised, first-in-man study. *Lancet Neurol.* **12**, 435-442 (2013).

204. Zamiri, B., Reddy, K., Macgregor, R. B. Jr. & Pearson, C. E. TMPyP4 porphyrin distorts RNA G-quadruplex structures of the disease-associated r(GGGGCC)n repeat of the *C9orf72* gene and blocks interaction of RNA-binding proteins. *J. Biol. Chem.* **289**, 4653-4659 (2014).

205. Simone, R. et al. G-quadruplex-binding small molecules ameliorate *C9orf72* FTD/ALS pathology in vitro and in vivo. *EMBO Mol. Med.* **10**, 22-31 (2018).

206. Alniss, H., Zamiri, B., Khalaj, M., Pearson, C. E. & Macgregor, R. B. Jr. Thermodynamic and spectroscopic investigations of TMPyP4 association with guanine- and cytosine-rich DNA and RNA repeats of *C9orf72*. *Biochem. Biophys. Res. Commun.* **495**, 2410-2417 (2018).

207. Kramer, N. J. et al. Spt4 selectively regulates the expression of *C9orf72* sense and antisense mutant transcripts. *Science* **353**, 708-712 (2016).

208. Hu, J., Rigo, F., Prakash, T. P. & Corey, D. R. Recognition of *c9orf72* mutant RNA by single-stranded silencing RNAs. *Nucleic Acid Ther.* **27**, 87-94 (2017).

209. Pinto, B. S. et al. Impeding transcription of expanded microsatellite repeats by deactivated Cas9. *Mol. Cell* **68**, 479-490.e5 (2017).

210. Batra, R. et al. Elimination of toxic microsatellite repeat expansion RNA by RNA-targeting Cas9. *Cell* **170**, 899-912.e10 (2017).

211. Barry, A. E. et al. Alzheimer's disease brain-derived amyloid-β-mediated inhibition of LTP in vivo is prevented by immunotargeting cellular prion protein. *J. Neurosci.* **31**, 7259-7263 (2011).

212. Asuni, A. A., Boutajangout, A., Quartermain, D. & Sigurdsson, E. M. Immunotherapy targeting pathological tau conformers in a tangle mouse model reduces brain pathology with associated functional improvements. *J. Neurosci.* **27**, 9115-9129 (2007).

213. Boutajangout, A., Ingadottir, J., Davies, P. & Sigurdsson, E. M. Passive immunization targeting pathological phospho-tau protein in a mouse model reduces functional decline and clears tau aggregates from the brain. *J. Neurochem.* **118**, 658-667 (2011).

214. Masliah, E. et al. Effects of α-synuclein immunization in a mouse model of Parkinson's disease. *Neuron* **46**, 857-868 (2005).

215. Masliah, E. et al. Passive immunization reduces behavioral and neuropathological deficits in an alpha-synuclein transgenic model of Lewy body disease. *PLoS ONE* **6**, e19338 (2011).

216. Zhou, Q. et al. Antibodies inhibit transmission and aggregation of *C9orf72* poly-GA dipeptide repeat proteins. *EMBO Mol. Med.* **9**, 687-702 (2017).

217. Yang, Y. M. et al. A small molecule screen in stem-cell-derived motor neurons identifies a kinase inhibitor as a candidate therapeutic for ALS. *Cell Stem Cell* **12**, 713-726 (2013).

Acknowledgements

R.B. is a Leonard Wolfson Clinical Research Training Fellow and is funded by a Wellcome Trust Research Training Fellowship (107196/Z/14/Z). A.M.I. is funded by the Motor Neuron Disease Association, Alzheimer's Research UK, the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (648716 — C9ND) and the UK Dementia Research Institute. We thank Dr Sarah Mizielinska for reviewing the manuscript and assistance with figures, and Mr Martino Guadalupi and Dr Rachele Saccon for assistance with and design of figures.

Author contributions

Both authors researched data for the article, discussed the content, wrote the article, and reviewed and edited the manuscript before submission.

Competing interests statement

The authors declare no competing interests.

Publisher's note

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

Key points

- **Rapid progress has been made in the understanding how repeat expansions in C9orf72 cause C9FTD/ALS**
- **Both loss of function of C9orf72 and gain of toxic function of the repeats are implicated**

- **A range of new models including mice, *Drosophila* and patient neurons have provided new insights**
- **Several cellular pathways are affected and could provide new options for treatment**
- **Targeted therapeutic strategies against the repeats themselves are most advanced and progressing towards clinical trials**

Box 1 | Epigenetic modification of *C9orf72*

Analysis of blood, frontal cortex and cervical spinal cord DNA from patients with *C9orf72*-associated frontotemporal dementia and/or amyotrophic lateral sclerosis (C9FTD/ALS) has revealed that in 20–40% of cases, the CpG island in the *C9orf72* promoter region upstream of the pathogenic repeats is hypermethylated[22,140,147,148]. This hypermethylation is associated with increased repeat length and reduced transcription of *C9orf72.* Moreover, analysis of blood DNA from patients with C9FTD/ALS showed that in 97% of cases, the expanded hexanucleotide repeat itself was methylated[148]. A subset of *C9orf72* bacterial artificial chromosome transgenic mice also demonstrate hypermethylation of the *C9orf72* promoter and increasing hexanucleotide repeat methylation with age[149]. Treatment of C9FTD/ALS patient fibroblasts with a DNA and histone demethylating agent led to an increase in *C9orf72* transcript levels[143]. Bromodomain-containing proteins are involved in epigenetic regulation, and bromodomain inhibitors increase *C9orf72* expression in cells from patients with *C9orf72*-associated ALS and control individuals[150].

Histone trimethylation is another epigenetic modification that can reduce gene expression. Chromatin immunoprecipitation experiments in frontal cortex and cerebellar tissue showed that the *C9orf72* promoter region is bound to trimethylated histones in *C9orf72* repeat expansion carriers[143].

Considerable evidence points to amelioration of disease phenotypes by *C9orf72* hypermethylation. Hypermethylation is associated with reductions in RNA foci and dipeptide repeat proteins (DPRs) in patients with C9FTD/ALS[41] and *C9orf72* cell models[151], as well as reduced neuronal and grey matter loss[152]. In addition, hypermethylation has been linked to longer survival in patients

with *C9orf72*-associated FTD[153] and a later age of onset in ALS and FTD[22]. However hypermethylation also correlates with reduced disease duration before death in patients with *C9orf72*-associated ALS[140]. One possibility is that some phenotypes are dependent on gain of function and others on loss of function, so hypermethylation could have pleiotropic effects.

One might predict that reducing the levels of transcript variants 1 and 3 (Fig. 1) which are responsible for producing RNA foci and DPRs, would be protective. However, it is unclear whether reduction of variant 2 would be beneficial — this is the most highly expressed variant so is likely to be the main contributor to the functional pool of C9orf72 protein.

Box 2 | RNA-binding protein sequestration

RNA-binding proteins (RBPs) have diverse roles in splicing, translational regulation and RNA transport and degradation. In repeat-mediated diseases, possibly including *C9orf72*-associated disease, RBPs can become sequestered by RNA foci, leading to downstream consequences.

A number of RBPs have been shown to interact with *C9orf72* repeat RNA in human tissue[12,66,84,169-174], *in vitro* assays[66,84,162,170,172,174], cell models[66,70,79,169,170,172,174,175], induced pluripotent stem cell-derived neurons[12,66,126] and *in vivo* models[79,90,169,172]. The most frequently identified RBPs that interact with *C9orf72* repeat RNA are the heterogeneous nuclear ribonucleoprotein (hnRNPs), in particular, hnRNP H, although hnRNP A1 and hnRNP A3 are also detected. Other RBPs include ALYREF, ASF/SF2, ADARB2, nucleolin, Pur-α and SRSF2. These RBPs are not found consistently across different studies, possibly reflecting the diversity of models and methodologies. However, several of these proteins have been identified in multiple independent studies, indicating that specific RBPs can be sequestered by GGGGCC repeat RNA. Only a small subset of RNA foci seem to colocalize with RBPs in human tissue[170] and cells[79], indicating that sequestration of these proteins is a dynamic process, or that diffuse *C9orf72* repeat RNA that is not contained within foci can sequester RBPs.

Overexpression of Pur-α, an RBP that is involved in transcription regulation[176,177], mRNA localization[178,179] and stress granule formation[77] and was shown to interact with *C9orf72* repeat

RNA[79,172], can ameliorate neurodegeneration in *C9orf72* repeat-expressing *Drosophila* and neuronal cell lines[172]. Overexpression of Zfp106 also supresses toxiciy in *C9orf72* repeat *Drosophila*[180]. To date, direct evidence regarding the effects of RBP sequestration in tissue from patients with C9FTD/ALS is limited, although some hnRNP H targets are altered[86,174]. A crucial next step is to establish how the implicated RBPs and their targets are mechanistically linked to pathogenesis.

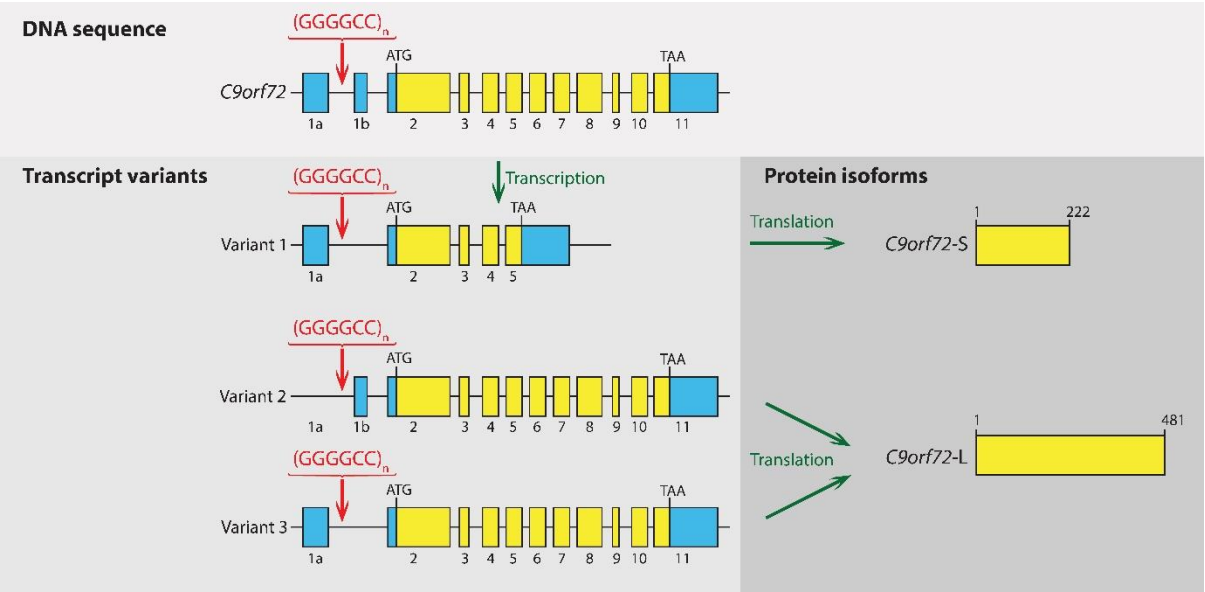

Figure 1 | *C9orf72* structure, transcript variants and protein isoforms. The *C9orf72* gene consists of 11 exons, has three main alternatively spliced transcript variants and produces two protein isoforms. In the figure, coding exons are indicated in yellow and noncoding exons in blue (not to scale). The GGGGCC hexanucleotide repeat expansion mutation is located in the first intron of variants 1 and 3 and within the promoter region of variant 2. Variant 1 encodes C9orf72-S (short), a 222-amino-acid protein of 24 kDa, and variants 2 and 3 encode C9orf72-L (long), a 481-amino-acid protein of 54 kDa

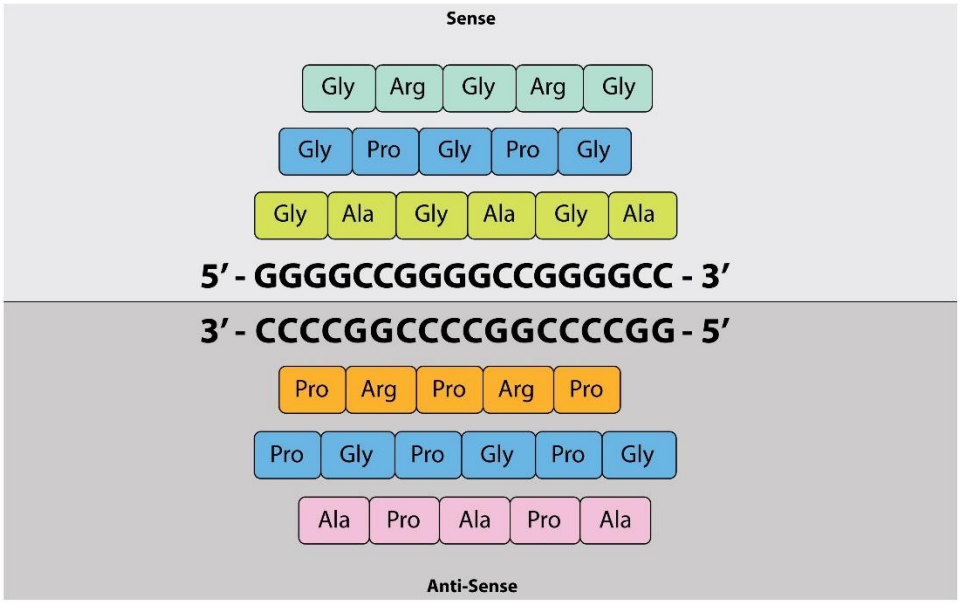

Figure 2 | Dipeptide repeat proteins. The figure shows the dipeptide repeat proteins that are generated by GGGGCC repeat-associated non-ATG (RAN) translation. The sense strand generates poly-GA, poly-GP and poly-GR and the antisense strand generates poly-GP, poly-PA and poly-PR.

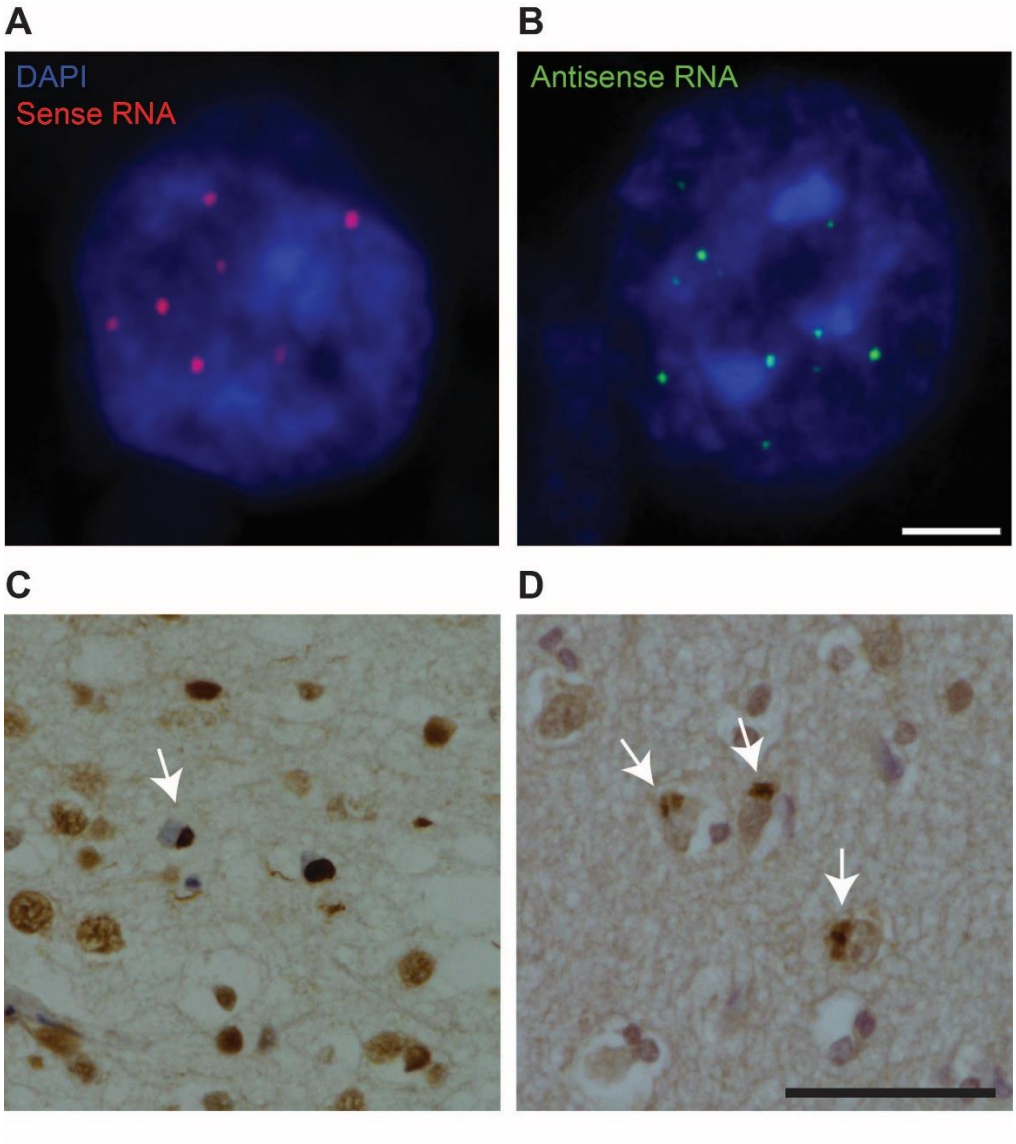

Figure 3 | C9FTD/ALS neuropathology. Sense and antisense RNA foci are a common feature in the brains of patients with *C9orf72*-associated frontotemporal dementia and/or amyotrophic lateral sclerosis (C9FTD/ALS). **a,b** |Representative images show neurons from the frontal cortex of a patient with C9FTD/ALS, containing multiple sense (red; part a) and antisense (green; part b) foci in nuclei (stained blue with DAPI). Scale bar: 2.5 µm. **c** | TAR DNA-binding protein 43 (TDP-43) pathology in a patient with C9FTD/ALS. Arrow indicates a neuronal cytoplasmic TDP-43 inclusion in the frontal cortex, with concomitant depletion of nuclear TDP-43. Scale bar: 50 µm. **d** | Dipeptide repeat protein (DPR) pathology in a patient with C9FTD/ALS. Inclusions consisting of sense and antisense DPRs are produced by repeat-associated non-ATG (RAN) translation. Arrows indicate neuronal cytoplasmic inclusions of poly-GA protein. Scale bar: 50 µm.

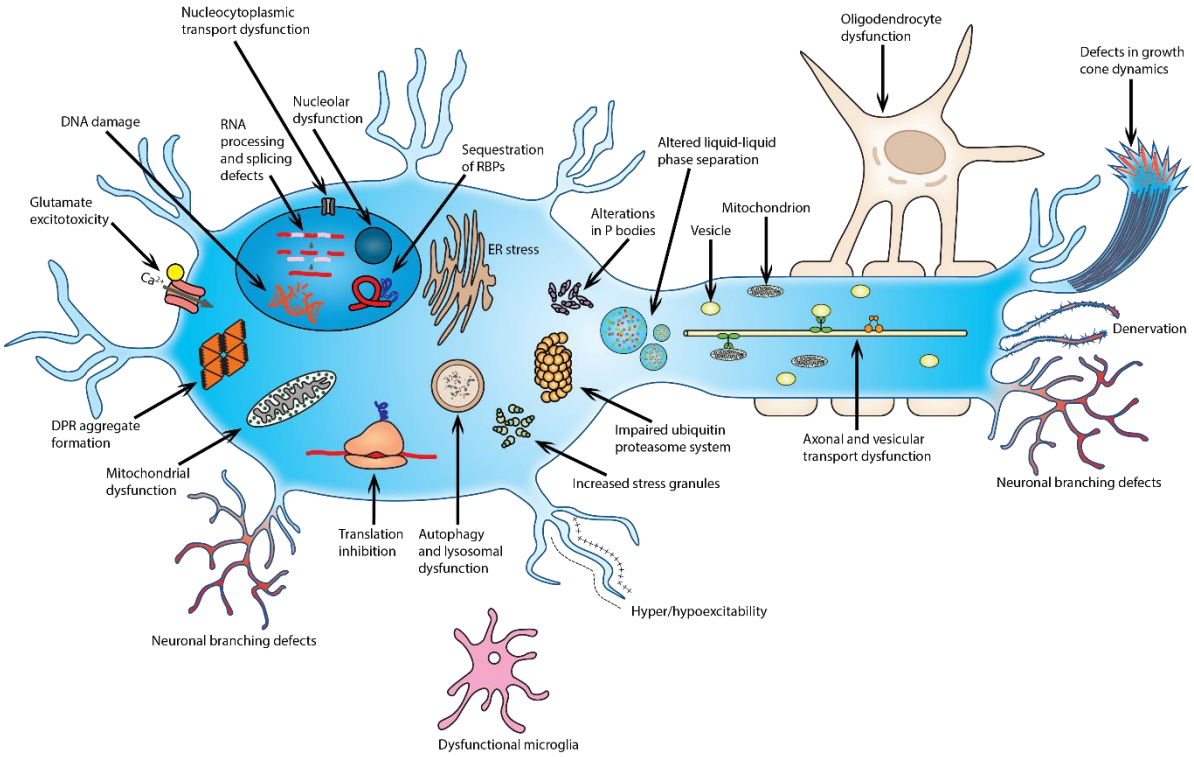

Figure 4 | Cellular processes implicated in *C9orf72*-associated FTD and ALS. A wide range of cellular pathways have been implicated in *C9orf72*-mediated disease, several of which have previously been linked to amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD). *C9orf72* loss-of-function and toxic gain-of-function mechanisms can both alter RNA processing and metabolism pathways, with alterations in stress granules and P-bodies, and *C9orf72* gain-of-function mechanisms can lead to nucleolar dysfunction, affect RNA splicing and transcription and cause DNA damage. Proteostasis pathways have also been implicated, with impairments in autophagy and lysosomal function, the unfolded protein response and the endoplasmic reticulum, and the ubiquitin–proteasome system. Other cellular processes including nucleocytoplasmic transport, vesicular trafficking and transport granule function, and mitochondrial function, can also be impaired. In addition, neuron-specific processes, including hyperexcitabiliy and hypoexcitability, glutamate excitotoxicity, axonal transport and neuronal branching defects, have been implicated in C9FTD/ALS. Finally, loss of *C9orf72* function alters immune system and microglial function (Table 3).

Table 1 | Mouse models of *C9orf72* loss of function

Study	Method(s)	Motor phenotypes in homozygotes	Cognitive and behavioural phenotypes in homozygotes	Other phenotypes in homozygotes	Survival in homozygotes
Clotilde Lagier-Tourenne et al. (2013)[13]	Somatic brain transgenesis with antisense oligonucleotide	Normal function	Normal function	None reported	Not reported
Panda et al. (2013)[49]	Non-conditional TALEN-mediated knockout	Not reported	Not reported	None reported	Not reported
Suzuki et al. (2013)[50]	Non-conditional knockout of exons 2–6	Not reported	Not reported	None reported	Not reported
Koppers et al. (2015)[51]	Conditional Cre–loxP-mediated knockout in neurons and glia	Normal function	Not reported	6% reduction in body weight in homozgotes compared with controls	Normal: oldest mice lived >24 months
Jiang et al. (2016)[52]	Non-conditional knockout of exons 2–6	Mild motor deficits on rotarod test	Mild social interaction and social recognition abnormalities	Reduced body weight, splenomegaly and cervical lymphadenopathy	Normal until 11 months, 7% of mice survived to 20 months
Atanasio et al. (2016)[53]	Non-conditional knockout of the full gene	Mild motor deficits at 40 weeks	Not reported	Lymphadenopathy at 12 months	9 of 17 mice survived to the end of the neurological assay period
O'Rourke et al. (2016)[54]	1. Non-conditional knockout of exons 2–6 2. Non-conditional knockout with zinc finger deletion	Normal function	Not reported	Cervical lymphadenopathy and splenomegaly	Normal lifespan
Sudria-Lopez et al. (2016)[55]	Not reported — full knockout in all tissues	Normal function	Not reported	Reduced body weight, lymphadenopathy and splenomegaly	Reduced survival: median lifespan ~500 days
Sullivan et al. (2016)[56]	CRISPR–Cas9-mediated non-conditional knockout	Not reported	Not reported	Lymphadenopathy and splenomegaly	Not reported
Burberry et al. (2016)[57]	1. Non-conditional knockout of exons 2–6 in a C57BL/6 inbred background (model 1) 2. Same as model 1, but on an outbred background (model 2) 3. CRISPR–Cas9-mediated non-conditional knockout (model 3)	Not reported	Not reported	Reduced body weight, splenomegaly, cervical lymphadenopathy and hepatomegaly	Model 1: 7% alive by 400 days Model 2: 64% alive by 300 days Model 3: reduced survival
Ugolino et al. (2017)[58]	Non-conditional knockout of exons 2–6	Not reported	Not reported	Splenomegaly and lethargy	Homozygotes: >50% dead in 600 days Heterozygotes: 20% dead in 600 days

A range of techniques to reduce expression of or knock out the mouse *C9orf72* orthologue have been used to investigate the normal function of C9orf72 protein. Homozygous *C9orf72* knockouts suggest roles for the protein in immune system function, autophagy and endosomal processes. TALEN, transcription activator-like effector nuclease.

Table 2 | Mouse models of *C9orf72* gain of function

Study	Mouse strain and methodology	RNA foci	DPRs detected by immunocytochemistry	DPRs detected by Immunoassay	TDP-43 pathology	Motor and cognitive phenotypes	Survival
Chew et al. (2015)[59]	C57BL/6J mice Somatic brain transgenesis: AAV-mediated expression of $(GGGGCC)_{66}$ repeats and 119 bp upstream and 100 bp downstream C9orf72 sequence	Sense foci throughout CNS	Poly-GA and poly-GP inclusions	Poly-GP expression in $(GGGGCC)_{66}$ mice	Nuclear and occasionally cytoplasmic pTDP-43 inclusions	Rotarod impairments from day 2 of testing onwards, plus anxiety-like behaviour and hyperactivity	Not reported
Peters et al. (2015)[60]	SJL/B6 mice BAC transgenic, 140.5 kb upstream *C9orf72* sequence, exons 1–5, and 300 or 500 repeats	Abundant sense foci throughout CNS, antisense foci more sparse	Poly-GP inclusions, increase with age	Poly-GP throughout brain, lower than in patients with C9FTD/ALS	No TDP-43 pathology	Normal rotarod and grip strength testing, normal social behaviour	Normal survival
O'Rourke et al. (2015)[61]	C57BL/6J mice BAC transgenic, 110 kb upstream *C9orf72* sequence, full gene and 100–1000 repeats, 20 kb downstream sequence	Sense and antisense RNA foci throughout CNS	Poly-GP inclusions, increase with age	Soluble and insoluble poly-GP, similar levels to patients with *C9orf72*-associatedFTLD	No TDP-43 pathology	Normal grip strength, rotarod and open field testing; normal behaviour	Not reported
Liu et al. (2016)[62]	FVB/NJ mice BAC transgenic, 52 kb upstream *C9orf72* sequence, full gene and different lines with repeat lengths and copy numbers up to 500 repeats, 19.4 kb downstream sequence	Sense and antisense RNA foci	Poly-GA aggregates throughout CNS, increase with age	Not reported	Nuclear and cytoplasmic TDP-43 aggregates	Acute rapidly progressive disease with motor phenotype; normal open field behaviour test	Mice with acute rapidly progressive disease have decreased survival
Jiang et al. (2016)[52]	C57BL/6 mice BAC transgenic, 140 kb upstream *C9orf72* sequence, exons 1–5 and ~110 repeats or ~450 repeats	Sense and antisense foci	Poly-GA, poly-GP and poly-GR inclusions	2% SDS-soluble poly-GP detected	Increased levels of pTDP-43, no mislocalization or aggregates	No motor deficits; spatial and working memory deficits and anxiety found	Not reported
Zhang et al. (2016)[63]	C57BL/6J mice Somatic brain transgenesis: AAV-mediated, GFP-$(GA)_{50}$ or GFP-$(GA)_{50}$-mutated	Not applicable	GFP-$(GA)_{50}$ inclusions	Poly-GA levels twofold higher than in $(GGGGCC)_{66}$ mice from the Chew et al. study[59]	Rare pTDP-43 inclusions	Motor phenotype, hyperactivity and anxiety	Not reported
Schludi et al. (2017)[64]	C57BL/6 mice Germline transgenesis: neuronal expression, $(GA)_{149}$-CFP + 31 carboxy-terminal amino acids	Not applicable	$(GA)_{149}$-CFP inclusions in brainstem, cerebellar nuclei and spinal cord,	Similar poly-GA levels in mouse spinal cord and motor	pTDP-43 levels higher in the urea-soluble fraction, as determined by	Progressive gait and balance deficits; muscle	Not reported

	from endogenous human locus		increase with age	neurons from patients with C9FTD/ALS	ELISA; no TDP-43 inclusions or mislocalization.	strength and spatial memory normal	
Herranz-Martin et al. (2017)[65]	C57BL/6J mice Somatic brain transgenesis: AAV-mediated expression of 10 pure or 102 interrupted (by TCGAG linker) GGGGCC repeats No *C9orf72* flanking sequence	Sense foci throughout CNS, less frequent in spinal cord	Poly-GA, primarily in cerebellum and brainstem	Not reported	Infrequent TDP-43 aggregates found equally in 10 pure or 102 interrupted repeats	Progressive gait and behavioural deficits on open field and novel object recognition	Not reported

Several mouse gain-of-function models are described. The approaches used have included somatic brain or germline transgenesis and expression of BAC constructs containing the full-length or partial human *C9orf72* gene and repeat expansion. AAV, adeno-associated virus; BAC, bacterial artificial chromosome; C9FTD/ALS, C9orf72-associated frontotemporal dementia and/or amyotrophic lateral sclerosis; CFP, cyan fluorescent protein; DPRs, dipeptide repeat proteins; ELISA, enzyme-linked immunosorbent assay; FTLD, frontotemporal lobar degeneration; GFP, green fluorescent protein; pTDP-43, phospho-TDP-43; SDS, sodium dodecyl sulphate; TDP-43, TAR DNA-binding protein 43.

Table 3 | Downstream mechanisms implicated in C9FTD/ALS

	C9orf72 human tissue/model			Other *C9orf72 in vitro* models		*C9orf72 in vivo* models			
	Patient-derived neurons /glia	Post-mortem tissue	Other system/ tissue/in patients	Cell lines	Primary neurons	Yeast	Worm and zebrafish	Fly	Mouse
RNA metabolism									
Nucleolar function/LLPS	Yes[66]	Yes[42,66,67]	Yes[42,66]	Yes[68-73]	Yes[36]	–	–	Yes[67,72]	–
Processing bodies	Yes[66]	–	–	Yes[74]	Yes[42,74]	–	–	–	–
Stress granules/LLPS	Yes[75]	Yes[76]	Yes[77]	Yes[69,72-74,78-83]	Yes[42,74]	–	–	Yes[72]	–
RNA processing, splicing, transcription and transport	Yes[84,85]	Yes[84-86]	–	Yes[68,84,85,87,88]	Yes[84]	Yes[89]	–	Yes[84]	–
Nuclear speckles/LLPS	–	–	–	Yes[72]	–	–	–	Yes[72]	–
Cajal bodies/LLPS	–	–	–	Yes[72]	–	–	–	Yes[72]	–
Proteostasis									
Autophagy and lysosomal function	Yes[75,90-94]	Yes[54]	Yes[93,95]	Yes[54,58,78,88,91,96-101]	Yes[91,96,102]	–	–	–	Yes[54,56,58]
Endoplasmic reticulum/UPR	Yes[66,75]	Yes[86,103,104]	–	Yes[80,83,88]	Yes[80,104]	–	–	–	–
Translational inhibition	–	–	–	Yes[79,80,87]	–	–	–	–	–
Ubiquitin-proteasome system	–	–	–	Yes[88,99,105,106]	Yes[63,102,104,107]	–	–	–	Yes[63]
Other cellular processes									
Nucleocytoplasmic transport	Yes[82,84,89,108,109]	Yes[86,109-111]	–	Yes[68,79,82,84,88,108,110-114]	Yes[84,89]	Yes[89]	–	Yes[82,84,108,109,115]	Yes[63]
Vesicle trafficking	Yes[93]	Yes[86]	Yes[93]	Yes[78,96]	–	–	–	Yes[116]	–
DNA damage	Yes[117]	Yes[118,119]	–	Yes[118,119]	Yes[118]	–	–	Yes[117]	–
Transport granule function	Yes[120]	Yes[120]	–	Yes[121]	Yes[120]	–	–	Yes[120]	–
Mitochondrial function	Yes[75,117]	–	Yes[95]	Yes[122]	–	–	–	–	–
Notch signalling	Yes[123]	Yes[123]	–	–	–	–	–	Yes[123]	–
Arginine methylation	–	–	–	Yes[115]	–	–	–	Yes[115]	–
Nervous system-specific processes									
Neuronal excitability and glutamate toxicity	Yes[12,94,124-127]	Yes[12,124-126]	Yes[128-131]	Yes[71]	–	–	–	–	Yes[94]
Neuronal branching and growth	–	Yes[105]	–	–	Yes[104,105,120]	–	Yes[132-134]	Yes[120,123]	–
Axonal transport	–	–	–	–	–	–	–	Yes[135]	–
Actin/growth cone dynamics	Yes[136]	–	–	–	Yes[136]	–	–	–	–
Glial dysfunction	Yes[137]	–	–	–	–	–	–	–	–
Systemic functions									
Immune system function	–	Yes[54]	–	Yes[54]	–	–	–	–	Yes (Table 1)

Several studies (referenced within the relevant cell of the table) have identified a spectrum of downstream processes involved in *C9orf72*-mediated disease, using patient-derived tissue, cells or induced pluripotent stem cell-derived neurons, and *in vitro* or *in vivo* disease models. Globally, these processes can be classified into those involved in RNA metabolism and processing, proteostasis, other cellular processes, nervous system-specific functions or systemic functions. LLPS, liquid–liquid phase separation; UPR, unfolded protein response.

Glossary

Repeat-associated non-ATG (RAN) translation
Translation is canonically dependent on an ATG start codon for initiation. RAN translation is a non-canonical form of translation that in the presence of repetitive sequences can start without the need for an ATG codon.

Hairpins
A secondary structure in which an RNA or DNA molecule folds back onto itself to resemble a hairpin.

G-quadruplexes
A secondary structure formed by guanine rich RNA or DNA molecules consisting of a stack of G-quartets (four guanine residues aligned in a square planar configuration).

i-motifs
A four-stranded secondary structure formed by cytosine-rich DNA or RNA molecules.

Frontotemporal lobar degeneration
Frontotemporal lobar degeneration (FTLD) describes the pathological findings observed in patients with frontotemporal dementia (FTD), however FTLD and FTD are also often used interchangeably to describe the clinical syndrome.

Bacterial artificial chromosome (BAC)
A vector for maintaining large pieces of DNA, often 50-200 kilobases in size.

Cryo-electron tomography
A high-resolution technique that involves collecting a series of tilted images of frozen hydrated samples using an electron microscope to produce a £D reconstruction of the sample

P-bodies
Processing bodies – membraneless organelles within the cytoplasm that are involved in translational repression of mRNAs and mRNA silencing and degradation