



RNA Misprocessing in *C9orf72*-Linked Neurodegeneration

Holly V. Barker*, Michael Niblock, Youn-Bok Lee, Christopher E. Shaw and Jean-Marc Gallo

Department of Basic and Clinical Neuroscience, Maurice Wohl Clinical Neuroscience Institute, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom

OPEN ACCESS

Edited by:

Ho Yin Edwin Chan,
The Chinese University of Hong
Kong, Hong Kong

Reviewed by:

Chris W. Lee,
Atlantic Health System/Biomedical
Research Institute of New Jersey,
United States

Katarzyna Gaweda-Walerych,
Mossakowski Medical Research
Centre, Poland
Kim A. Staats,
University of Southern California,
United States

*Correspondence:

Holly V. Barker
holly.v.barker@kcl.ac.uk

Received: 07 February 2017

Accepted: 21 June 2017

Published: 11 July 2017

Citation:

Barker HV, Niblock M, Lee Y-B,
Shaw CE and Gallo J-M (2017) RNA
Misprocessing in *C9orf72*-Linked
Neurodegeneration.
Front. Cell. Neurosci. 11:195.
doi: 10.3389/fncel.2017.00195

A large GGGGCC hexanucleotide repeat expansion in the first intron or promoter region of the *C9orf72* gene is the most common genetic cause of familial and sporadic Amyotrophic lateral sclerosis (ALS), a devastating degenerative disease of motor neurons, and of Frontotemporal Dementia (FTD), the second most common form of presenile dementia after Alzheimer's disease. *C9orf72*-associated ALS/FTD is a multifaceted disease both in terms of its clinical presentation and the misregulated cellular pathways contributing to disease progression. Among the numerous pathways misregulated in *C9orf72*-associated ALS/FTD, altered RNA processing has consistently appeared at the forefront of *C9orf72* research. This includes bidirectional transcription of the repeat sequence, accumulation of repeat RNA into nuclear foci sequestering specific RNA-binding proteins (RBPs) and translation of RNA repeats into dipeptide repeat proteins (DPRs) by repeat-associated non-AUG (RAN)-initiated translation. Over the past few years the true extent of RNA misprocessing in *C9orf72*-associated ALS/FTD has begun to emerge and disruptions have been identified in almost all aspects of the life of an RNA molecule, including release from RNA polymerase II, translation in the cytoplasm and degradation. Furthermore, several alterations have been identified in the processing of the *C9orf72* RNA itself, in terms of its transcription, splicing and localization. This review article aims to consolidate our current knowledge on the consequence of the *C9orf72* repeat expansion on RNA processing and draws attention to the mechanisms by which several aspects of *C9orf72* molecular pathology converge to perturb every stage of RNA metabolism.

Keywords: *C9orf72*, amyotrophic lateral sclerosis, frontotemporal dementia, RNA, repeats, splicing

INTRODUCTION

Amyotrophic lateral sclerosis (ALS) is a fatal progressive neuromuscular disease resulting from the degeneration of motor neurons in the brain and spinal cord for which no effective treatments are currently available. Frontotemporal Dementia (FTD), the second most common form of dementia after Alzheimer's disease, is characterized by behavioral and language deficits and manifests pathologically by neuronal atrophy in the frontal and anterior temporal lobes in the brain. A hexanucleotide G₄C₂ repeat expansion in the first intron or promoter region of the *C9orf72* gene on chromosome 9p21 represents the most common familial cause of ALS and FTD (Dejesus-Hernandez et al., 2011; Renton et al., 2011). *C9orf72* repeat expansions manifest pathologically as proteinaceous inclusions of the RNA/DNA binding protein TDP-43. *C9orf72* repeat expansion as a common genetic cause of ALS of FTD further

emphasizes the extensive clinical, genetic and pathological overlap between these two conditions, suggesting that both diseases represent opposite ends of a continuous clinical spectrum. *C9orf72* lies at the heart of this spectrum, with the risk of developing FTD and ALS concurrently increasing from 7% to 30% in expansion carriers (van der Zee et al., 2013). The number of repeats in the normal population ranges from 2 to 23. Expansion sizes vary considerably between individual cases and expansion sizes as high as 3500 repeats have been reported (Chen et al., 2016), however the lower threshold required to initiate pathogenesis is poorly defined. A report described an individual with an intermediate repeat length (30 units) exhibiting some *C9orf72*-associated pathological phenotypes in the absence of clinical manifestations (Gami et al., 2015) suggesting that 30 repeats may lie on the border between initial cellular abnormalities and full-blown disease. Expansions of less than 30 repeats are not typically associated with disease; however a minority of ALS cases with 20–22 repeats have been described (Byrne et al., 2014). Furthermore, such intermediate repeat sizes significantly increase the risk of developing ALS (Chen et al., 2016) and are associated with decreased *C9orf72* promoter activity (Gijssels et al., 2016).

Differential use of transcription alternative start and termination sites generates three RNA transcripts from *C9orf72* DNA. These encode two protein isoforms consisting of a long isoform (isoform A) of approximately 54 kDa derived from variants 2 (NM_018325.4) and 3 (NM_001256054.2), and a short isoform (isoform B) of approximately 24 kDa derived from variant 1 (NM_145005.6; **Figure 1**). Several studies have shown a reduction in *C9orf72* mRNA (Dejesus-Hernandez et al., 2011; Renton et al., 2011; Gijssels et al., 2012) and protein levels (Waite et al., 2014) in expansion carriers, suggesting that toxicity may be mediated by loss of function of the *C9orf72*-encoded protein. However, neural-specific ablation of *C9orf72* or knockdown using antisense oligonucleotides (ASOs) in mice has failed to recapitulate pathology (Lagier-Tourenne et al., 2013; Koppers et al., 2015; Atanasio et al., 2016; Jiang et al., 2016; O'Rourke et al., 2016). *C9orf72* knockdown in murine models produces an altered immune response (Atanasio et al., 2016) characterized by the accumulation of lysosomal vesicles within macrophages, implicating a role for the *C9orf72* protein in the regulation of late endosomal/lysosomal trafficking in macrophages and microglia (O'Rourke et al., 2016). The *C9orf72* protein functions as a guanine exchange factor (GEF) for Rab GTPases to regulate vesicular trafficking and autophagy (Webster et al., 2016). Several independent reports have shown that the *C9orf72* long isoform forms a complex with WDR41 and SMCR8 proteins, both regulators of autophagy (Blokhuys et al., 2016; Sellier et al., 2016; Sullivan et al., 2016; Xiao et al., 2016).

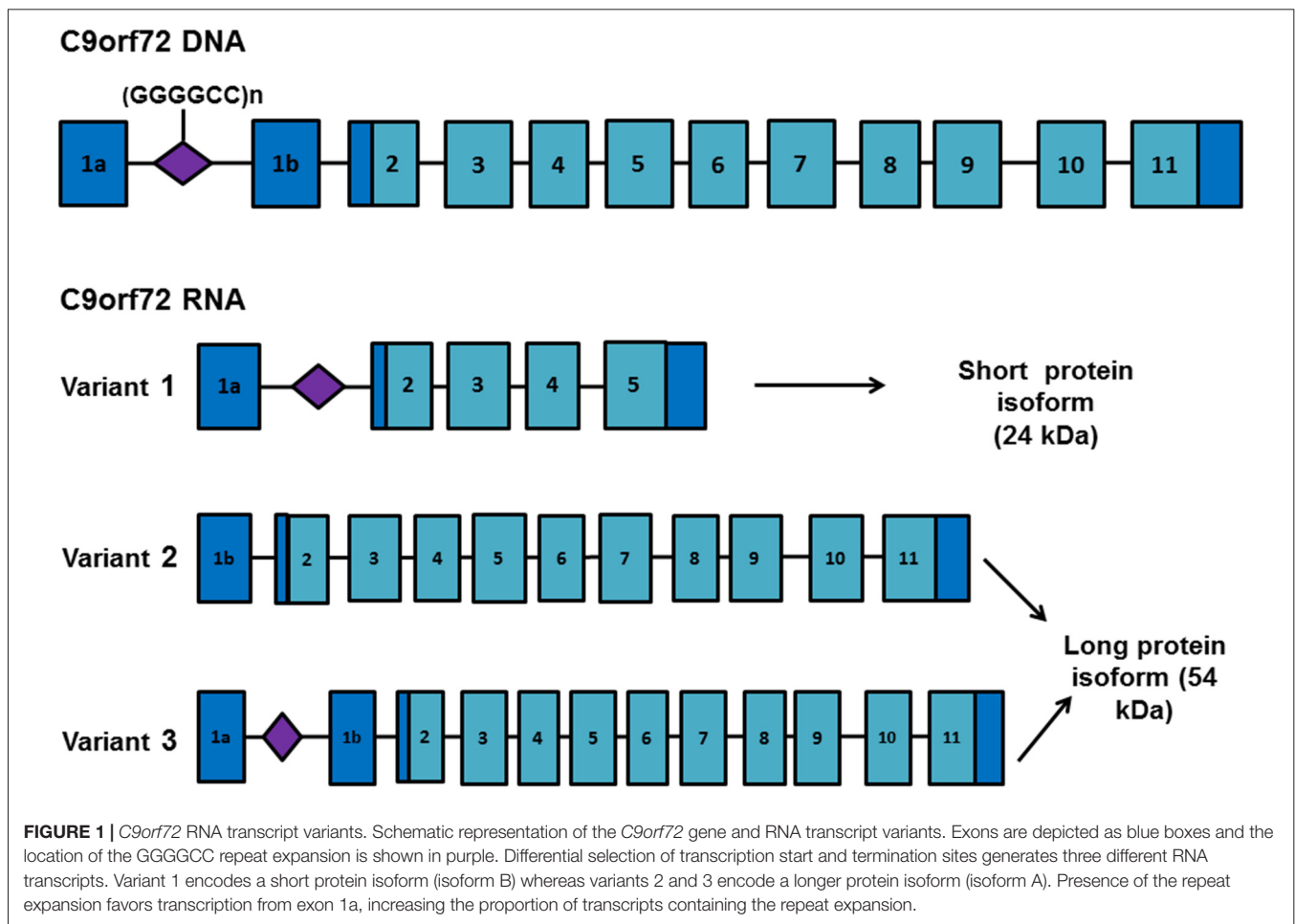
It is becoming increasingly apparent that altered RNA processing plays a key role in *C9orf72*-mediated toxicity through two separate, albeit related, aspects of RNA processing (**Figure 2**). The first is altered processing of the expanded *C9orf72* transcript itself, in terms of altered transcription, splicing defects, nuclear aggregation and non-conventional translation. The second represents downstream and indirect

changes in RNA processing of other transcripts. Far from being disparate entities in their role in *C9orf72* toxicity, we highlight how RNA and protein frequently interact and drive disturbances in RNA metabolism. In this review article we describe aspects of the pathomechanism of *C9orf72*-associated ALS/FTD linked to disturbances at the RNA level and how these changes may drive further RNA processing abnormalities.

TDP-43 IS AN ESSENTIAL MEDIATOR OF RNA METABOLISM

A characteristic pathological hallmark of ~90% of ALS and ~60% of FTD cases, including c9ALS/FTD cases is the presence of cytoplasmic, ubiquitin-positive, inclusions of the DNA/RNA binding protein, TDP-43, in affected neurons (Neumann et al., 2006; Mackenzie et al., 2014). In disease, TDP-43 aggregates in the cytoplasm where it undergoes a series of post-translational modifications, including ubiquitination, phosphorylation and C-terminal cleavage (Neumann et al., 2006). TDP-43 is mainly a nuclear protein and its mislocalization to the cytoplasm and aggregation are associated with a reduction and subsequent loss of function. Mislocalization of TDP-43 from the nucleus to the cytoplasm is likely to be a key mediator of pathogenesis, as mutations in *TARDBP*, the gene encoding TDP-43, are causative of some forms of familial ALS (Kabashi et al., 2008; Sreedharan et al., 2008).

TDP-43 is a highly conserved RNA-binding protein (RBP), possessing two RNA recognition motif (RRM) domains. TDP-43 regulates splicing, RNA turnover and transport and therefore interacts with a myriad of proteins involved in RNA metabolism (Buratti and Baralle, 2010; Blokhuys et al., 2016). Downregulation of TDP-43 expression in the CNS alters the expression of ~600 transcripts and splicing of over 900 genes, the majority of which are implicated in synaptic transmission, neuronal development and RNA metabolism (Polymenidou et al., 2011; Tollervey et al., 2011). Depletion of TDP-43 reduces the expression of long intron-containing transcripts, the majority of these essential for maintaining neuronal integrity (Lagier-Tourenne et al., 2012). ALS-causing TDP-43 mutations can also disrupt the splicing pattern of its mRNA targets in the absence of TDP-43 mislocalization and aggregation, resulting in abnormal splicing of transcripts involved in neurological function (Arnold et al., 2013). One function of TDP-43 is that it serves to repress splicing of nonconserved cryptic exons (Ling et al., 2015). Cryptic exons are stretches of nucleotides surrounded by sequences similar to authentic splice sites but which are not normally recognized as such so are not spliced into the wild-type RNA. Inclusion of a cryptic exon often results in the introduction of a premature stop codon, resulting in either protein truncation or mRNA degradation through nonsense-mediated decay. Cryptic exon inclusion occurs in Alzheimer's brain exhibiting TDP-43 nuclear clearance in the absence of inclusion formation, suggesting that cryptic exon inclusion occurs as an early event among TDP-43 proteinopathies (Sun et al., 2017).



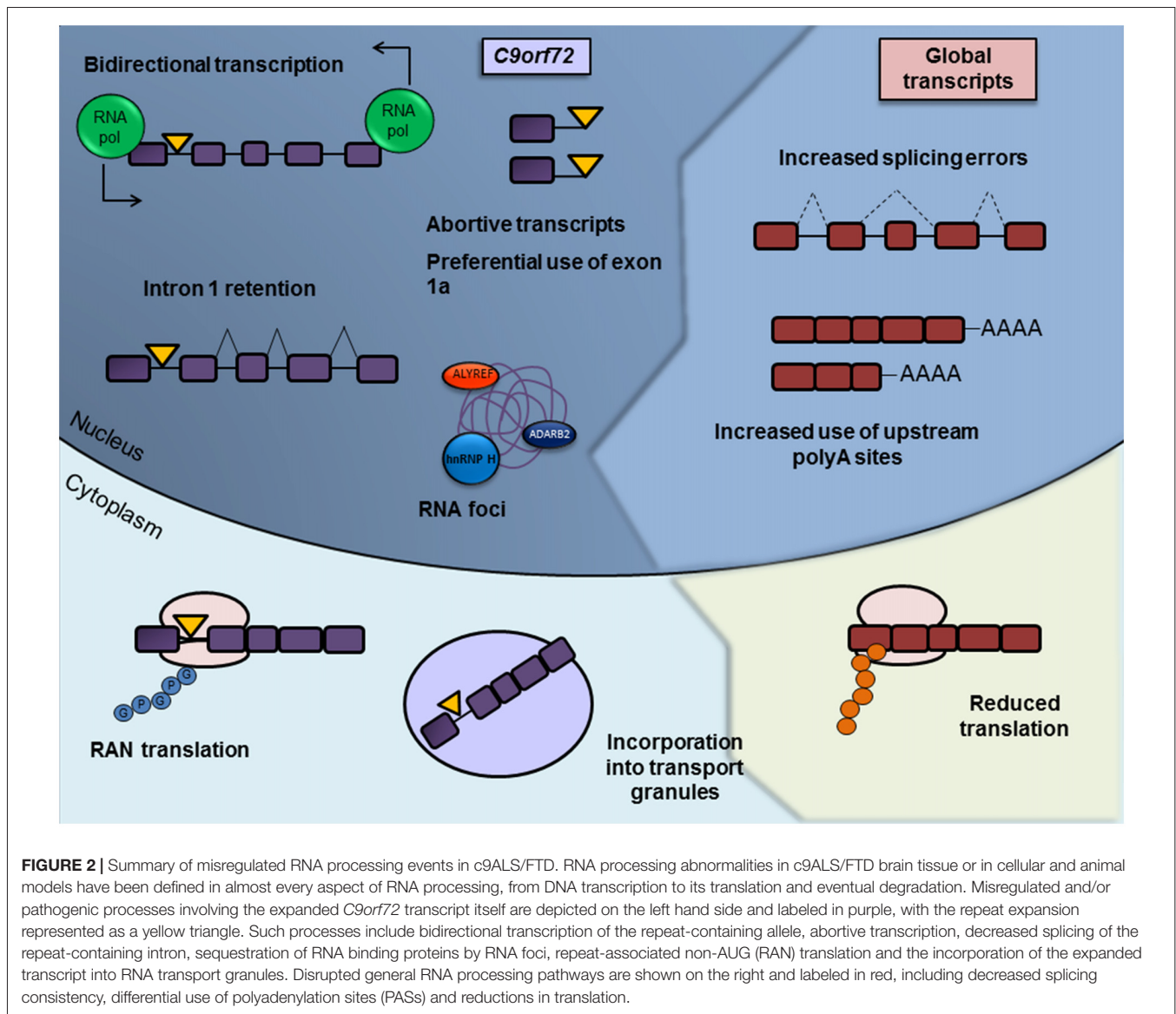
TDP-43 not only serves to regulate coding RNAs but is also implicated in the regulation of non-coding RNA expression, including microRNA (miRNA) and long noncoding RNA (lncRNA). TDP-43 associates with both microprocessor complexes and its depletion results in reduced binding efficiency of both complexes to a subset of miRNAs (Kawahara and Mieda-Sato, 2012). In addition to facilitating miRNA biogenesis on a global level, TDP-43 can directly bind miRNAs or their precursors, including let-7b, miR-663 (Buratti et al., 2010) and miR-9 (Zhang et al., 2013). TDP-43 depletion results in downregulation of let-7b which in turn is associated with altered expression levels of several let-7b targets. (Buratti et al., 2010). Included in these targets are STX3 and VAMP3 which are involved in synaptic vesicle formation and synaptic exocytosis, respectively (Darios and Davletov, 2006; Deák et al., 2006). Another transcript affected by let-7b downregulation is DYRK1A, upregulation of which is associated with neuronal deficits in individuals with Down syndrome (Lepagnol-Bestel et al., 2009).

lncRNAs make up an extensive family of long (several kb) RNA molecules which serve key regulatory roles and have been implicated in numerous cellular pathways. In cases of sporadic frontotemporal lobar degeneration with TDP-43 pathology, TDP-43 exhibits significantly increased binding to

MALAT1 and *NEAT1* lncRNAs (Tollervey et al., 2011). *MALAT1* influences the phosphorylation patterns of the SR-family of splicing proteins and recruits splicing factors to nuclear speckles (Tripathi et al., 2010). *NEAT1* is responsible for paraspeckle assembly (Clemson et al., 2009), nuclear domains implicated in transcriptional regulation, pre-mRNA splicing and mRNA nuclear retention. Downregulation of TDP-43 results in a corresponding decrease in *MALAT1* expression (Guo et al., 2015), suggesting that loss of TDP-43 nuclear function could result in splicing abnormalities through altered expression of *MALAT1* lncRNA.

Furthermore, TDP-43 facilitates transport and local translation of its mRNA targets within distal neuronal regions. In *Drosophila* motor neurons, mouse cortical neurons and motor neurons derived from human induced pluripotent stem cells (iPSCs), TDP-43 associates with bidirectionally transported messenger ribonucleoprotein complex (mRNP) granules. In neuronal cells harboring ALS-causing TDP-43 mutations, the anterograde transport of mRNP granules is selectively impaired (Alami et al., 2014), suggesting that impaired axonal transport of TDP-43 target mRNAs may contribute to ALS-associated pathogenesis.

Due to the multitudinous roles of TDP-43 in the metabolism of not only messenger RNAs, but also miRNAs and lncRNAs, it is



perhaps unsurprising that TDP-43 proteinopathies present with disruptions in RNA at so many levels. Defective RNA processing may represent a common pathway connecting FTD and ALS, although the mechanism by which TDP-43 nuclear clearance and aggregation results in such clinical heterogeneity remains ambiguous.

RNA TOXICITY: THE ROLE OF RNA-BINDING PROTEIN SEQUESTRATION IN ALTERED RNA PROCESSING EVENTS

C9orf72 repeat RNA is bidirectionally transcribed, from both the sense $(G_4C_2)_n$ and the antisense $(G_2C_4)_n$ DNA strand (Zu et al., 2013), a phenomenon which appears to be more widespread than previously recognized (Pelechano and Steinmetz, 2013). Repeat RNA forms punctate nuclear aggregates termed RNA foci. Multiple nuclear, and more rarely, cytoplasmic, sense

and antisense RNA foci have been identified in patient-derived cells and throughout the CNS in expansion carriers (Gendron et al., 2013; Lagier-Tourenne et al., 2013; Mizielinska et al., 2013; Zu et al., 2013). Analyses of the distribution and abundance of RNA foci initially revealed a significant correlation with clinical phenotype (Cooper-Knock et al., 2014) and an inverse correlation with age of onset (Mizielinska et al., 2013), implicating RNA foci formation as a key driver of *C9orf72* toxicity. However, extensive characterization in a larger cohort found no association between RNA foci distribution and clinical phenotype. RNA foci load is highest in cerebellar Purkinje cells, the loss of which has not been described in *C9orf72* expansion carriers. Interestingly, a higher percentage of antisense foci in the frontal cortex appear to correlate with a later age of onset, particularly in patients diagnosed with FTL (DeJesus-Hernandez et al., 2017).

BOX 1 | The mechanism of repeat-associated non-AUG (RAN) translation.

The type of repeat-containing RNA species that are exported to the cytoplasm and subject to RAN translation is still to be elucidated. However several recent results support the notion that the substrate for RAN translation of G_4C_2 repeats into dipeptide repeat proteins (DPRs) is a mature mRNA. First G_4C_2 repeats have been shown to cause overt neurotoxicity in *Drosophila* only if there are in the context of an mRNA (Tran et al., 2015). Second, initiation of RAN translation at CGG repeats, causing Fragile X tremor ataxia syndrome (FXTAS), requires a 5' 7-methylguanosine (m^7G) cap (Kearse et al., 2016). Addition of an m^7G cap is part of the final processing of mRNA. In patient tissues, the repeat-containing intron has been shown to be retained in a proportion of polyadenylated *C9orf72* RNA species in which downstream exons were spliced correctly resulting in a *C9orf72* mRNA with an enlarged 5'-untranslated region (5'-UTR) containing the G_4C_2 repeat domain (Niblock et al., 2016). These species would be exported to the cytoplasm through the conventional pathway of mRNA export and have a 5' m^7G cap and are therefore the prime candidates for the template for RAN translation of G_4C_2 repeats into DPRs.

Finally, the study of Kearse et al. (2016) revealed that the mechanism of RAN translation may more closely resemble the canonical mode of translation than previously thought. The canonical mode of translation initiation occurs by recruitment of various initiation factors to the 5' m^7G cap, which scan the mRNA in search of an AUG initiation codon. In the case of microsatellite repeat disorders in which translation occurs in the absence of an AUG codon, RAN translation may occur in a similar fashion to internal ribosomal entry site (IRES)-mediated translation initiation (Chiang et al., 2001), in which secondary structures directly recruit initiation factors. Indeed, previous studies have identified a requirement for the formation of hairpin secondary structures to initiate RAN translation, as repeat sequences unable to form such secondary structures are not subject to RAN translation (Zu et al., 2011). However, RAN translation of CGG repeats is dependent on the m^7G cap and eIF4E and eIF4A initiation factors, utilizing a scanning mechanism to initiate translation at codons which slightly deviate from the AUG sequence, located upstream of the repeats or within the repeat sequence itself. The secondary structure formed by CGG repeats may stall the initiation complex as it scans the mRNA transcript, resulting in the utilization of near-AUG sequences upstream of the repeats as alternative initiation codons (Kearse et al., 2016).

Since the discovery of RNA foci in c9ALS/FTD patient tissue, RNA toxicity has emerged as a leading hypothesis for *C9orf72*-associated pathogenesis. The RNA toxicity hypothesis postulates that G_4C_2 repeats accumulate into nuclear RNA foci and sequester essential RBPs, impairing their ability to regulate their RNA targets and culminating in a range of RNA misprocessing events. RNA pull-down and colocalization experiments in c9ALS brain and neurons derived from iPSCs (iPSNs) have identified a panel of RBPs binding to G_4C_2 repeat RNA, including hnRNP A1, hnRNP A3, hnRNP H, ADARB2, Pur- α , ASF/SF2, ALYREF and nucleolin (Donnelly et al., 2013; Lee et al., 2013; Mori et al., 2013a; Sareen et al., 2013; Xu et al., 2013; Cooper-Knock et al., 2014; Haeusler et al., 2014). Sense and antisense RNA foci sequester a similar panel of RBPs, specifically SC35, hnRNP K, hnRNP A1, ALYREF and hnRNP H (Cooper-Knock et al., 2014). Interestingly, only antisense foci demonstrate a significant positive correlation with TDP-43 mislocalization in motor neurons of c9ALS patients (Cooper-Knock et al., 2015b), which in turn is associated with neurodegeneration (Davidson et al., 2016). ASOs targeting the *C9orf72* transcript downstream of the repeats reduce RNA foci levels (Donnelly et al., 2013; Jiang et al., 2016), attenuate

sequestration of specific RBPs and normalize gene expression changes (Donnelly et al., 2013), suggesting that transcriptome alterations represent a direct downstream consequence of G_4C_2 RNA toxicity. Upregulation of *C9orf72* mRNA expression is associated with concomitant downregulation of genes enriched for functions in RNA metabolism, such as genes encoding tRNA synthetases (Nataf and Pays, 2015) emphasizing the role of tRNA metabolism in motor neuron degeneration (Weitzer et al., 2015). Such downregulation of genes involved in the regulation of RNA metabolism is consistent with RBP sequestration by the *C9orf72* hexanucleotide repeat expansion, as the reduced pool of RBPs is less able to promote the expression of such transcripts.

DPR ACCUMULATION IS DETRIMENTAL TO NUMEROUS CELLULAR PATHWAYS

Sense and antisense *C9orf72* RNA repeats are translated by an unconventional form of translation, repeat-associated non-AUG initiated translation (RAN translation) (Box 1). RAN translation was first described in the context of spinocerebellar ataxia type 8 (SCA8) caused by a CTG repeat expansion. RAN translation of $(G_4C_2)_n(C_4G_2)_n$ results in the synthesis of five dipeptide repeat protein (DPR) species; poly-GA, poly-GP, poly-GR, from the sense transcript and poly-GP, poly-PR and poly-PA, from the antisense transcript. All five DPR species form ubiquitinated inclusions in cerebellum, hippocampus and other brain regions in expansion carriers (Ash et al., 2013; Gendron et al., 2013; Mann et al., 2013). Several studies have demonstrated the toxicity of DPRs *in vitro* and *in vivo* (Kwon et al., 2014; Mizielinska et al., 2014; Wen et al., 2014; Zhang et al., 2014). Poly-GA expression in primary neurons reduces dendritic branching, increases apoptosis through caspase-3 activation (May et al., 2014) and induces ER stress (Zhang et al., 2014). The arginine-rich DPRs, poly-GR and poly-PR, exhibit the most robust toxicity in transgenic fly models and motor neuron cultures (Mizielinska et al., 2014; Wen et al., 2014). Expression of $(PR)_{50}$ and $(GR)_{50}$ dramatically decreases survival of primary cortical and motor neurons, presumably through a pathway of translational dysregulation (Wen et al., 2014). Poly-PR and poly-GR irreversibly bind nucleoli and can alter the splicing patterns of specific RNAs. For instance, when applied to cultured human astrocytes $(PR)_{20}$ penetrates cells and causes exon-skipping in RAN and *PTX3* RNA, resulting in the production of mRNAs encoding truncated proteins and in impaired biogenesis of ribosomal RNA, suggestive of nucleolar dysfunction (Kwon et al., 2014). A yeast genetic screen demonstrated that deletion of specific genes involved in ribosomal RNA processing, such as *NOB1* and *NSR1*, suppressed poly-PR toxicity (Jovičić et al., 2015). Consistent with these findings, a separate study identified nuclear inclusions of poly-PR and poly-GR tightly bound to nucleolar proteins, resulting in enlargement of the nucleus and induction of apoptosis (Wen et al., 2014).

While analyses of DPR interactomes have revealed preferential interactions between poly-GA and proteins involved

in proteasomal degradation (May et al., 2014; Zhang et al., 2016), the arginine-rich DPRs bind to numerous interactors enriched in RBPs and components of membrane-less organelles (Lee et al., 2016; Lin et al., 2016; Boeynaems et al., 2017). A significant number of these proteins possess low complexity sequence domains (LCDs) which mediate the assembly of membrane-less organelles such as RNA granules, nucleoli, spliceosomes and the nuclear pore complex (NPC). In order for membrane-less organelles to form they must separate from the liquid cytoplasm, which is achieved by concentrating the organelle components and the formation of a network of weak multi-valent interactions, a process known as liquid-liquid phase separation (LLPS) (Hyman et al., 2014). hnRNPs including FUS, hnRNP A1 and TIA1 undergo LLPS and aid the formation of stress granules (SGs) (Murakami et al., 2015; Patel et al., 2015; Lin et al., 2016). Arginine-rich DPRs are able to interact with the LCD of these proteins, disrupting LLPS and altering their biophysical properties (Lee et al., 2016). Therefore, DPRs may contribute to toxicity by upsetting the composition of membrane-less organelles.

Although the precise role of DPR toxicity in *C9orf72*-associated pathogenesis remains controversial, it is evident that their accumulation has detrimental effects on nucleolar function and RNA metabolism. Specific effects of DPR aggregation on RNA processing are discussed in further detail in the following sections.

ASSESSING THE RELATIVE CONTRIBUTION OF RNA FOCI AND DPR ACCUMULATION TO *C9orf72* TOXICITY

Since the discovery of RNA foci and RAN translation products in cells derived from expansion carriers, the relative contribution of these disease parameters to the overall pathomechanism remains a controversial issue. The main argument against the role of DPRs as a main driver of toxicity comes from neuropathological analysis using human brain tissue, which has revealed that DPR distribution is not spatially correlated to severity of degeneration in ALS or FTD (Mackenzie et al., 2013; Davidson et al., 2014, 2016; Gomez-Deza et al., 2015; Schuldi et al., 2015). In fact, DPR load is lower in vulnerable regions (e.g., motor cortex) and higher in unaffected areas (e.g., cerebellum) (Mackenzie et al., 2013), suggestive of a possible neuroprotective role for insoluble DPR aggregates. On the other hand, extensive quantification of RNA foci in the frontal cortex and cerebellum of *C9orf72* expansion carriers revealed limited association between clinic-pathological phenotypes (DeJesus-Hernandez et al., 2017), suggesting that neither RAN translation nor RNA foci can in isolation account for the phenotypic heterogeneity among expansion carriers. Even together, their presence is insufficient to cause disease, as evidenced by a patient with 30× repeats exhibiting both RNA foci and DPRs in the absence of TDP-43 pathology or clinical symptoms (Gami et al., 2015). This suggests that longer repeat lengths overcome a pathological threshold, enabling these initial molecular abnormalities to trigger full-blown disease.

In order to elucidate the relative contributions of RNA toxicity and RAN translation, several experimental strategies employing the use of randomized codon constructs and G_4C_2 repeats containing stop codon interruptions have been developed to determine the effects of RNA or DPRs in isolation. The results of these studies are summarized in **Table 1**. Taken together, these studies suggest that G_4C_2 toxicity is mediated through the expression of arginine-rich DPRs, particularly poly-PR, which induce nucleolar stress. However, the fact that we cannot conclusively prove that RNA or DPRs alone act as a main driver of *C9orf72* toxicity indicates that it is a combination of the two, perhaps in parallel with other disease factors, which is ultimately responsible for disease.

PROCESSING OF THE EXPANDED *C9orf72* ALLELE TRANSCRIPT

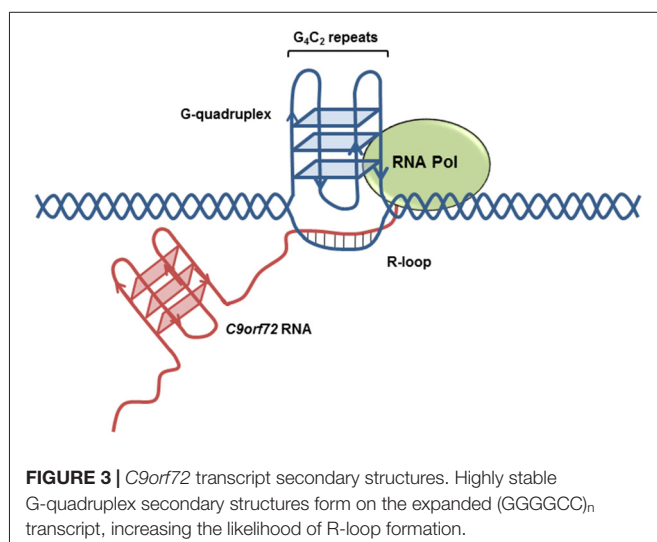
Owing to the high guanine content of the sense transcript, the G_4C_2 repeat expansion forms highly stable G-quadruplex structures (**Figure 3**). G-quadruplexes consist of stacks of planar tetramers in which four guanines interact through Hoogsteen hydrogen bonds. Furthermore, the antisense transcript adopts i-motif secondary structure, consisting of two parallel duplexes held by hemi-protonated C^+-C pairs, under near physiological conditions (Kovanda et al., 2015). Highly sensitive gene expression profiling identified increased levels of potentially truncated transcripts compared to controls in the frontal cortex, an effect most prominent among patients diagnosed with FTD (van Blitterswijk et al., 2015). Furthermore, a substantial increase in the number of sense and antisense transcripts containing intron 1 (the location of expansion) has been reported among *C9orf72* expansion carriers relative to controls in both cerebellum (Mori et al., 2013b; Niblock et al., 2016) and patient-derived iPSCs (Sareen et al., 2013). This suggests that the presence of the repeat expansion results in preferential use of exon 1a, resulting in a transcriptional bias towards the expanded transcript (Sareen et al., 2013). It is of interest to note that levels of intron 1 containing transcripts are significantly higher in the frontal cortex of FTD patients compared to individuals with ALS (Mori et al., 2013b; van Blitterswijk et al., 2015).

Analysis of *C9orf72* intron 1 splicing demonstrated that a proportion of polyadenylated *C9orf72* transcripts retained intron 1. Intron retention was significantly increased in the frontal cortex from *C9orf72* expansion carriers compared to controls. This effect was especially pronounced in a case homozygous for the *C9orf72* repeat expansion (Niblock et al., 2016). These findings could reflect the aforementioned transcriptional bias towards exon 1a, thereby increasing the total number of transcripts containing the expanded sequence. Of note, a relationship has been established between high intronic GC content and splicing abnormalities (Wong et al., 2013). Furthermore, reduced rates of transcription across the intron owing to complex secondary structure may contribute to perturbed splicing of intron 1.

TABLE 1 | Assessing the relative contributions of RNA and dipeptide repeat protein (DPR) toxicity in *C9orf72* model organisms.

Experimental strategy	Effect in model system	References
Compared “pure repeats” (produced RNA foci and DPRs) with “RNA-only repeats” containing stop-codon interruptions (produced only RNA foci) in <i>Drosophila</i>	Eye degeneration occurs only when pure repeats are expressed. Pure repeats have no effect.	Mizielinska et al. (2014)
“Protein-only” constructs employing alternative codons to generate DPRs in <i>Drosophila</i>	Arginine-rich DPRs cause eye degeneration and lethality.	Mizielinska et al. (2014)
Alternative codons to generate DPRs expressed in cortical and motor neurons	PR ₅₀ expression decreased cortical neuron survival. Both PR ₅₀ and GR ₅₀ toxic in motor neurons.	Wen et al. (2014)
Constructs containing intronic GGGGCC repeats (which do not initiate RAN translation) expressed in cortical, motor and primary cortical neurons	Increased death of cortical and motor neurons expressing 42× repeats compared to 21× repeats and controls.	Wen et al. (2014); Zhang et al. (2014)
Synthetic GR ₂₀ and PR ₂₀ peptides produced using a peptide synthesizer applied to U2OS cells and human astrocytes	PR ₂₀ and GR ₂₀ bind nucleoli and kill U2OS cells. Disrupt splicing and ribosomal RNA biogenesis in astrocytes.	Kwon et al. (2014)
GFP-GA ₅₀ construct consisting of G4C2 repeats with stop codon interruptions expressed in primary neurons	GA ₅₀ expression stimulates caspase-3 activation, ER stress and decreased proteasome activity	Zhang et al. (2014)
Alternative codons to express poly-GA, poly-GR and poly-GP in <i>Drosophila</i>	GR ₅₀ causes eye degeneration and death. GA and GP expression has no effect.	Freibaum et al. (2015)
Alternative codons to express poly-GR, poly-PR, poly-GA, poly-GP and poly-PA in <i>Drosophila</i>	GR ₅₀ and PR ₅₀ causes severe eye degeneration. GA ₅₀ , GP ₅₀ and PA ₅₀ has no effect.	Lee et al. (2016)
G4C2 repeats flanked by <i>C9orf72</i> intronic and exonic sequences which form RNA foci but low levels of DPRs compared to G4C2 repeats with a poly(A) tail producing high levels of RAN translation expressed in <i>Drosophila</i>	Intronic construct causes modest toxicity whereas poly(A) construct produces severe eye degeneration.	Tran et al. (2015)
Alternative codons to express poly-GR, poly-PR, poly-GA, poly-GP and poly-PA in <i>S. cerevisiae</i>	PR ₅₀ expression is sufficient to disrupt nuclear pore trafficking	Jovčić et al. (2015)

Taken together, these results suggest that the presence of the expanded G₄C₂ sequence and the formation of complex secondary structure results in numerous transcriptional defects, including abortive transcription, increased use of exon 1a and ineffective splicing of the repeat-containing intron.



GLOBAL SPLICING ALTERATION

C9orf72 RNA foci sequester several members of the hnRNP family of splicing factors (Lee et al., 2013; Mori et al., 2013a; Sareen et al., 2013; Cooper-Knock et al., 2014), resulting in altered splicing patterns of their RNA targets. The production of certain alternatively spliced mRNAs depends on the relative abundance of antagonistically acting splicing factors, highlighting the potential implications of hnRNP sequestration within G₄C₂ foci. Multiple lines of evidence point towards perturbations in constitutive and alternative splicing events in *C9orf72* expansion carriers. Gene expression profiling has revealed an increased occurrence of splicing errors in cell lines derived from c9ALS patients, an effect which was most evident among patients with faster disease progression (Cooper-Knock et al., 2015a). Transcriptome analysis of differentially expressed genes between c9ALS and control groups revealed an enrichment of upregulated transcripts involved in RNA splicing (Cooper-Knock et al., 2015a), consistent with a compensatory mechanism for RBP sequestration by RNA foci. Many of the differentially expressed genes have also been independently identified as candidate G₄C₂ binding proteins, including hnRNP A3 and hnRNP H (Lee et al., 2013; Mori et al., 2013a). These results suggest that the increased splicing error rate in expansion carriers is a consequence of RBP sequestration

into foci, which in turn contributes to disease progression and severity. Consistent with this hypothesis, analysis of abundant RBP motifs among transcripts exhibiting splicing abnormalities revealed an enrichment for binding motifs for hnRNP H (Prudencio et al., 2015). Furthermore, an additional transcriptome analysis reported a three-fold increase in the number of splicing events in c9ALS cases compared to sporadic ALS in both cerebellum and frontal cortex (Prudencio et al., 2015). Intron retention events embodied a significant proportion of such splicing events, revealing a staggering 40-fold increase among c9ALS compared to sporadic ALS within the frontal cortex. Such differences in the extent of intron retention in c9ALS may contribute to the reported increase in nuclear accumulation of mRNA transcripts in *C9orf72*-ALS models (Freibaum et al., 2015; Jovičić et al., 2015; Rossi et al., 2015).

Splicing abnormalities also arise from pathogenic events other than RBP sequestration into foci, such as DPR toxicity. Exposure of cultured cells to arginine-rich DPRs results in their localization to the nucleus where they bind nucleoli, resulting in disruptions in ribosomal RNA biogenesis (Kwon et al., 2014). Furthermore, PR₂₀ exposure produced altered splicing patterns of specific mRNA transcripts, including exon skipping and intron retention events. PR₂₀ expression caused exon 2 skipping of the mRNA encoding Ran GTPase, producing a protein of reduced length. Ran GTPase functions as a regulator of protein nucleocytoplasmic transport, which has been implicated in *C9orf72*-mediated toxicity and is discussed in further detail below. Exposure to PR₂₀ also promotes intron retention within the *GADD45* mRNA, altering the open reading frame and thereby encoding an inactive version of the *GADD45* protein (Kwon et al., 2014). Such findings suggest a possible contribution of DPR proteins to the alternative splicing abnormalities documented in c9ALS. Indeed, impaired ribosomal RNA biogenesis and RNA granule formation due to DPR expression suggest that DPRs may perturb RNA processing on a global level (Kwon et al., 2014; Tao et al., 2015).

Thus, TDP-43 loss-of-function, RBP sequestration and the production of DPR species converge to disrupt splicing patterns of mRNAs in various cellular pathways. Perturbations in numerous signaling pathways may explain why so many processes, from autophagy to the stress response, are impaired in *C9orf72*-disease.

GLOBAL CHANGES IN POLYADENYLATION SITE SELECTION

In addition to alterations in alternative splicing, variations in polyadenylation site (PAS) selection of a large number of transcripts have been identified in the cerebellum of individuals with *C9orf72*-ALS. Analogous to alternative splicing, alternative PAS selection contributes to the complexity of the human transcriptome and has been detected in over 50% of human genes (Tian et al., 2005). Alternative polyadenylation generates mRNA isoforms which vary in their 3'-untranslated regions (3'-UTRs), thus determining the stability, localization and

translational efficiency. c9ALS patients show increased use of upstream PASs compared to downstream PASs in the cerebellum compared to sporadic ALS patients (Prudencio et al., 2015). Transcripts exhibiting shifts in PAS selection are enriched for functions in RNA processing, including transcripts encoding exosomal proteins, components of the RISC complex and splicing factors, such as *TARBP2*, *ATXN2* and *EXOSC7* (Prudencio et al., 2015). 3'-UTRs frequently harbor miRNA and RBP binding sites, therefore the selection of upstream PASs, which reduce the length of the 3'-UTR, are likely to result in the loss of such binding sites. mRNA transcripts with shorter 3'-UTRs have an increased stability, resulting in higher protein expression and thereby generating distortions in downstream pathways (Matoulkova et al., 2012). Altered 3'-UTR length may also affect the stability, localization and transport of the mRNA. In addition to its role in alternative splicing, hnRNP H has also been implicated in alternative PAS selection (Chou et al., 1999), suggesting a role for RBP sequestration in disrupted alternative polyadenylation events in *C9orf72*-mediated toxicity.

NUCLEOCYTOPLASMIC TRANSPORT

Mature mRNA transcripts are transported to the cytoplasm from the nucleus through large proteinaceous assemblies embedded in the nuclear envelope, termed NPCs. The mRNA transcript, together with proteins involved in pre-mRNA processing, exist as a large mRNP which interacts with the NPC through nuclear transport receptors. RNA processing reactions generate signals in the form of proteins docked onto the mature mRNA which interact with the nuclear export machinery, ensuring that only fully matured mRNAs are transported to the cytoplasm for translation (Brodsky and Silver, 2000). For example, the UAP56 helicase is retained onto the mRNA after splicing and recruits the nuclear export protein, ALYREF, thereby coupling splicing to nuclear export (Luo et al., 2001). Conversely, following translation, nuclear proteins are transported from the cytoplasm to the nucleus. The vast majority of nuclear proteins are transported from the cytoplasm via the Ran-mediated pathway, which is an energy-dependent directional process reliant on the asymmetric distribution of Ran-GTP in the nucleus and Ran-GDP in the cytoplasm. RanGAP functions as a molecular switch, converting Ran-GTP to Ran-GDP, enabling Ran to shuttle proteins into the cytoplasm and maintaining the gradient of cytoplasmic Ran-GDP (Figure 4). Ran-mediated nucleocytoplasmic trafficking is facilitated by the karyopherin protein family (consisting of importins and exportins which mediate nuclear import and export, respectively) which form a complex with Ran, cargo proteins and components of the NPC.

A genetic screen in *Drosophila* expressing (G₄C₂)₃₀ repeats identified RanGAP as a potent suppressor of the rough eye phenotype (Zhang et al., 2015), suggesting that upregulation of protein nuclear import can alleviate *C9orf72*-associated toxicity. Furthermore, RanGAP was shown to physically interact with G₄C₂ RNA repeats and exhibited mislocalization from the nuclear periphery into large perinuclear aggregates in *Drosophila*

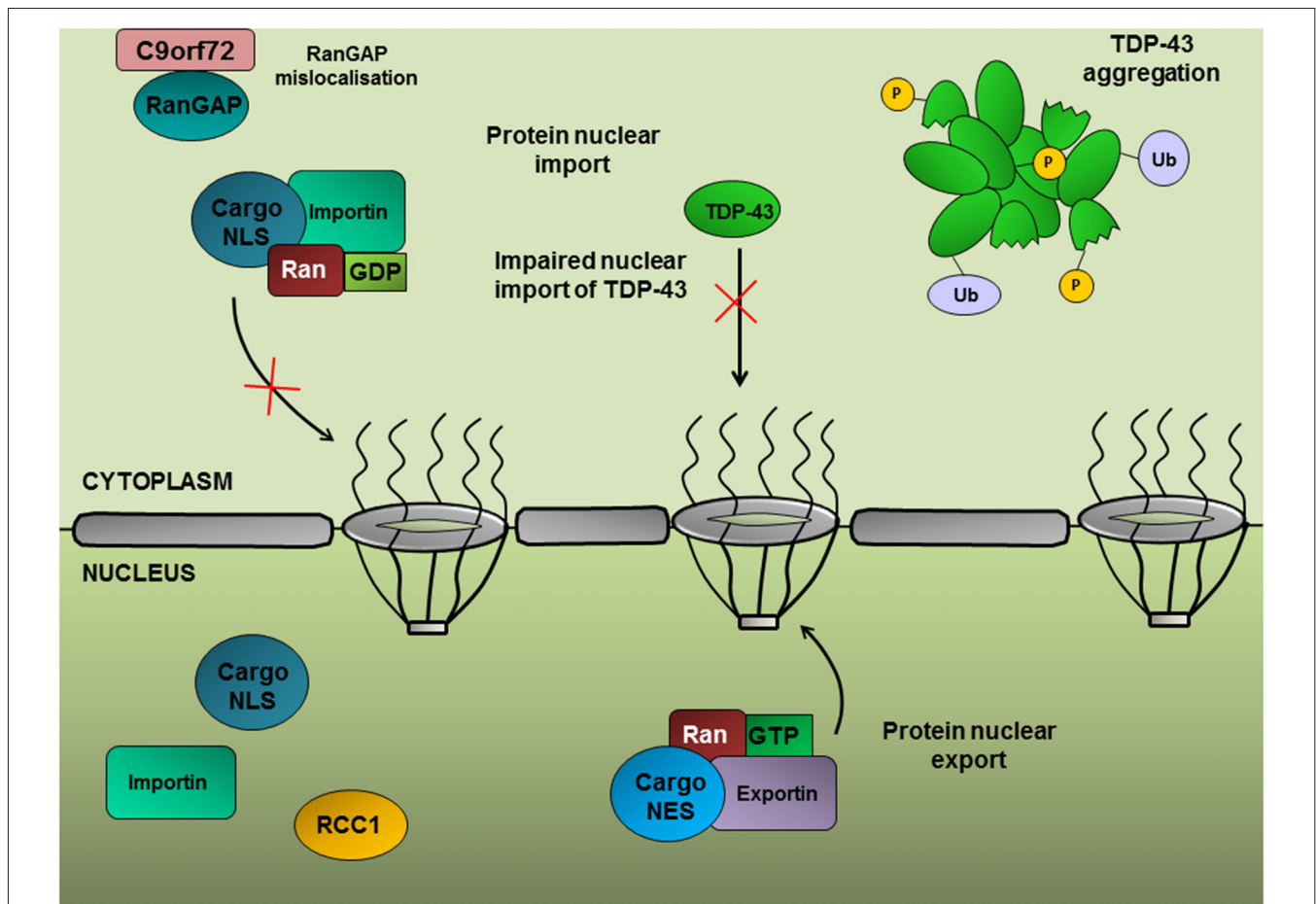


FIGURE 4 | Ran-mediated protein nuclear import. The import of proteins into the nucleus is dependent on a concentration gradient of Ran-GDP:Ran-GTP in the cytoplasmic and nuclear compartments, respectively. RanGAP converts Ran to its GDP-bound form, enabling Ran to shuttle proteins possessing a nuclear localization signal (NLS) into the nucleus. Nuclear import of TDP-43 depends on the Ran-mediated pathway; therefore RanGAP mislocalization could contribute to TDP-43 nuclear depletion.

expressing $(G_4C_2)_{30}$, in patient-derived iPSCs and in the motor cortex from c9ALS patients. A significant correlation between the nuclear to cytoplasmic ratios of Ran and TDP-43 has been established (Ward et al., 2014; Zhang et al., 2015), implying a role for Ran in the nuclear import of TDP-43. Indeed, siRNA knockdown of Ran accessory proteins, such as importin- β , which function in the nuclear import of proteins with a nuclear localization signal (NLS), results in the cytoplasmic accumulation of TDP-43 (Nishimura et al., 2010). Furthermore, TDP-43 regulates Ran expression through interaction with the 3'-UTR of *Ran* mRNA and TDP-43 inhibition leads to a decrease in *Ran* mRNA and protein expression (Sephton et al., 2011; Ward et al., 2014).

Xiao et al. (2015) reported a direct interaction between the C9orf72 protein with Ran and importin- β , implicating a potential role for the C9orf72 protein in nucleocytoplasmic trafficking. This group demonstrated redistribution of the C9orf72 short isoform (isoform B) from the nuclear membrane in healthy neurons to the plasma membrane in neurons from c9ALS patients. Loss of C9orf72 isoform B, Ran and

importin- β 1 from the nuclear membrane correlated with TDP-43 mislocalization and aggregation (Xiao et al., 2015).

Another genetic screen in *Drosophila* expressing $(G_4C_2)_{58}$ specifically identified disruptions in the protein and mRNA nuclear export pathways (Freibaum et al., 2015). $(G_4C_2)_{58}$ toxicity is exacerbated by disruptions to the CRM1-mediated nuclear export pathway, the pathway responsible for the nuclear export for the majority of proteins and RNA transcripts. Depletion of the nuclear pore components Nup107 or Nup160 which function in RNA nuclear export exacerbates degeneration caused by $(G_4C_2)_{58}$. Furthermore, downregulation of *Ref1*, the *Drosophila* ortholog of *ALYREF*, was identified as the strongest suppressor of toxicity (Freibaum et al., 2015). Immunohistochemistry and UV crosslinking have previously demonstrated colocalization to foci and direct binding of *ALYREF* with both sense and antisense RNA in c9ALS motor neurons (Cooper-Knock et al., 2014, 2015b). Through its interaction with the TREX complex, *ALYREF* protects the mRNP from exosomal degradation (Chang et al., 2013), therefore the suppressive effect of *Ref1* loss of function results

from increased exosomal degradation of nuclear transcripts. A higher proportion of nuclear to cytoplasmic mRNA transcripts were identified in G_4C_2 -expressing *Drosophila* and c9ALS iPSCs, an effect rescued by *Ref1* knockdown. Moreover, loss of exosomal function was demonstrated to enhance toxicity (Freibaum et al., 2015) and mutations in *EXOSC3* which encodes an exosomal component are responsible for a congenital form of motor neuron disease, pontocerebellar hypoplasia (Wan et al., 2012).

To elucidate the contribution of RNA and DPR toxicity to nucleocytoplasmic defects, a yeast genetic screen was used to generate poly-PR employing codon-optimized constructs without using the G_4C_2 repetitive sequence. The strongest suppressors of toxicity were identified as members of the karyopherin family of nuclear import proteins and MTR10, a nuclear import receptor which mediates the import of the serine-arginine (SR) family of splicing factors (Jovičić et al., 2015). It has been previously suggested that DPRs may compete with SR proteins for binding to ribonucleoproteins, as both the SR-domains of splicing factors and synthetic constructs containing 20 GR or PR repeats both bind hnRNP A2 hydrogels (Kwon et al., 2014). Upregulation of MTR10 could alleviate this competitive effect, thereby increasing the influx of SR proteins into the nucleus. This would restore alterations in RNA splicing and processes, such as mRNA nuclear export and nonsense-mediated decay. Consistent with the *Drosophila* genetic screens, the same group also identified components of the NPC and Ran-associated proteins as genetic modifiers of (PR)₅₀ toxicity in yeast.

Taken together, the results of these studies suggest that abnormal nucleocytoplasmic transport is a pathological signature of c9ALS/FTD. Disruptions in nucleocytoplasmic transport have also been documented in a mutant SOD1 mouse model of ALS (Zhang et al., 2006), suggestive of an underlying defect common to ALS with different genetic etiologies. Nuclear import and export factors therefore represent a potential therapeutic target for G_4C_2 repeat toxicity. ASOs targeting the *C9orf72* transcript have already been shown to normalize Ran and TDP-43 localization in neurons derived from c9ALS iPSCs, presumably through the reduction of G_4C_2 RNA foci (Zhang et al., 2015). In addition, small molecules which disrupt the G-quadruplex conformation of G_4C_2 RNA rescue nuclear import deficits in the (G_4C_2)₃₀ *Drosophila* model (Zhang et al., 2015), highlighting the therapeutic amenability of nucleocytoplasmic transport.

mRNA LOCALIZATION AND TRANSLATION

Expanded G_4C_2 RNA localizes to neurites where it is subsequently incorporated into active transport granules in *Drosophila* neurons and iPSCs from expansion carriers (Burguete et al., 2015). RNA transport granules are ribonucleoprotein complexes which serve to transport mRNA along microtubules for local protein synthesis. For highly polarized cells, such as neurons, efficient mRNA transport and local translation is crucial for the maintenance of synaptic plasticity. Localization of the G_4C_2 repeat RNA

to neuronal transport granules is associated with neuritic branching deficits which are modulated by the components of transport granules and translational regulators, FMRP and CPEB3 (Burguete et al., 2015), suggesting that loss of neuritic branches may result from transport of G_4C_2 repeat RNA in neurites. Secondary structural conformations, such as G-quadruplexes, appear to promote neuritic localization of G_4C_2 repeat RNA and incorporation into RNA transport particles (Subramanian et al., 2011; Burguete et al., 2015), therefore targeting specific structural conformations may promote neuritic branching of neurons in repeat expansion carriers. Of note, binding partners and modifiers of G_4C_2 repeat toxicity, including hnRNP A3, hnRNP A2/B1 and Pur- α , are responsible for maintaining transport granule function (Jin et al., 2007; Sofola et al., 2007; Xu et al., 2013).

Impairment in the mRNA nuclear export pathway is predicted to result in reduced translation of mRNA into proteins. Indeed, downregulation of protein translation has been documented in cells expressing (G_4C_2)₃₁. An RNA pulldown assay using (G_4C_2)₃₁ revealed binding proteins enriched for functions in translational control, such as EF1 α , eIF2 α , eIF2 β and Pur- α , indicating that protein sequestration by the repeat expansion may also contribute to impaired translation (Rossi et al., 2015). Expression of the G_4C_2 sequence in cultured mammalian cells initiates a stress response, characterized by the formation of SGs and a subsequent reduction in the rate of global translation (Rossi et al., 2015). SGs serve to protect the cell during stress, transiently storing mRNAs encoding housekeeping genes and prioritizing the translation of stress-response proteins, such as heat shock proteins and chaperones. SGs are composed of RBPs and poly(A)⁺ mRNAs, which are either transiently stored to resume translation upon alleviation of the stress response or are degraded in processing bodies (P-bodies; Bentmann et al., 2013). Persistent SG formation results in a chronic stress response and prolonged translational repression. Increasing lines of evidence suggest that SG-associated translational repression may play a role in c9ALS/FTD. For example, TDP-43, FUS and Ataxin-2, mutations in which are associated with neurodegenerative disease, all localize to SGs upon cellular stress and may function to regulate SG assembly (Wolozin, 2012; Monahan et al., 2016). Furthermore, genes involved in SG formation are potent modifiers of TDP-43 toxicity in yeast and *Drosophila* (Kim et al., 2014). TDP-43 expression promotes eIF2 α phosphorylation, indicative of SG formation and translational repression. Inhibiting SG formation using a small molecule inhibitor of eIF2 α alleviates TDP-43 toxicity (Kim et al., 2014). eIF2 α is crucial for translation initiation, facilitating the interaction between the ribosome and Met-tRNA. Phosphorylation of eIF2 α results in its inactivation, inhibiting global protein synthesis. SG formation can be stimulated by limited availability of eIF2 α (Mokas et al., 2009), therefore eIF2 α sequestration into RNA foci (Rossi et al., 2015) has the potential to initiate SG formation and translational inhibition. Arginine-rich DPR species impede protein translation in a dose-dependent manner through a mechanism independent of eIF2 α phosphorylation (Kanekura

et al., 2016). Immunoprecipitation analysis revealed that poly-(PR)₂₀ interacts with ribosomal proteins and translation initiation and elongation factors. Furthermore, both poly-(PR)₂₀ and poly-(GR)₂₀ were shown to aggregate in the presence of mRNA transcripts, obstructing access of the eIF4E and eIF4G translation initiation factors and disrupting proteins translation (Kanekura et al., 2016).

Consistent with reduced nuclear export, nuclear accumulation of poly(A)⁺ mRNA transcripts was detected in (G₄C₂)₃₁-expressing cells and was strongly associated with relocalization of cytoplasmic poly(A)-binding protein (PABPc) to the nucleus. PABPc directly binds G₄C₂ RNA and accumulates into RNA foci (Rossi et al., 2015). Of note, targeting PABPc to the nucleus is sufficient to cause nuclear accumulation of mRNAs (Kumar and Glaunsinger, 2010), therefore nuclear sequestration of PABPc by G₄C₂ foci may represent a potential link between RNA toxicity and nuclear retention of mRNAs. Furthermore, PABPc plays a key role in regulating mRNA translation by facilitating the interaction between translation initiation factors and the mRNA 5' cap, enhancing ribosomal recruitment. PABPc nuclear sequestration may reduce global translational efficiency of the depleted pool of mRNAs present in the cytoplasm. Thus, sequestration of translational regulators by RNA foci may result in the nuclear retention of mRNAs and a concomitant reduction of translation in the cytoplasm.

IMPAIRED RNA PROCESSING IS ASSOCIATED WITH R-LOOP FORMATION

The occurrence of the G-quadruplex or i-motif structures from the sense or antisense transcript, respectively, displaces the complementary DNA strand, favoring the formation of R-loop secondary structures. R-loops form during transcription when the nascent RNA molecule interacts with displaced DNA, resulting in the formation of a three stranded nucleic acid structure consisting of an RNA:DNA hybrid and single stranded DNA (ssDNA) (Figure 3), taking advantage of the increased thermodynamic stability of RNA:DNA hybrids compared to the DNA duplex. Conversely, perturbations in RNA processing can result in increased formation of R-loop structure (Huertas and Aguilera, 2003; Wahba et al., 2011). R-loops may reduce transcriptional efficiency resulting in the formation of aborted transcripts. An *in vitro* transcription assay using G₄C₂ repeats that form G-quadruplex structures demonstrated a repeat-length dependent accumulation of transcripts truncated within the repeat expansion region. RNase H treatment, which specifically digests the RNA transcript within DNA-RNA duplexes and resolves R-loop structures, reduces abortive transcripts (Haeusler et al., 2014). Of note, mutations in senataxin, an RNA helicase which resolves R-loop structures, are causative of ALS4, a juvenile form of ALS (Chen et al., 2004). Depletion of the splicing factor ASF/SF2, a protein shown to bind G₄C₂ repeat sequences (Reddy et al., 2013; Rossi et al., 2015), results in a concomitant increase in R-loop formation (Li and Manley, 2005). Furthermore, R-loops are a known source of DNA damage (Huertas and Aguilera,

2003) and are likely to contribute to the upregulation of DNA damage markers identified in ALS motor neurons (Farg et al., 2017).

In addition to *C9orf72*-linked disease R-loop formation has been characterized in several microsatellite disorders, such as Fragile X Syndrome (FXS) and Friedrich's Ataxia (FRDA) (Groh et al., 2014). R-loops were found to be enriched over GAA repeats associated with FRDA and the abundance of R-loops correlated with repeat expansion length. Furthermore, increased formation of R-loops resulted in upregulation of repressive chromatin, promoting transcriptional silencing of the *FXN* gene (Groh et al., 2014). Therefore, R-loop formation may contribute to some of the processes characteristic of repeat expansion disorders, including repeat instability, antisense transcription and impaired transcription. Incubation of transcriptionally induced R-loops with cell lysates stimulated variations in repeat length within the DNA template (Reddy et al., 2014), suggesting that R-loop formation can have direct consequences on repeat instability. It remains unclear how R-loops promote repeat instability, however it has been proposed that R-loop structures favor a non-B-DNA structure of the ssDNA, resulting in the recruitment of various DNA repair factors which mediate error-prone DNA repair synthesis (Lin et al., 2010). R-loops also have the ability to induce transcriptional pausing or stalling, perhaps by forming a structural block or by triggering mechanisms of transcriptional repression, such as histone methylation (Groh et al., 2014; Skourti-Stathaki et al., 2014). Additionally, R-loops may induce expression of the antisense transcript (Skourti-Stathaki et al., 2014). The fact that R-loop formation is associated with so many features of *C9orf72* toxicity makes it an imperative area of future research and a potential attractive therapeutic target applicable to a range of microsatellite repeat disorders. Small molecules able to intercalate within G-quadruplex or hairpin structures have demonstrated promising suppression of R-loops arising from trinucleotide expansions (Colak et al., 2014).

CONCLUDING REMARKS

It is becoming clear that impaired RNA processing is a central feature of c9ALS/FTD pathogenesis which is mediated via several interconnected mechanisms, including sequestration of RBPs, RAN translation and TDP-43 mislocalization. It is likely that these mechanisms all contribute to RNA processing abnormalities to different extents, generating numerous small defects at the level of RNA metabolism. Initially the cell may be able to compensate for these defects, however over time these abnormalities may accumulate beyond a certain threshold at which point the aging cell loses viability. The idea of c9ALS/FTD as a disorder mediated by an accumulation of several minor disruptions, rather than a few isolated RNA misprocessing events as has been described for myotonic dystrophy type 1 (Timchenko et al., 2001; Jiang et al., 2004), may in part account for the extreme variability in clinical phenotypes exhibited by *C9orf72* expansion carriers. Given the importance of RNA processing in cellular function, RNA metabolism can be considered an essential target for therapeutic intervention to improve the lives of those living

not only with c9ALS/FTD but for a range of neurodegenerative disorders.

AUTHOR CONTRIBUTIONS

HVB wrote the manuscript. MN, Y-BL and CES reviewed the manuscript. J-MG reviewed and edited the manuscript.

REFERENCES

- Alami, N. H., Smith, R. B., Carrasco, M. A., Williams, L. A., Winborn, C. S., Han, S. S., et al. (2014). Axonal transport of TDP-43 mRNA granules is impaired by ALS-causing mutations. *Neuron* 81, 536–543. doi: 10.1016/j.neuron.2013.12.018
- Arnold, E. S., Ling, S. C., Huelga, S. C., Lagier-Tourenne, C., Polymenidou, M., Ditsworth, D., et al. (2013). ALS-linked TDP-43 mutations produce aberrant RNA splicing and adult-onset motor neuron disease without aggregation or loss of nuclear TDP-43. *Proc. Natl. Acad. Sci. U S A* 110, E736–E745. doi: 10.1073/pnas.1222809110
- Ash, P. E., Bieniek, K. F., Gendron, T. F., Caulfield, T., Lin, W. L., DeJesus-Hernandez, M., et al. (2013). Unconventional translation of *C9ORF72* GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron* 77, 639–646. doi: 10.1016/j.neuron.2013.02.004
- Atanasio, A., Decman, V., White, D., Ramos, M., Ikiz, B., Lee, H. C., et al. (2016). *C9orf72* ablation causes immune dysregulation characterized by leukocyte expansion, autoantibody production, and glomerulonephropathy in mice. *Sci. Rep.* 6:23204. doi: 10.1038/srep23204
- Bentmann, E., Haass, C., and Dormann, D. (2013). Stress granules in neurodegeneration—lessons learnt from TAR DNA binding protein of 43 kDa and fused in sarcoma. *FEBS J.* 280, 4348–4370. doi: 10.1111/febs.12287
- Blokhuis, A. M., Koppers, M., Groen, E. J., van den Heuvel, D. M., Dini Modigliani, S., Anink, J. J., et al. (2016). Comparative interactomics analysis of different ALS-associated proteins identifies converging molecular pathways. *Acta Neuropathol.* 132, 175–196. doi: 10.1007/s00401-016-1575-8
- Boeynaems, S., Bogaert, E., Kovacs, D., Konijnenberg, A., Timmerman, E., Volkov, A., et al. (2017). Phase separation of *C9orf72* dipeptide repeats perturbs stress granule dynamics. *Mol. Cell* 65, 1044.e5–1055.e5. doi: 10.1016/j.molcel.2017.02.013
- Brodsky, A. S., and Silver, P. A. (2000). Pre-mRNA processing factors are required for nuclear export. *RNA* 6, 1737–1749. doi: 10.1017/s135583820001059
- Buratti, E., and Baralle, F. E. (2010). The multiple roles of TDP-43 in pre-mRNA processing and gene expression regulation. *RNA Biol.* 7, 420–429. doi: 10.4161/rna.7.4.12205
- Buratti, E., De Conti, L., Stuardi, C., Romano, M., Baralle, M., and Baralle, F. (2010). Nuclear factor TDP-43 can affect selected microRNA levels. *FEBS J.* 277, 2268–2281. doi: 10.1111/j.1742-4658.2010.07643.x
- Burguete, A. S., Almeida, S., Gao, F. B., Kalb, R., Akins, M. R., and Bonini, N. M. (2015). GGGGCC microsatellite RNA is neuriticly localized, induces branching defects, and perturbs transport granule function. *Elife* 4:e08881. doi: 10.7554/eLife.08881
- Byrne, S., Heverin, M., Elamin, M., Walsh, C., and Hardiman, O. (2014). Intermediate repeat expansion length in *C9orf72* may be pathological in amyotrophic lateral sclerosis. *Amyotroph. Lateral Scler. Frontotemporal Degener.* 15, 148–150. doi: 10.3109/21678421.2013.838586
- Chen, Y. Z., Bennett, C. L., Huynh, H. M., Blair, I. P., Puls, I., Irobi, J., et al. (2004). DNA/RNA helicase gene mutations in a form of juvenile amyotrophic lateral sclerosis (ALS4). *Am. J. Hum. Genet.* 74, 1128–1135. doi: 10.1086/421054
- Chen, Y., Lin, Z., Chen, X., Cao, B., Wei, Q., Ou, R., et al. (2016). Large *C9orf72* repeat expansions are seen in Chinese patients with sporadic amyotrophic lateral sclerosis. *Neurobiol. Aging* 38, 217.e215–217.e22. doi: 10.1016/j.neurobiolaging.2015.11.016
- Chang, C. T., Hautbergue, G. M., Walsh, M. J., Viphakone, N., van Dijk, T. B., Philipsen, S., et al. (2013). Chtop is a component of the dynamic TREX mRNA export complex. *EMBO J.* 32, 473–486. doi: 10.1038/emboj.2012.342
- Chiang, P. W., Carpenter, L. E., and Hagerman, P. J. (2001). The 5'-untranslated region of the FMR1 message facilitates translation by internal ribosome entry. *J. Biol. Chem.* 276, 37916–37921. doi: 10.1074/jbc.M105584200
- Chou, M. Y., Rooke, N., Turck, C. W., and Black, D. L. (1999). hnRNP H is a component of a splicing enhancer complex that activates a c-src alternative exon in neuronal cells. *Mol. Cell. Biol.* 19, 69–77. doi: 10.1128/mcb.19.1.69
- Clemson, C. M., Hutchinson, J. N., Sara, S. A., Ensminger, A. W., Fox, A. H., Chess, A., et al. (2009). An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol. Cell* 33, 717–726. doi: 10.1016/j.molcel.2009.01.026
- Colak, D., Zaninovic, N., Cohen, M. S., Rosenwaks, Z., Yang, W. Y., Gerhardt, J., et al. (2014). Promoter-bound trinucleotide repeat mRNA drives epigenetic silencing in fragile X syndrome. *Science* 343, 1002–1005. doi: 10.1126/science.1245831
- Cooper-Knock, J., Bury, J. J., Heath, P. R., Wyles, M., Higginbottom, A., Gelthorpe, C., et al. (2015a). *C9ORF72* GGGGCC expanded repeats produce splicing dysregulation which correlates with disease severity in amyotrophic lateral sclerosis. *PLoS One* 10:e0127376. doi: 10.1371/journal.pone.0127376
- Cooper-Knock, J., Higginbottom, A., Stopford, M. J., Highley, J. R., Ince, P. G., Wharton, S. B., et al. (2015b). Antisense RNA foci in the motor neurons of *C9ORF72*-ALS patients are associated with TDP-43 proteinopathy. *Acta Neuropathol.* 130, 63–75. doi: 10.1007/s00401-015-1429-9
- Cooper-Knock, J., Walsh, M. J., Higginbottom, A., Robin Highley, J., Dickman, M. J., Edbauer, D., et al. (2014). Sequestration of multiple RNA recognition motif-containing proteins by *C9orf72* repeat expansions. *Brain* 137, 2040–2051. doi: 10.1093/brain/awu120
- Darios, F., and Davletov, B. (2006). Omega-3 and omega-6 fatty acids stimulate cell membrane expansion by acting on syntaxin 3. *Nature* 440, 813–817. doi: 10.1038/nature04598
- Davidson, Y. S., Barker, H., Robinson, A. C., Thompson, J. C., Harris, J., Troakes, C., et al. (2014). Brain distribution of dipeptide repeat proteins in frontotemporal lobar degeneration and motor neurone disease associated with expansions in *C9ORF72*. *Acta Neuropathol. Commun.* 2:70. doi: 10.1186/2051-5960-2-70
- Davidson, Y., Robinson, A. C., Liu, X., Wu, D., Troakes, C., Rollinson, S., et al. (2016). 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. *Neuropath. Appl. Neurobiol.* 42, 242–254. doi: 10.1111/nan.12292
- Deák, F., Shin, O.-H., Kavalali, E. T., and Südhof, T. C. (2006). Structural determinants of synaptobrevin 2 function in synaptic vesicle fusion. *J. Neurosci.* 26, 6668–6676. doi: 10.1523/jneurosci.5272-05.2006
- DeJesus-Hernandez, M., Finch, N. A., Wang, X., Gendron, T. F., Bieniek, K. F., Heckman, M. G., et al. (2017). In-depth clinico-pathological examination of RNA foci in a large cohort of *C9ORF72* expansion carriers. *Acta Neuropathol.* doi: 10.1007/s00401-017-1725-7 [Epub ahead of print].
- DeJesus-Hernandez, M., Mackenzie, I. R., Boeve, B. F., Boxer, A. L., Baker, M., Rutherford, N. J., et al. (2011). Expanded GGGGCC hexanucleotide repeat in noncoding region of *C9ORF72* causes chromosome 9p-linked FTD and ALS. *Neuron* 72, 245–256. doi: 10.1016/j.neuron.2011.09.011 [Epub ahead of print].
- Donnelly, C. J., Zhang, P.-W., Pham, J. T., Heusler, A. R., Mistry, N. A., Vidensky, S., et al. (2013). RNA toxicity from the ALS/FTD *C9ORF72* expansion is mitigated by antisense intervention. *Neuron* 80, 415–428. doi: 10.1016/j.neuron.2013.10.015
- Farg, M. A., Konopka, A., Ying Soo, K., Ito, D., and Atkin, J. D. (2017). The DNA damage response (DDR) is induced by the *C9orf72* repeat expansion

FUNDING

Wellcome Trust/Medical Research Council (089701/Z/09/Z), the Motor Neurone Disease Association (Gallo/Apr14/827-791), Alzheimer's Research UK (ARUK-PhD2014-8). Open access for this article was funded by King's College London.

- in amyotrophic lateral sclerosis. *Hum. Mol. Genet.* doi: 10.1093/hmg/ddx170 [Epub ahead of print].
- Freibaum, B. D., Lu, Y., Lopez-Gonzalez, R., Kim, N. C., Almeida, S., Lee, K. H., et al. (2015). GGGGCC repeat expansion in *C9orf72* compromises nucleocytoplasmic transport. *Nature* 525, 129–133. doi: 10.1038/nature14974
- Gami, P., Murray, C., Schottlaender, L., Bettencourt, C., De Pablo Fernandez, E., Mudanohwo, E., et al. (2015). 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. doi: 10.1007/s00401-015-1473-5
- Gendron, T. F., Bieniek, K. F., Zhang, Y. J., Jansen-West, K., Ash, P. E., Caulfield, T., et al. (2013). 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. doi: 10.1007/s00401-013-1192-8
- Gijssels, I., Van Langenhove, T., van der Zee, J., Slegers, K., Philtjens, S., Kleinberger, G., et al. (2012). 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. doi: 10.1016/S1474-4422(11)70261-7
- Gijssels, I., Van Mossevelde, S., van der Zee, J., Sieben, A., Engelborghs, S., De Bleecker, J., et al. (2016). The *C9orf72* repeat size correlates with onset age of disease, DNA methylation and transcriptional downregulation of the promoter. *Mol. Psychiatry* 21, 1112–1124. doi: 10.1038/mp.2015.159
- Gomez-Deza, J., Lee, Y. B., Troakes, C., Nolan, M., Al-Sarraj, S., Gallo, J.-M., et al. (2015). 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. doi: 10.1186/s40478-015-0218-y
- Groh, M., Lufino, M. M., Wade-Martins, R., and Gromak, N. (2014). R-loops associated with triplet repeat expansions promote gene silencing in Friedreich ataxia and fragile X syndrome. *PLoS Genet.* 10:e1004318. doi: 10.1371/journal.pgen.1004318
- Guo, F., Jiao, F., Song, Z., Li, S., Liu, B., Yang, H., et al. (2015). Regulation of MALAT1 expression by TDP43 controls the migration and invasion of non-small cell lung cancer cells *in vitro*. *Biochem. Biophys. Res. Commun.* 465, 293–298. doi: 10.1016/j.bbrc.2015.08.027
- Haeusler, A. R., Donnelly, C. J., Periz, G., Simko, E. A. J., Shaw, P. G., Kim, M.-S., et al. (2014). *C9orf72* nucleotide repeat structures initiate molecular cascade of disease. *Nature* 507, 195–200. doi: 10.1038/nature13124
- Huertás, P., and Aguilera, A. (2003). Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. *Mol. Cell* 12, 711–721. doi: 10.1016/j.molcel.2003.08.010
- Hyman, A. A., Weber, C. A., and Jülicher, F. (2014). Liquid-liquid phase separation in biology. *Annu. Rev. Cell Dev. Biol.* 30, 39–58. doi: 10.1146/annurev-cellbio-100913-013325
- Jiang, H., Mankodi, A., Swanson, M. S., Moxley, R. T., and Thornton, C. A. (2004). Myotonic dystrophy type 1 is associated with nuclear foci of mutant RNA, sequestration of muscleblind proteins and deregulated alternative splicing in neurons. *Hum. Mol. Genet.* 13, 3079–3088. doi: 10.1093/hmg/ddh327
- Jiang, J., Zhu, Q., Gendron, T. F., Saberi, S., McAlonis-Downes, M., Seelman, A., et al. (2016). Gain of toxicity from ALS/FTD linked repeat expansions in *C9ORF72* is alleviated by antisense oligonucleotides targeting GGGGCC-containing RNAs. *Neuron* 90, 535–550. doi: 10.1016/j.neuron.2016.04.006
- Jin, P., Duan, R., Qurashi, A., Qin, Y., Tian, D., Rosser, T. C., et al. (2007). Pur α binds to rCGG repeats and modulates repeat-mediated neurodegeneration in a *Drosophila* model of fragile X tremor/ataxia syndrome. *Neuron* 55, 556–564. doi: 10.1016/j.neuron.2007.07.020
- Jovičić, A., Mertens, J., Boeynaems, S., Bogaert, E., Chai, N., Yamada, S. B., et al. (2015). Modifiers of *C9orf72* dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS. *Nat. Neurosci.* 18, 1226–1229. doi: 10.1038/nn.4085
- Kabashi, E., Valdmanis, P. N., Dion, P., Spiegelman, D., McConkey, B. J., Vande, V. C., et al. (2008). *TARDBP* mutations in individuals with sporadic and familial amyotrophic lateral sclerosis. *Nat. Genet.* 40, 572–574. doi: 10.1038/ng.132
- Kanekura, K., Yagi, T., Cammack, A. J., Mahadevan, J., Kuroda, M., Harms, M. B., et al. (2016). Poly-dipeptides encoded by the *C9ORF72* repeats block global protein translation. *Hum. Mol. Genet.* 25, 1803–1813. doi: 10.1093/hmg/ddw052
- Kawahara, Y., and Mieda-Sato, A. (2012). TDP-43 promotes microRNA biogenesis as a component of the Drosha and Dicer complexes. *Proc. Natl. Acad. Sci. U S A* 109, 3347–3352. doi: 10.1073/pnas.1112427109
- Kearse, M. G., Green, K. M., Krans, A., Rodriguez, C. M., Linsalata, A. E., Goldstrohm, A. C., et al. (2016). CGG repeat-associated non-AUG translation utilizes a cap-dependent scanning mechanism of initiation to produce toxic proteins. *Mol. Cell* 62, 314–322. doi: 10.1016/j.molcel.2016.02.034
- Kim, H. J., Raphael, A. R., LaDow, E. S., McGurk, L., Weber, R. A., Trojanowski, J. Q., et al. (2014). Therapeutic modulation of eIF2 α phosphorylation rescues TDP-43 toxicity in amyotrophic lateral sclerosis disease models. *Nat. Genet.* 46, 152–160. doi: 10.1038/ng.2853
- Koppers, M., Blokhuis, A. M., Westeneng, H. J., Terpstra, M. L., Zundel, C. A., Vieira de Sa, R., et al. (2015). *C9orf72* ablation in mice does not cause motor neuron degeneration or motor deficits. *Ann. Neurol.* 78, 426–438. doi: 10.1002/ana.24453
- Kovanda, A., Zalar, M., Sket, P., Plavec, J., and Rogelj, B. (2015). Anti-sense DNA d(GGCCCC)_n expansions in *C9ORF72* form i-motifs and protonated hairpins. *Sci. Rep.* 5:17944. doi: 10.1038/srep17944
- Kumar, G. R., and Glaunsinger, B. A. (2010). Nuclear import of cytoplasmic poly(A) binding protein restricts gene expression via hyperadenylation and nuclear retention of mRNA. *Mol. Cell. Biol.* 30, 4996–5008. doi: 10.1128/MCB.00600-10
- Kwon, I., Xiang, S., Kato, M., Wu, L., Theodoropoulos, P., Wang, T., et al. (2014). Poly-dipeptides encoded by the *C9orf72* repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science* 345, 1139–1145. doi: 10.1126/science.1254917
- Lagier-Tourenne, C., Baughn, M., Rigo, F., Sun, S., Liu, P., Li, H. R., et al. (2013). Targeted degradation of sense and antisense *C9orf72* RNA foci as therapy for ALS and frontotemporal degeneration. *Proc. Natl. Acad. Sci. U S A* 110, E4530–E4539. doi: 10.1073/pnas.1318835110
- Lagier-Tourenne, C., Polymenidou, M., Hutt, K. R., Vu, A. Q., Baughn, M., Huelga, S. C., et al. (2012). Divergent roles of ALS-linked proteins FUS/TLS and TDP-43 intersect in processing long pre-mRNAs. *Nat. Neurosci.* 15, 1488–1497. doi: 10.1038/nn.3230
- Lee, Y.-B., Chen, H.-J., Peres, J. N., Gomez-Deza, J., Attig, J., Stalekar, M., et al. (2013). Hexanucleotide repeats in ALS/FTD form length-dependent RNA foci, sequester RNA binding proteins, and are neurotoxic. *Cell Rep.* 5, 1178–1186. doi: 10.1016/j.celrep.2013.10.049
- Lee, K.-H., Zhang, P., Kim, H. J., Mitrea, D. M., Sarkar, M., Freibaum, B. D., et al. (2016). *C9orf72* dipeptide repeats impair the assembly, dynamics and function of membrane-less organelles. *Cell* 167, 774.e17–788.e17. doi: 10.1016/j.cell.2016.10.002
- Lepagnol-Bestel, A. M., Zvara, A., Maussion, G., Quignon, F., Ngimbous, B., Ramoz, N., et al. (2009). DYRK1A interacts with the REST/NRSF-SWI/SNF chromatin remodelling complex to deregulate gene clusters involved in the neuronal phenotypic traits of Down syndrome. *Hum. Mol. Genet.* 18, 1405–1414. doi: 10.1093/hmg/ddp047
- Li, X., and Manley, J. L. (2005). Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. *Cell* 122, 365–378. doi: 10.1016/j.cell.2005.06.008
- Lin, Y., Dent, S. Y., Wilson, J. H., Wells, R. D., and Napierala, M. (2010). R loops stimulate genetic instability of CTG/CAG repeats. *Proc. Natl. Acad. Sci. U S A* 107, 692–697. doi: 10.1073/pnas.0909740107
- Lin, Y., Mori, E., Kato, M., Xiang, S., Wu, L., Kwon, I., et al. (2016). Toxic PR poly-dipeptides encoded by the *C9orf72* repeat expansion target LC domain polymers. *Cell* 167, 789.e12–802.e12. doi: 10.1016/j.cell.2016.10.003
- Ling, J. P., Pletnikova, O., Troncoso, J. C., and Wong, P. C. (2015). TDP-43 repression of nonconserved cryptic exons is compromised in ALS-FTD. *Science* 349, 650–655. doi: 10.1126/science.aab0983
- Luo, M. L., Zhou, Z., Magni, K., Christoforides, C., Rappsilber, J., Mann, M., et al. (2001). Pre-mRNA splicing and mRNA export linked by direct interactions between UAP56 and Aly. *Nature* 413, 644–647. doi: 10.1038/35098106
- Mackenzie, I. R., Arzberger, T., Kremmer, E., Troost, D., Lorenzl, S., Mori, K., et al. (2013). Dipeptide repeat protein pathology in *C9ORF72* mutation cases: clinico-pathological correlations. *Acta Neuropathol.* 126, 859–879. doi: 10.1007/s00401-013-1181-y

- Mackenzie, I. R., Frick, P., and Neumann, M. (2014). The neuropathology associated with repeat expansions in the *C9ORF72* gene. *Acta Neuropathol.* 127, 347–357. doi: 10.1007/s00401-013-1232-4
- Mann, D. M., Rollinson, S., Robinson, A., Bennion Callister, J., Thompson, J. C., Snowden, J. S., et al. (2013). 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. doi: 10.1186/2051-5960-1-68
- Matoulkova, E., Michalova, E., Vojtesek, B., and Hrstka, R. (2012). The role of the 3' untranslated region in post-transcriptional regulation of protein expression in mammalian cells. *RNA Biol.* 9, 563–576. doi: 10.4161/rna.20231
- May, S., Hornburg, D., Schludi, M. H., Arzberger, T., Rentzsch, K., Schwenk, B. M., et al. (2014). *C9orf72* FTL/ALS-associated Gly-Ala dipeptide repeat proteins cause neuronal toxicity and Unc119 sequestration. *Acta Neuropathol.* 128, 485–503. doi: 10.1007/s00401-014-1329-4
- Mizielinska, S., Grönke, S., Niccoli, T., Ridler, C. E., Clayton, E. L., Devoy, A., et al. (2014). *C9orf72* repeat expansions cause neurodegeneration in *Drosophila* through arginine-rich proteins. *Science* 345, 1192–1194. doi: 10.1126/science.1256800
- Mizielinska, S., Lashley, T., Norona, F. E., Clayton, E. L., Ridler, C. E., Fratta, P., et al. (2013). *C9orf72* frontotemporal lobar degeneration is characterised by frequent neuronal sense and antisense RNA foci. *Acta Neuropathol.* 126, 845–857. doi: 10.1007/s00401-013-1200-z
- Mokas, S., Mills, J. R., Garreau, C., Fournier, M. J., Robert, F., Arya, P., et al. (2009). Uncoupling stress granule assembly and translation initiation inhibition. *Mol. Biol. Cell* 20, 2673–2683. doi: 10.1091/mbc.e08-10-1061
- Monahan, Z., Shewmaker, F., and Pandey, U. B. (2016). Stress granules at the intersection of autophagy and ALS. *Brain Res.* 1649, 189–200. doi: 10.1016/j.brainres.2016.05.022
- Mori, K., Lammich, S., Mackenzie, I. R., Forné, I., Zilow, S., Kretzschmar, H., et al. (2013a). 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. doi: 10.1007/s00401-013-1088-7
- Mori, K., Weng, S. M., Arzberger, T., May, S., Rentzsch, K., Kremmer, E., et al. (2013b). The *C9orf72* GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTL/ALS. *Science* 339, 1335–1338. doi: 10.1126/science.1232927
- Murakami, T., Qamar, S., Lin, J. Q., Schierle, G. S., Rees, E., Miyashita, A., et al. (2015). ALS/FTD mutation-induced phase transition of FUS liquid droplets and reversible hydrogels into irreversible hydrogels impairs RNP granule function. *Neuron* 88, 678–690. doi: 10.1016/j.neuron.2015.10.030
- Nataf, S., and Pays, L. (2015). Gene co-expression analysis unravels a link between *C9orf72* and RNA metabolism in myeloid cells. *Acta Neuropathol. Commun.* 3:64. doi: 10.1186/s40478-015-0242-y
- Neumann, M., Sampathu, D. M., Kwong, L. K., Truax, A. C., Micsenyi, M. C., Chou, T. T., et al. (2006). Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* 314, 130–133. doi: 10.1126/science.1134108
- Niblock, M., Smith, B. N., Lee, Y. B., Sardone, V., Topp, S., Troakes, C., et al. (2016). Retention of hexanucleotide repeat-containing intron in *C9orf72* mRNA: implications for the pathogenesis of ALS/FTD. *Acta Neuropathol. Commun.* 4:18. doi: 10.1186/s40478-016-0289-4
- Nishimura, A. L., Zupunski, V., Troakes, C., Kathe, C., Fratta, P., Howell, M., et al. (2010). Nuclear import impairment causes cytoplasmic trans-activation response DNA-binding protein accumulation and is associated with frontotemporal lobar degeneration. *Brain* 133, 1763–1771. doi: 10.1093/brain/awq111
- O'Rourke, J. G., Bogdanik, L., Yáñez, A., Lall, D., Wolf, A. J., Muhammad, A. K., et al. (2016). *C9orf72* is required for proper macrophage and microglial function in mice. *Science* 351, 1324–1329. doi: 10.1126/science.aaf1064
- Patel, A., Lee, H. O., Jawerth, L., Maharana, S., Jahnel, M., Hein, M. Y., et al. (2015). A liquid-to-solid phase transition of the ALS protein FUS accelerated by disease mutation. *Cell* 163, 1066–1077. doi: 10.1016/j.cell.2015.07.047
- Pelechano, V., and Steinmetz, L. M. (2013). Gene regulation by antisense transcription. *Nat. Rev. Genet.* 14, 880–893. doi: 10.1038/nrg3594
- Polymenidou, M., Lagier-Tourenne, C., Hutt, K. R., Huelga, S. C., Moran, J., Liang, T. Y., et al. (2011). Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. *Nat. Neurosci.* 14, 459–468. doi: 10.1038/nn.2779
- Prudencio, M., Belzil, V. V., Batra, R., Ross, C. A., Gendron, T. F., Pregent, L. J., et al. (2015). Distinct brain transcriptome profiles in *C9orf72*-associated and sporadic ALS. *Nat. Neurosci.* 18, 1175–1182. doi: 10.1038/nn.4065
- Reddy, K., Schmidt, M. H., Geist, J. M., Thakkar, N. P., Panigrahi, G. B., Wang, Y. H., et al. (2014). Processing of double-R-loops in (CAG)_n(CTG)_n and *C9orf72* (GGGGCC)_n(GGCCCC) repeats causes instability. *Nucleic Acids Res.* 42, 10473–10487. doi: 10.1093/nar/gku658
- Reddy, K., Zamiri, B., Stanley, S. Y., Macgregor, R. B. Jr., and Pearson, C. E. (2013). The disease-associated r(GGGGCC)_n repeat from the *C9orf72* gene forms tract length-dependent uni- and multi-molecular RNA G-quadruplex structures. *J. Biol. Chem.* 288, 9860–9866. doi: 10.1074/jbc.C113.452532
- Renton, A. E., Majounie, E., Waite, A., Simon-Sanchez, J., Rollinson, S., Gibbs, J. R., et al. (2011). A hexanucleotide repeat expansion in *C9ORF72* is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* 72, 257–268. doi: 10.1016/j.neuron.2011.09.010
- Rossi, S., Serrano, A., Gerbino, V., Giorgi, A., Di Francesco, L., Nencini, M., et al. (2015). Nuclear accumulation of mRNAs underlies G4C2 repeat-induced translational repression in a cellular model of *C9orf72* ALS. *J. Cell Sci.* 128, 1787–1799. doi: 10.1242/jcs.165332
- Sareen, D., O'Rourke, J. G., Meera, P., Muhammad, A. K., Grant, S., Simpkinson, M., et al. (2013). Targeting RNA foci in iPSC-derived motor neurons from ALS patients with a *C9ORF72* repeat expansion. *Sci. Transl. Med.* 5:208ra149. doi: 10.1126/scitranslmed.3007529
- Schuldi, M. H., May, S., Grässer, F. A., Rentzsch, K., Kremmer, E., Kupper, C., et al. (2015). Distribution of dipeptide repeat proteins in cellular models and *C9orf72* mutation cases suggests link to transcriptional silencing. *Acta Neuropathol.* 4, 537–555. doi: 10.1007/s00401-015-1464-6
- Sellier, C., Campanari, M. L., Julie Corbier, C., Gaucherot, A., Kolb-Cheynel, I., Oulad-Abdelghani, M., et al. (2016). Loss of *C9ORF72* impairs autophagy and synergizes with polyQ Ataxin-2 to induce motor neuron dysfunction and cell death. *EMBO J.* 35, 1276–1297. doi: 10.15252/emboj.201593350
- Sephton, C. F., Cenik, C., Kucukural, A., Dammer, E. B., Cenik, B., Han, Y., et al. (2011). Identification of neuronal RNA targets of TDP-43-containing ribonucleoprotein complexes. *J. Biol. Chem.* 286, 1204–1215. doi: 10.1074/jbc.M110.190884
- Skourti-Stathaki, K., Kamieniarz-Gdula, K., and Proudfoot, N. J. (2014). R-loops induce repressive chromatin marks over mammalian gene terminators. *Nature* 516, 436–439. doi: 10.1038/nature13787
- Sofola, O. A., Jin, P., Qin, Y., Duan, R., Liu, H., de Haro, M., et al. (2007). RNA-binding proteins hnRNP A2/B1 and CUGBP1 suppress fragile X CCG premutation repeat-induced neurodegeneration in a *Drosophila* model of FXTAS. *Neuron* 55, 565–571. doi: 10.1016/j.neuron.2007.07.021
- Sreedharan, J., Blair, I. P., Tripathi, V. B., Hu, X., Vance, C., Rogelj, B., et al. (2008). TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. *Science* 319, 1668–1672. doi: 10.1126/science.1154584
- Subramanian, M., Rage, F., Tabet, R., Flatter, E., Mandel, J. L., and Moine, H. (2011). G-quadruplex RNA structure as a signal for neurite mRNA targeting. *EMBO Rep.* 12, 697–704. doi: 10.1038/embor.2011.76
- Sullivan, P. M., Zhou, X., Robins, A. M., Paushter, D. H., Kim, D., Smolka, M. B., et al. (2016). The ALS/FTLD associated protein *C9orf72* associates with SMCR8 and WDR41 to regulate the autophagy-lysosome pathway. *Acta Neuropathol. Commun.* 4:51. doi: 10.1186/s40478-016-0324-5
- Sun, M., Bell, W., LaClair, K. D., Ling, J. P., Han, H., Kageyama, Y., et al. (2017). Cryptic exon incorporation occurs in Alzheimer's brain lacking TDP-43 inclusion but exhibiting nuclear clearance of TDP-43. *Acta Neuropathol.* 133, 923–931. doi: 10.1007/s00401-017-1701-2
- Tao, Z., Wang, H., Xia, Q., Li, K., Li, K., Jiang, X., et al. (2015). Nucleolar stress and impaired stress granule formation contribute to *C9orf72* RAN translation-induced cytotoxicity. *Hum. Mol. Genet.* 24, 2426–2441. doi: 10.1093/hmg/ddv005
- Tian, B., Hu, J., Zhang, H., and Lutz, C. S. (2005). A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Res.* 33, 201–212. doi: 10.1093/nar/gki158
- Timchenko, N. A., Cai, Z. J., Welm, A. L., Reddy, S., Ashizawa, T., and Timchenko, L. T. (2001). RNA CUG repeats sequester CUGBP1 and alter

- protein levels and activity of CUGBP1. *J. Biol. Chem.* 276, 7820–7826. doi: 10.1074/jbc.M005960200
- Tollervy, J. R., Curk, T., Rogelj, B., Briese, M., Cereda, M., Kaykici, M., et al. (2011). Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nat. Neurosci.* 14, 452–458. doi: 10.1038/nn.2778
- Tran, H., Almeida, S., Moore, J., Gendron, T. F., Chalasani, U., Lu, Y., et al. (2015). Differential toxicity of nuclear RNA foci versus dipeptide repeat proteins in a *Drosophila* model of *C9ORF72* FTD/ALS. *Neuron* 87, 1207–1214. doi: 10.1016/j.neuron.2015.09.015
- Tripathi, V., Ellis, J. D., Shen, Z., Song, D. Y., Pan, Q., Watt, A. T., et al. (2010). The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol. Cell* 39, 925–938. doi: 10.1016/j.molcel.2010.08.011
- van Blitterswijk, M., Gendron, T. F., Baker, M. C., DeJesus-Hernandez, M., Finch, N. A., Brown, P. H., et al. (2015). Novel clinical associations with specific *C9ORF72* transcripts in patients with repeat expansions in *C9ORF72*. *Acta Neuropathol.* 130, 863–876. doi: 10.1007/s00401-015-1480-6
- van der Zee, J., Gijssels, I., Dillen, L., Van Langenhove, T., Theuns, J., Engelborghs, S., et al. (2013). A pan-European study of the *C9orf72* repeat associated with FTL: geographic prevalence, genomic instability, and intermediate repeats. *Hum. Mutat.* 34, 363–373. doi: 10.1002/humu.22244
- Wahba, L., Amon, J. D., Koshland, D., and Vuica-Ross, M. (2011). RNase H and multiple RNA biogenesis factors cooperate to prevent RNA:DNA hybrids from generating genome instability. *Mol. Cell* 44, 978–988. doi: 10.1016/j.molcel.2011.10.017
- Waite, A. J., Bäumer, D., East, S., Neal, J., Morris, H. R., Ansoorge, O., et al. (2014). 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. doi: 10.1016/j.neurobiolaging.2014.01.016
- Wan, J., Yourshaw, M., Mamsa, H., Rudnik-Schöneborn, S., Menezes, M. P., Hong, J. E., et al. (2012). Mutations in the RNA exosome component gene *EXOSC3* cause pontocerebellar hypoplasia and spinal motor neuron degeneration. *Nat. Genet.* 44, 704–708. doi: 10.1038/ng.2254
- Ward, M. E., Taubes, A., Chen, R., Miller, B. L., Sephton, C. F., Gelfand, J. M., et al. (2014). Early retinal neurodegeneration and impaired Ran-mediated nuclear import of TDP-43 in progranulin-deficient FTL. *J. Exp. Med.* 211, 1937–1945. doi: 10.1084/jem.20140214
- Webster, C. P., Smith, E. F., Bauer, C. S., Moller, A., Hautbergue, G. M., Ferraiuolo, L., et al. (2016). The *C9orf72* protein interacts with Rab1a and the ULK1 complex to regulate initiation of autophagy. *EMBO J.* 35, 1656–1676. doi: 10.15252/embj.201694401
- Weitzer, S., Hanada, T., Penninger, J. M., and Martinez, J. (2015). CLP1 as a novel player in linking tRNA splicing to neurodegenerative disorders. *Wiley Interdiscip. Rev. RNA* 6, 47–63. doi: 10.1002/wrna.1255
- Wen, X., Tan, W., Westergard, T., Krishnamurthy, K., Markandaiah, S. S., Shi, Y., et al. (2014). 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. doi: 10.1016/j.neuron.2014.12.010
- Wolozin, B. (2012). Regulated protein aggregation: stress granules and neurodegeneration. *Mol. Neurodegener.* 7:56. doi: 10.1186/1750-1326-7-56
- Wong, J. J., Ritchie, W., Ebner, O. A., Selbach, M., Wong, J. W., Huang, Y., et al. (2013). Orchestrated intron retention regulates normal granulocyte differentiation. *Cell* 154, 583–595. doi: 10.1016/j.cell.2013.06.052
- Xiao, S., MacNair, L., McGoldrick, P., McKeever, P. M., McLean, J. R., Zhang, M., et al. (2015). Isoform specific antibodies reveal distinct subcellular localizations of *C9orf72* in amyotrophic lateral sclerosis. *Ann. Neurol.* 78, 568–583. doi: 10.1002/ana.24469
- Xiao, S., MacNair, L., McLean, J., McGoldrick, P., McKeever, P., Soleimani, S., et al. (2016). *C9orf72* isoforms in amyotrophic lateral sclerosis and frontotemporal lobar degeneration. *Brain Res.* 1647, 43–49. doi: 10.1016/j.brainres.2016.04.062
- Xu, Z., Poidevin, M., Li, X., Li, Y., Shu, L., Nelson, D. L., et al. (2013). Expanded GGGGCC repeat RNA associated with amyotrophic lateral sclerosis and frontotemporal dementia causes neurodegeneration. *Proc. Natl. Acad. Sci. U S A* 110, 7778–7783. doi: 10.1073/pnas.1219643110
- Zhang, Z., Almeida, S., Lu, Y., Nishimura, A. L., Peng, L., Sun, D., et al. (2013). Downregulation of microRNA-9 in iPSC-derived neurons of FTD/ALS patients with TDP-43 mutations. *PLoS One* 8:e76055. doi: 10.1371/journal.pone.0076055
- Zhang, K., Donnelly, C. J., Haeusler, A. R., Grima, J. C., Machamer, J. B., Steinwald, P., et al. (2015). The *C9orf72* repeat expansion disrupts nucleocytoplasmic transport. *Nature* 525, 56–61. doi: 10.1038/nature14973
- Zhang, Y.-J., Gendron, T. F., Grima, J. C., Sasaguri, H., Jansen-West, K., Xu, Y. F., et al. (2016). *C9ORF72* poly(GA) aggregates sequester and impair HR23 and nucleocytoplasmic transport proteins. *Nat. Neurosci.* 5, 668–677. doi: 10.1038/nn.4272
- Zhang, J., Ito, H., Wate, R., Ohnishi, S., Nakano, S., and Kusaka, H. (2006). Altered distributions of nucleocytoplasmic transport-related proteins in the spinal cord of a mouse model of amyotrophic lateral sclerosis. *Acta Neuropathol.* 112, 673–680. doi: 10.1007/s00401-006-0130-4
- Zhang, Y. J., Jansen-West, K., Xu, Y. F., Gendron, T. F., Bieniek, K. F., Lin, W. L., et al. (2014). Aggregation-prone c9FTD/ALS poly(GA) RAN-translated proteins cause neurotoxicity by inducing ER stress. *Acta Neuropathol.* 128, 505–524. doi: 10.1007/s00401-014-1336-5
- Zu, T., Gibbens, B., Doty, N. S., Gomes-Pereira, M., Huguet, A., Stone, M. D., et al. (2011). Non-ATG-initiated translation directed by microsatellite expansions. *Proc. Natl. Acad. Sci. U S A* 108, 260–265. doi: 10.1073/pnas.1013343108
- Zu, T., Liu, Y., Bañez-Coronel, M., Reid, T., Pletnikova, O., Lewis, J., et al. (2013). RAN proteins and RNA foci from antisense transcripts in *C9ORF72* ALS and frontotemporal dementia. *Proc. Natl. Acad. Sci. U S A* 110, E4968–E4977. doi: 10.1073/pnas.1315438110

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Barker, Niblock, Lee, Shaw and Gallo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.