



OPEN ACCESS

EDITED BY

Tetsuya Takano,
PRESTO, Japan Science and Technology
Agency (JST), Japan

REVIEWED BY

Bilal Malik,
University College London, United Kingdom

*CORRESPONDENCE

Qixu Cai
✉ qxcai@xmu.edu.cn

RECEIVED 18 October 2023

ACCEPTED 04 January 2024

PUBLISHED 22 January 2024

CITATION

Geng Y and Cai Q (2024) Role of C9orf72
hexanucleotide repeat expansions in
ALS/FTD pathogenesis.
Front. Mol. Neurosci. 17:1322720.
doi: 10.3389/fnmol.2024.1322720

COPYRIGHT

© 2024 Geng and Cai. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Role of C9orf72 hexanucleotide repeat expansions in ALS/FTD pathogenesis

Yanyan Geng¹ and Qixu Cai^{2*}

¹Clinical Research Institute of the First Affiliated Hospital of Xiamen University, Fujian Key Laboratory of Brain Tumors Diagnosis and Precision Treatment, Xiamen Key Laboratory of Brain Center, the First Affiliated Hospital of Xiamen University, School of Medicine, Xiamen University, Xiamen, Fujian, China, ²State Key Laboratory of Vaccines for Infectious Diseases, School of Public Health, Xiamen University, Xiamen, Fujian, China

Amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) are progressive neurological disorders that share neurodegenerative pathways and features. The most prevalent genetic causes of ALS/FTD is the GGGGCC hexanucleotide repeat expansions in the first intron region of the chromosome 9 open reading frame 72 (C9orf72) gene. In this review, we comprehensively summarize the accumulating evidences elucidating the pathogenic mechanism associated with hexanucleotide repeat expansions in ALS/FTD. These mechanisms encompass the structural polymorphism of DNA and transcribed RNA, the formation of RNA foci via phase separation, and the cytoplasmic accumulation and toxicities of dipeptide-repeat proteins. Additionally, the formation of G-quadruplex structures significantly impairs the expression and normal function of the C9orf72 protein. We also discuss the sequestration of specific RNA binding proteins by GGGGCC RNA, which further contributes to the toxicity of C9orf72 hexanucleotide repeat expansions. The deeper understanding of the pathogenic mechanism of hexanucleotide repeat expansions in ALS/FTD provides multiple potential drug targets for these devastating diseases.

KEYWORDS

C9orf72, hexanucleotide repeat expansions, G-quadruplex, phase separation, amyotrophic lateral sclerosis, frontotemporal dementia, neurodegenerative diseases

1 Introduction

Amyotrophic lateral sclerosis (ALS) is a neurological disorder mainly characterized by the degeneration of motor neurons, leading to symptoms including muscle weakness, muscle atrophy and ultimately death from respiratory failure (Rowland and Shneider, 2001), while frontotemporal dementia (FTD) refers to a broad spectrum of neurological disorders marked by progressive damage to the temporal and frontal lobes of the brain (Graff-Radford and Woodruff, 2007; Rademakers et al., 2012). Presently, both ALS and FTD are recognized as part of a broad neurodegenerative continuum that shares common neurodegenerative pathways and features. This includes the presence of TAR DNA-binding protein 43 (TDP-43) cytoplasmic inclusions within the central nervous system (Neumann et al., 2006; Lillo and Hodges, 2009), which was considered as a potential biomarker for

ALS/FTD (Majumder et al., 2018). Furthermore, both ALS and FTD were genetically linked to the expanded GGGGCC (G4C2) hexanucleotide repeats in the first intron region of the chromosome 9 open reading frame 72 (C9orf72) gene (Figure 1A). This genetic abnormality was identified as the most common genetic cause of familial ALS and familial FTD (DeJesus-Hernandez et al., 2011; Renton et al., 2011; Majounie et al., 2012).

Tremendous progress and hypothesis have greatly contributed to our understanding of the pathologic mechanism of C9orf72 hexanucleotide repeat expansions in ALS/FTD (Edbauer and Haass, 2016; Gitler and Tsuiji, 2016; Haeusler et al., 2016; Goodman et al., 2019). Both the C9orf72 hexanucleotide repeat expansions and the transcribed RNA, were reported to adopt polymorphic unusual secondary structures, such as G-quadruplexes and hairpin, which were involved in the ALS/FTD pathogenesis via several potential mechanisms (Fratta et al., 2012; Reddy et al., 2013; Haeusler et al., 2014). Moreover, the non-ATG translation of the hexanucleotide repeat expansions leads to the accumulated production of cytoplasmic dipeptide-repeat (DPR) proteins, whose toxicity is also considered as a contributing factor to the pathology of diseases (Ash et al., 2013; Gendron et al., 2013; Mackenzie et al., 2013; Mori et al., 2013a,c). These findings represent crucial steps in unraveling the complex pathogenic mechanisms associated with C9orf72 hexanucleotide repeat expansions in ALS/FTD.

In this mini review, we focus on the pivotal role played by C9orf72 hexanucleotide repeat expansions in the pathogenesis of ALS/FTD across multiple dimensions, as visually represented in Figure 1. We present a comprehensive summary of recent progress, highlighting the structural polymorphism observed in both the DNA and transcribed RNA of C9orf72 hexanucleotide repeat expansions. Additionally, we explore the mechanism of the formation of RNA foci via phase separation, elucidating its profound implications for the pathogenesis of ALS/FTD. Moreover, we delve into the adverse effects of G-quadruplex formation, which significantly impairs the expression and normal function of C9orf72 protein. Meanwhile, we discuss the cytoplasmic accumulation and toxicity of DPR proteins, as well as the role of sequestered and enriched RNA binding proteins (RBPs) by G4C2 RNA, all of which contribute substantially to the overall toxicity related to hexanucleotide repeat expansions in ALS/FTD.

2 Structural polymorphism of G4C2 repeat expansions of DNA

G-quadruplex refers to a unique four-stranded helical structure which is formed by guanosine-rich DNA or RNA sequences (Sen and Gilbert, 1988). The fundamental structural unit of a G-quadruplex, known as a G-tetrad, is composed of four guanine molecules arranged in a cyclic Hoogsteen hydrogen-bonded square planar structure (Sannohe and Sugiyama, 2010). Two or more G-tetrads stack to form a G-quadruplex, which is stabilized by monovalent cations (Huppert, 2008). In the context of ALS/FTD, it was reported that the guanosine-rich C9orf72 G4C2 hexanucleotide repeat expansions can fold into polymorphic G-quadruplex *in vitro* (Haeusler et al., 2014; Brcic and Plavec, 2015, 2018; Sket et al., 2015; Zamiri et al., 2015; Zhou et al., 2015, 2018; Figure 1A). Notably, the highly variable size of the hexanucleotide repeat expansions

(DeJesus-Hernandez et al., 2011; Renton et al., 2011; Haeusler et al., 2016), as well as the complexity of the G-quadruplex formed by G4C2 repeat sequence, give rise to a dramatic structural diversity of G-quadruplex structures folded by the C9orf72 G4C2 repeat expansions (Haeusler et al., 2014; Zhou et al., 2015, 2018).

In recent years, various biophysical methods have been employed to elucidate the folding and structures of G4C2 repeat expansions with varying length, highlighting the structural polymorphism of G4C2 repeat expansions and the potential existence of multiple architectures. Using circular dichroism (CD) and nuclear magnetic resonance (NMR) spectroscopy, Zhou et al. (2015) investigated the structural heterogeneity of two, three, four and five repeats of G4C2 DNA (referred to as d(G4C2)₂, d(G4C2)₃, d(G4C2)₄ and d(G4C2)₄, respectively). They further employed NMR to determine the topology of the G-quadruplex formed by d(G4C2)₄, revealing a monomeric chair-type G-quadruplex with a four-layer antiparallel G-tetra core and three edgewise loops (Zhou et al., 2015).

Brcic and Plavec (2015) carried out an *in vitro* investigation involving several mutational variants of d(G4C2)₄ oligonucleotides. Notably, they observed that d[(G4C2)₃G₄] could adopt two predominant structures with antiparallel topologies (Brcic and Plavec, 2015). To enhance structural stability, they introduced a substitution of dG₂₁ with 8-bromoguanine (8Br-dG), a residue that favors a syn glycosidic conformation (Sarma et al., 1974). Interestingly, they found that the folding conditions, including solution pH and the rate of cooling, exert a substantial influence on the structure. They solved the structure of d[(G4C2)₃G₄] formed in a neutral pH solution with slow cooling (referred to as NAN for neutral and annealing) (Brcic and Plavec, 2015), as well as an alternative structure formed in an acidic pH solution with fast cooling (termed AQU for acidic and quenching) using NMR technique (Brcic and Plavec, 2018). These investigations collectively demonstrate the structural polymorphism of G-quadruplex formed by G4C2 repeat expansions (Figure 1A).

Due to the structural diversity inherent in G-quadruplex formed by G4C2 repeat expansions, Zhou et al. (2018) employed biochemical purification methods to isolate and prepare these G-quadruplex *in vitro*. Initial screening using CD and NMR methods revealed that d(G4C2)₂ could form mixed parallel and antiparallel topologies (Zhou et al., 2015). Subsequently, using anion exchange chromatography, they successfully separated different conformations of d(G4C2)₂. The various purified fractions were subjected to CD melting experiments and NMR spectroscopy, revealing distinct properties of each fractions (Zhou et al., 2018). Notably, one of the major fractions, designated as the F5 fraction, was successfully crystallized in a K⁺ solution (Zhou et al., 2018). To solve the phase problem, d(G4C2)₂ was annealed in Ba²⁺ solution, leading to the crystallization of d(G4C2)₂-Ba. With the help of the anomalous signal of Ba²⁺, the structure of d(G4C2)₂-Ba was solved and subsequently used as the template to solve the structure of d(G4C2)₂-K by molecular replacement method. The structure revealed that d(G4C2)₂ folds into a novel eight-layer parallel tetrameric G-quadruplex (Geng et al., 2021). Interestingly, the crystal structure displayed two distinct forms of tetrameric G-quadruplexes, referred to as Form-1/7 (where the dG₁ base in one dimeric block stacks with the dG₇ base in the opposite dimeric block) and Form-1/1 (where the dG₁ base

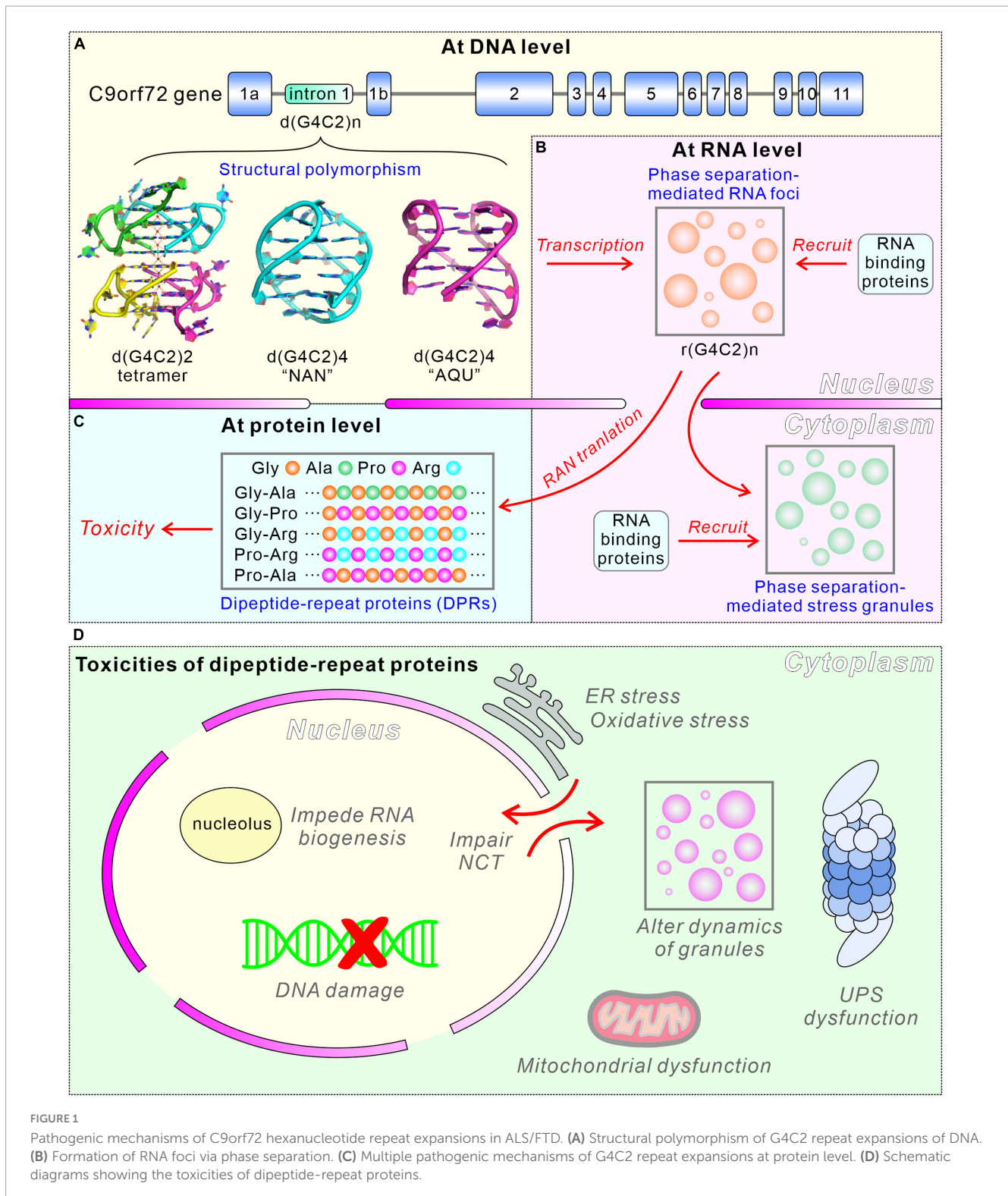


FIGURE 1

Pathogenic mechanisms of C9orf72 hexanucleotide repeat expansions in ALS/FTD. (A) Structural polymorphism of G4C2 repeat expansions of DNA. (B) Formation of RNA foci via phase separation. (C) Multiple pathogenic mechanisms of G4C2 repeat expansions at protein level. (D) Schematic diagrams showing the toxicities of dipeptide-repeat proteins.

in one dimeric block stacks with the dG1 base in the opposite dimeric block) respectively. These novel multimeric G-quadruplex structures formed by d(G4C2)₂ underscore the complexity and potential for oligomerization of G-quadruplexes generated by G4C2 repeat expansions (Figure 1A).

In order to comprehend the functional consequence and pathological implications of the structural polymorphism of

G4C2 repeat expansions, Haeusler et al. (2014) examined the transcription of the G4C2 repeat expansions using an *in vitro* transcription assay. Their findings unveiled a repeat-length-dependent accumulation of transcripts that were prematurely terminated within the repeat expansions. The abortive transcription was also characterized in patient cells with C9orf72 G4C2 repeat expansions. Moreover, they identified an

essential nucleolar protein, nucleolin, which specifically binds to these transcripts. This interaction led to impaired nucleolar function and nucleolar stress in patient cells, thus it is considered as one of the pathogenic mechanisms underlying C9orf72-linked ALS/FTD (Haeusler et al., 2014) (see below). Nevertheless, it is still unclear how the diverse structures formed by G4C2 repeat expansions impact the specific pathogenic mechanisms.

3 Formation of RNA foci via phase separation

It has been proved that the transcribed RNA from G4C2 repeat expansions [termed r(G4C2)_n] can indeed form various structures, including G-quadruplexes (Fratta et al., 2012; Reddy et al., 2013; Haeusler et al., 2014), hairpin structures (Su et al., 2014) and duplex structures (Maity et al., 2021). These diverse structures have been confirmed through topological investigations using gel-shift assay, CD spectroscopy and NMR spectroscopy. Furthermore, it has been observed that the equilibrium between G-quadruplex, hairpin structure and duplex structure was regulated by factors like pH and temperature (Bozic et al., 2020). This phenomenon is reminiscent of the distinct G-quadruplex structures formed by d[(G4C2)₃G4], as previously discussed, influenced by pH and rate of cooling (Brcic and Plavec, 2015, 2018). However, it's important to note that the detailed molecular structure of G-quadruplex formed by r(G4C2)_n has not been resolved yet.

The accumulation of RNA foci in the nucleus, formed by transcripts containing the G4C2 repeat expansions, is a recognized pathological hallmark shared by ALS/FTD. Phase separation of biological macromolecules have been extensively studied, and plays a crucial role in the formation and regulation of membraneless organelles (Banani et al., 2017; Shin and Brangwynne, 2017; Wu et al., 2020; Roden and Gladfelter, 2021), particularly in the neurodegenerative diseases (Zbinden et al., 2020). Jain and Vale (2017) made a significant discovery by observing that G4C2 repeat expansions undergo a solution-gel phase separation *in vitro* at a critical repeat number similar to what is observed in ALS/FTD patients. The G4C2 repeat-containing RNA can form RNA foci via phase separation in human cells, suggesting that the gelation of G4C2 repeat-containing RNA contributes to the pathology of ALS/FTD disease via phase separation (Fay et al., 2017; Jain and Vale, 2017; Figure 1B).

Phase separation of biological macromolecules depends on the multivalent intermolecular interactions (Banani et al., 2017; Shin and Brangwynne, 2017). Given the tetrameric structure of G-quadruplex formed by d(G4C2)₂ (Geng et al., 2021), it is reasonable to speculate that G-quadruplex formed by r(G4C2)₂ is also multimeric. The multimeric organization provides the specific multivalent intermolecular interactions, but not the weak and unspecific interactions between RNA, necessary for the phase separation of G4C2 repeat-containing RNA (Shen et al., 2023). Furthermore, the phase separation of RNA occurs at a boundary condition of increasing effective valence contributed by the increasing repeat number and specific intermolecular interaction. This could explain why ALS/FTD disease is triggered after the G4C2 repeat expansions reach a certain threshold of repeat number (Haeusler et al., 2016; Jain and Vale, 2017). Additionally, the

highly enriched condensates of r(G4C2)_n can further recruit RNA-binding proteins through phase separation. This contributes to the formation of RNA granules, including the nuclear RNA foci and cytoplasmic stress granules, all of which play a role in the pathogenesis of the diseases (Fay et al., 2017; Figure 1B).

4 Multiple pathogenic mechanisms of G4C2 repeat expansions at protein level

4.1 Loss-of-function pathogenic mechanisms

The C9orf72 protein interacts with Smith-Magenis syndrome chromosome region candidate gene 8 (SMCR8) and WD repeat-containing protein 41 (WDR41) (Amick et al., 2016; Sullivan et al., 2016), forming a stable ARF GTPase-activating protein (GAP) complex (Su et al., 2020, 2021; Tang et al., 2020; Norpel et al., 2021). This ARF GAP complex was reported to be involved in regulation of autophagy (Amick et al., 2016; Sullivan et al., 2016; Yang et al., 2016), and is critical for microglial function (O'Rourke et al., 2016) and the modulation of actin dynamic in motor neurons (Sivadasan et al., 2016). The unusual secondary structures formed by the G4C2 repeat expansions in the first intron region of C9orf72 gene significantly reduce the expression of C9orf72 protein by the abortive transcription (DeJesus-Hernandez et al., 2011; Renton et al., 2011; Haeusler et al., 2014; Shi et al., 2018). Additionally, the epigenetic mechanisms, including the methylation of cytosine and histone, also contribute to the reduction of C9orf72 expression (Belzil et al., 2013; Xi et al., 2014). For the human induced motor neurons, the exogenous restoration of C9orf72 expression can rescue the survival of neurons (Shi et al., 2018), suggesting that the decrease of C9orf72 protein contributes to the pathogenesis of ALS/FTD. However, the C9orf72 loss-of-function mouse model studies have found that the knock down or knockout of C9orf72 gene does not recapitulate ALS/FTD (Lagier-Tourenne et al., 2013; Koppers et al., 2015; Atanasio et al., 2016; Burberry et al., 2016; Jiang et al., 2016; O'Rourke et al., 2016; Sudria-Lopez et al., 2016; Sullivan et al., 2016; Ugolino et al., 2016; Liu et al., 2022), although some models showed mild motor or cognitive phenotypes or reduced survival (Table 1; Atanasio et al., 2016; Burberry et al., 2016; Jiang et al., 2016; Sudria-Lopez et al., 2016; Ugolino et al., 2016; Batra and Lee, 2017; Balendra and Isaacs, 2018; Liu et al., 2022). This indicates that C9orf72 loss of function is not sufficient to lead to ALS/FTD, further highlighting the complex nature of these neurodegenerative diseases.

4.2 Toxicity of dipeptide-repeat proteins

The transcribed r(G4C2)_n can undergo a mechanism called repeat-associated non-ATG dependent translation (RAN translation), leading to the production of dipeptide-repeat (DPR) proteins (Ash et al., 2013; Mori et al., 2013a,c; Zu et al., 2013). This non-ATG initiated translation of the G4C2 repeat expansions from all six potential reading frames can generate five different DPR

TABLE 1 C9orf72 mouse models.

	Methods	Phenotypes
Loss of function	C9orf72 depletion in the adult mouse CNS by ASO injection (Lagier-Tourenne et al., 2013)	Normal
	Neural-specific ablation of C9orf72 (Koppers et al., 2015)	Reduced body weight; Normal motor function.
	Remove exons 2-6 of C9orf72 (Jiang et al., 2016)	Reduced body weight; Splenomegaly and enlarged lymph nodes; Decreased hemoglobin and packed cell volume; Decreased percentage of lymphocytes; Increased percentage of neutrophils in blood; Mild social interaction and social recognition abnormalities; Mild motor deficits.
	Knockout of C9orf72 full gene (Atanasio et al., 2016)	Mild motor deficits; Lymphadenopathy and splenomegaly; Mixed inflammatory infiltrates in multiple organs; Systemic lupus erythematosus-like disease; Reduced survival.
	Model 1: Remove exons 2-6 of C9orf72 Model 2: ZFN-mediated knockout of C9orf72 (O'Rourke et al., 2016)	Progressive splenomegaly and lymphadenopathy; Age-related neuroinflammation
	Full ablation of C9orf72 (Sudria-Lopez et al., 2016)	Reduced survival; Reduced body weight; Normal motor function; Enlarged lymph nodes and splenomegaly.
	CRISPR/Cas9-mediated knockout of C9orf72 (Sullivan et al., 2016)	No growth defects; Lymph node and spleen enlargement.
	Model 1: Remove exons 2-6 of C9orf72 Model 2: Remove exons 2-6 of C9orf72 without selection cassette Model 3: CRISPR/Cas9-mediated knockout of C9orf72 (Burberry et al., 2016)	Splenomegaly; Neutrophilia; Thrombocytopenia; High mortality rate.
	Remove exons 2-6 of C9orf72 without selection cassette (Ugolino et al., 2016)	Decreased life span; Splenomegaly; No obvious neuronal cell death in brain or spinal cord.
	Jackson Lab Stock No. 027068 (O'Rourke et al., 2016); ZFN-mediated knockout of C9orf72 (Liu et al., 2022)	Motor deficits; Hyperactivity of Purkinje cells; Enhanced BK channel in the cerebellum.
Gain of function	Expression of (G4C2) ₆₆ throughout CNS by AAV-mediated somatic brain transgenesis (Chew et al., 2015)	Nuclear RNA foci; Inclusions of poly(GP), poly(GA) and poly(GR) DPR inclusions; TDP-43 pathology; Cortical neuron and cerebellar Purkinje cell loss; Astrogliosis; Reduced body weight; Behavioral abnormalities.
	BAC containing exons 1-6 of human C9orf72 gene with ~500 repeats of G4C2 motif (Peters et al., 2015)	Sense and antisense RNA foci; Poly(GP) DPR inclusions; No TDP-43 pathology; Survival, motor and cognitive systems are normal.
	BAC containing human C9orf72 gene with ~100-1000 repeats of G4C2 motif (O'Rourke et al., 2016)	Widespread RNA foci; Poly(GP) DPR inclusions; Normal behavior; No neurodegeneration.
	BAC containing human C9orf72 gene with up to 500 repeats of G4C2 motif (Liu et al., 2016)	Decreased survival; Kyphosis, reduced activity, hyperactivity when provoked, claspings and intermittent seizures; Motor neuron disease; Neurodegenerative changes; Sense and antisense RNA foci; Poly(GA) and poly(GP) aggregates; Nuclear and cytoplasmic TDP-43 aggregates.
	BAC containing exons 1-5 of human C9orf72 gene with ~110 or ~450 repeats of G4C2 motif (Jiang et al., 2016)	Sense and antisense RNA foci; Cytoplasmic poly(GP), poly(GR) and poly(GA) aggregates; Age-dependent cognitive impairment and anxiety-like behavior; No TDP-43 mislocalization or aggregation with increased phosphorylated TDP-43.
	Expression of GFP-(GA) ₅₀ or GFP-(GA) ₅₀ -mut in the CNS by AAV-mediated somatic brain transgenesis (Zhang et al., 2016)	Poly(GA) aggregation; Brain atrophy and neurotoxicity; Hyperactivity, anxiety-like behavior, motor and cognitive deficits; rare phosphorylated TDP-43 inclusions;
	Expression of (GA) ₁₄₉ , 31 amino acids corresponding to the 3' region of the poly(GA) reading frame in patients and a C-terminal CFP tag (Schludi et al., 2017)	Poly(GA) inclusions; Mild TDP-43 phosphorylation; No TDP-43 inclusions and mislocalization; Microglia activation without astrogliosis; Progressive motor deficits.
	Expression of 10 pure or 102 interrupted G4C2 repeats in the brain by AAV-mediated somatic brain transgenesis (Herranz-Martin et al., 2017)	RNA foci; Extensive DPR pathology; Disease-related NMJ pathology; Progressive gait and behavioral deficits.
	Conditional expression of GFP-(PR) ₂₈ in neurons (Hao et al., 2019)	Poly(PR) aggregation; No TDP-43 inclusions; Decreased survival; Smaller body size; Smaller brain volumes; Motor deficits, motor-related neurodegeneration and gliosis for heterozygous mice; No cytoplasmic TDP-43 inclusions.
	Knockin of 80 G4C2 repeats with human flanking fragments within exon1a and exon1b at the rat C9orf72 locus (Dong et al., 2020)	Motor deficits; Loss of motor neurons; Hind limb paralysis for females.
	Tet-on inducible expression of 36 × pure G4C2 repeats with 100-bp upstream and downstream human flanking regions (Riemsлагh et al., 2021)	Sporadic sense DPR aggregates; No apparent neurodegeneration; No RNA foci; No pTDP-43 pathology; Locomotor phenotype, rapid muscular dystrophy and neuromuscular junction abnormalities.

(Continued)

TABLE 1 (Continued)

	Methods	Phenotypes
	Expression of ATG-driven FLAG-(GR)50-eGFP using FAST cassette (Verdone et al., 2022)	No TDP-43 pathology; No reduction in survival; No motor and cognitive impairments; Mild motor neuron loss in males.
	AAV-mediated expression of (G4C2)149 (Jambeau et al., 2022)	Poly(GA), Poly(GR) and Poly(GP) aggregation; Normal survival; Increases in distance traveled, velocity of movement and time spent moving on an open-field assay.

proteins: Gly-Ala (GA), Gly-Pro (GP) and Gly-Arg (GR) translated by sense (G4C2)_n RNA, as well as Pro-Gly (PG) (equivalent to GP), Pro-Arg (PR) and Pro-Ala (PA) translated by antisense (G4C2)_n RNA. The accumulation of these unconventionally translated DPR proteins leads to the formation of cytoplasmic insoluble inclusions in neurons, which can be identified by specific DPR antibodies (Ash et al., 2013; Mori et al., 2013c; Zu et al., 2013). Thus, DPR proteins are considered as a pathognomonic feature of ALS/FTD. Especially, the formation of polymorphic RNA G-quadruplex formed by r(G4C2)_n can trigger ribosomal frame shifting during translation (Yu et al., 2014), leading to the unconventional translation of DPR proteins.

Two of the Arg-containing DPR proteins, GR and PR, were reported to bind to nucleoli, impede RNA biogenesis and induce cell death, illustrating the toxic effects of DPR proteins (Kwon et al., 2014; Figures 1C, D). Additionally, the Arg-containing DPR proteins directly bind to proteins harboring low complexity domains and impair the assembly, dynamics and function of phase separation-mediated membraneless organelles (Lee et al., 2016; Lin et al., 2016). In iPSC-derived motor neurons, expression of poly(GR) increases oxidative stress and DNA damage, and causes mitochondrial dysfunction (Lopez-Gonzalez et al., 2016). Another DPR, poly(GA), forms ubiquitin/p62-positive inclusions in neuronal cells, indicating the dysfunction of ubiquitin-proteasome system (UPS) leads to the cytotoxicity of DPR proteins (Yamakawa et al., 2015). Poly(GA) also forms abundant inclusions in cells and cerebellar tissue of ALS/FTD patients, and causes impairment of neurite outgrowth, endoplasmic reticulum (ER) stress and neuronal death (Zhang et al., 2014). The mouse model generated by the overexpression of poly(GA) showed the sequestration of proteins involved in proteasomal degradation and nucleocytoplasmic transport (NCT), leading to the neurodegeneration and behavioral abnormalities (Zhang et al., 2016). Overall, these toxic DPR proteins are a key pathological aspect of ALS/FTD, adding to the complexity of these neurodegenerative diseases (Figures 1C, D).

4.3 Other downstream pathogenic mechanisms

Subsequently, the transcribed r(G4C2)_n may sequester specific proteins, which is a potential mechanism of toxicity associated with hexanucleotide repeat expansions at the protein level. Extensive studies have identified a series of proteins that bind to the transcribed G4C2 repeat expansions through various approaches, including adenosine deaminase RNA-specific B2 (ADARB2), nucleolin, purα, TDP-43 and several heterogeneous nuclear ribonucleoproteins (hnRNPs) (Donnelly et al., 2013; Mori et al., 2013b; Xu et al., 2013; Cooper-Knock et al., 2014;

Haeusler et al., 2014). In particular, many of these binding proteins can interact specifically with r(G4C2)_n with special secondary structures, such as G-quadruplex and hairpin structures, underscoring the structural specificity of RNA-protein interactions. The interactions may lead to the accumulation of nuclear or cytoplasmic RNA-protein aggregates (Ramaswami et al., 2013) and RNA granules mediated by liquid-liquid phase separation (Ramaswami et al., 2013; Maharana et al., 2018; Hallegger et al., 2021; Yu et al., 2021; Lu et al., 2022), contributing to the toxicity associated with hexanucleotide repeat expansions (Figure 1B).

Using the stable isotope labeling by amino acids in cell culture (SILAC) method, Haeusler et al. (2014) identified nucleolin specifically recognize the RNA G-quadruplex, which was further confirmed by RNA pull down experiments. Nucleolin is a principal component of the nucleolus (Abdelmohsen and Gorospe, 2012), which was found to mislocalize with the G4C2 RNA foci in the neurons of the motor cortex of C9orf72 ALS patients, leading to nucleolar stress and impaired nucleolar function in patient cells. Importantly, treatment of wild type cells with the 21-repeat-containing abortive transcripts recapitulates the nucleolin pathology, indicating that the specific interaction between nucleolin and RNA G-quadruplex is a fundamental determinant of pathogenic mechanism for ALS/FTD (Haeusler et al., 2014).

Another prominent downstream pathogenic mechanism for ALS/FTD is the dysfunctional nucleocytoplasmic transport (Prpar Mihevc et al., 2017). Using a *Drosophila* model-based screening, Zhang et al. (2015) characterized that RNA G-quadruplex formed by r(G4C2)_n directly interacts with RanGAP and impairs nuclear import, leading to the nuclear pore pathology in the *Drosophila* model and ALS patient-derived iPSC cells. In the other study, Freibaum et al performed a large-scale genetic screen in a *Drosophila* model, and identified 18 genetic modifiers that encode components of the nuclear pore complex and the protein machinery for nucleocytoplasmic transport, demonstrating the compromised nucleocytoplasmic transport in the *Drosophila* model and C9orf72-linked patient-derived iPSC cells (Freibaum et al., 2015).

4.4 C9orf72 gain-of-function mouse models

Besides the C9orf72 loss-of-function mouse models (see above and Table 1), adeno-associated virus (AAV) or bacterial artificial chromosome (BAC)-mediated expression of C9orf72 gene with large G4C2 repeat expansions or DPR proteins was used to generate a number of gain-of-function mouse models (Table 1). With the expression of C9orf72 gene with large G4C2 repeat expansions, most of the mouse model developed both RNA foci and DPR

inclusions (Chew et al., 2015; Peters et al., 2015; O'Rourke et al., 2016; Herranz-Martin et al., 2017), while the mouse model with tet-on inducible expression of 36x pure G4C2 repeats only have sporadic sense DPR aggregates, but not RNA foci (Riemslogh et al., 2021). The mouse models with the overexpression of DPR proteins showed extensive DPR pathology, but not RNA foci (Zhang et al., 2016; Schludi et al., 2017; Hao et al., 2019). TDP-43 inclusions were not usually observed in the gain-of-function mouse models. Only two of the mouse models, (G4C2)⁶⁶ mice and C9-500 mice, displayed TDP-43 inclusion pathology and neurodegenerative phenotypes (Chew et al., 2015; Liu et al., 2016). It was noted that another group reported the absence of survival and motor deficits for C9-500 mice (Mordes et al., 2020; Nguyen et al., 2020a). Overall, the gain-of-function mouse models indicate that large G4C2 repeat expansions greatly contribute to the pathogenesis of ALS/FTD.

5 Potential therapeutic strategies

Understanding the pathogenic mechanism of hexanucleotide repeat expansions in ALS/FTD not only sheds light on the underlying biology of these diseases but also provides multiple potential drug targets for the therapeutic interventions. Various approaches have shown promise as potential treatments for ALS/FTD targeting C9orf72.

5.1 Targeting genomic C9orf72 hexanucleotide repeat expansions

The most straight-forward target is the genomic C9orf72 hexanucleotide repeat expansions. CRISPR/Cas9-mediated excision of G4C2 repeat expansions in neurons and mouse models resulted in reduction of RNA foci and DPR inclusions (Meijboom et al., 2022). Since CRISPR/Cas9-mediated genome editing may result in the risk of creating indels (Selvaraj et al., 2018), the combination of CRISPR/Cas9 genome editing and homology-directed repair (HDR) completely repaired the C9orf72 G4C2 repeat expansions to the wild-type repeat size in iPSC cells derived from ALS/FTD patient, and finally abolished pathological phenotypes (Ababneh et al., 2020). The key points for the development of CRISPR/Cas9-based therapeutic strategy should be the efficient and effective delivery, as well as accurate editing of CRISPR/Cas9.

5.2 Targeting transcribed RNA

As the accumulation of RNA foci is the pathological hallmark and contributes to the pathology of ALS/FTD, the transcribed r(G4C2)_n as a potential drug target for C9orf72-linked ALS/FTD has garnered considerable attention from researchers. One of the therapeutic strategies is degradation of the RNA transcripts using antisense oligonucleotides (ASOs) (Donnelly et al., 2013; Lagier-Tourenne et al., 2013). ASO-mediated degradation of repeat RNA decreased RNA foci and DPR inclusions, as well as ameliorated behavioral deficits (Jiang et al., 2016). Another promising development targeting transcribed r(G4C2)_n is that a small molecule, TMPyP4, was characterized to be able to distort the

G-quadruplex formation of r(G4C2)₈, and ablate the interaction between the G-quadruplex and its binding proteins (Zamiri et al., 2014). Su et al. (2014) further designed and screened small molecules targeting r(G4C2)_n to inhibit repeat-associated non-ATG translation and the formation of RNA foci. The findings from their work suggest that small molecules targeting r(G4C2)_n hold promise as a potential therapeutic approach for C9orf72-linked ALS/FTD (Su et al., 2014).

5.3 Targeting DPR proteins

The toxicity of DPR proteins is the essential part for ALS/FTD pathogenesis, thus the immunotherapeutic approach targeting DPR proteins is another potential therapeutic strategy. In cell and mouse models, the antibodies targeting poly(GA) could reduce poly(GA), poly(GP) and poly(GR) inclusions, improve behavioral deficits, decrease neuroinflammation and neurodegeneration, and increase survival (Zhou et al., 2017, 2020; Nguyen et al., 2020b; Jambeau et al., 2022). On the other hand, it was reported that RAN translation is highly regulated by PKR. The FDA-approved drug, metformin, inhibited PKR, leading to decrease DPR proteins and improve behavior deficits (Zu et al., 2020).

5.4 Targeting other downstream mechanisms

Dysfunctional nucleocytoplasmic transport has been characterized as a critical downstream mechanism for ALS/FTD pathogenesis (Prpar Mihevc et al., 2017). KPT-276, the exportin 1 inhibitor, inhibits nuclear export to compensate for disrupted import by the interaction between r(G4C2)_n and RanGAP, and suppresses the neurodegeneration in the fly eye (Zhang et al., 2015). Thus, the modulation of nucleocytoplasmic transport is also a potential therapeutic strategy for C9orf72-linked neurodegenerative diseases. ER stress induced by DPR proteins is one of the key downstream pathogenic mechanism. It was reported the inhibitors of ER stress could provide rescue against poly(GA)-induced ER stress and neurotoxicity in neurons (Zhang et al., 2014). Furthermore, pharmacological or genetic suppression of oxidative stress and cellular toxicity induced by poly(GR) decreases DNA damage, indicating reducing oxidative stress is a potential therapeutic strategy for C9orf72-linked ALS/FTD (Lopez-Gonzalez et al., 2016). These advancements mark important steps toward the development of targeted therapies for these debilitating neurodegenerative diseases.

6 Conclusion

The hexanucleotide repeat expansions located in the first intron of C9orf72 gene are the most common genetic cause of ALS/FTD. Accumulating research emphasizes the critical role of these hexanucleotide repeat expansions in the pathogenesis of the diseases. The unique sequence of the hexanucleotide, GGGGCC, facilitates the formation of polymorphic secondary structures known as G-quadruplexes, both in the DNA and transcribed RNA. These multimeric G-quadruplexes further promote the formation

of RNA foci via phase separation, when the repeat number exceeds a certain threshold. The formation of unusual secondary structures in hexanucleotide repeat expansions greatly impairs the expression of C9orf72 proteins. Furthermore, the accumulation of toxic DPR proteins produced by unconventional translation (i.e., RAN translation) generates inclusions in neuron, which is considered as a distinctive pathognomonic feature of ALS/FTD. Finally, the transcribed G4C2 RNA and its interaction with binding proteins can trigger the formation of RNA granules via phase separation and sequester specific proteins, contributing to the toxicity associated with hexanucleotide repeat expansions.

Author contributions

YG: Funding acquisition, Writing – original draft, Writing – review and editing. QC: Funding acquisition, Supervision, Writing – original draft, Writing – review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work

References

- Ababneh, N. A., Scaber, J., Flynn, R., Douglas, A., Barbagallo, P., Candalija, A., et al. (2020). Correction of amyotrophic lateral sclerosis related phenotypes in induced pluripotent stem cell-derived motor neurons carrying a hexanucleotide expansion mutation in C9orf72 by CRISPR/Cas9 genome editing using homology-directed repair. *Hum. Mol. Genet.* 29, 2200–2217. doi: 10.1093/hmg/ddaa106
- Abdelmohsen, K., and Gorospe, M. (2012). RNA-binding protein nucleolin in disease. *RNA Biol.* 9, 799–808. doi: 10.4161/rna.19718
- Amick, J., Rocznik-Ferguson, A., and Ferguson, S. M. (2016). C9orf72 binds SMCR8, localizes to lysosomes, and regulates mTORC1 signaling. *Mol. Biol. Cell* 27, 3040–3051. doi: 10.1091/mbc.E16-01-0003
- 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
- Balendra, R., and Isaacs, A. M. (2018). C9orf72-mediated ALS and FTD: Multiple pathways to disease. *Nat. Rev. Neurol.* 14, 544–558. doi: 10.1038/s41582-018-0047-2
- Banani, S. F., Lee, H. O., Hyman, A. A., and Rosen, M. K. (2017). Biomolecular condensates: Organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* 18, 285–298. doi: 10.1038/nrm.2017.7
- Batra, R., and Lee, C. W. (2017). Mouse models of C9orf72 hexanucleotide repeat expansion in amyotrophic lateral sclerosis/ frontotemporal dementia. *Front. Cell Neurosci.* 11:196. doi: 10.3389/fncel.2017.00196
- Belzil, V. V., Bauer, P. O., Prudencio, M., Gendron, T. F., Stetler, C. T., Yan, I. K., et al. (2013). Reduced C9orf72 gene expression in c9FTD/ALS is caused by histone trimethylation, an epigenetic event detectable in blood. *Acta Neuropathol.* 126, 895–905. doi: 10.1007/s00401-013-1199-1
- Bozic, T., Zalar, M., Rogelj, B., Plavec, J., and Sket, P. (2020). Structural diversity of sense and antisense RNA hexanucleotide repeats associated with ALS and FTL. *Molecules* 25:525. doi: 10.3390/molecules25030525
- Brcic, J., and Plavec, J. (2015). Solution structure of a DNA quadruplex containing ALS and FTD related GGGGCC repeat stabilized by 8-bromodeoxyguanosine substitution. *Nucleic Acids Res.* 43, 8590–8600. doi: 10.1093/nar/gkv815
- Brcic, J., and Plavec, J. (2018). NMR structure of a G-quadruplex formed by four d(G4C2) repeats: Insights into structural polymorphism. *Nucleic Acids Res.* 46, 11605–11617. doi: 10.1093/nar/gky886
- Burberry, A., Suzuki, N., Wang, J. Y., Moccia, R., Mordes, D. A., Stewart, M. H., et al. (2016). Loss-of-function mutations in the C9ORF72 mouse ortholog cause fatal autoimmune disease. *Sci. Transl. Med.* 8:347ra393. doi: 10.1126/scitranslmed.aaf6038
- Chew, J., Gendron, T. F., Prudencio, M., Sasaguri, H., Zhang, Y. J., Castanedes-Casey, M., et al. (2015). Neurodegeneration. C9ORF72 repeat expansions in mice cause TDP-43 pathology, neuronal loss, and behavioral deficits. *Science* 348, 1151–1154. doi: 10.1126/science.aaa9344
- 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(Pt 7), 2040–2051. doi: 10.1093/brain/awu120
- 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
- Dong, W., Zhang, L., Sun, C., Gao, X., Guan, F., Li, J., et al. (2020). Knock in of a hexanucleotide repeat expansion in the C9orf72 gene induces ALS in rats. *Animal Model. Exp. Med.* 3, 237–244. doi: 10.1002/ame2.12129
- Donnelly, C. J., Zhang, P. W., Pham, J. T., Haeusler, 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
- Edbauer, D., and Haass, C. (2016). An amyloid-like cascade hypothesis for C9orf72 ALS/FTD. *Curr. Opin. Neurobiol.* 36, 99–106. doi: 10.1016/j.conb.2015.10.009
- Fay, M. M., Anderson, P. J., and Ivanov, P. (2017). ALS/FTD-associated C9ORF72 repeat RNA promotes phase transitions in vitro and in cells. *Cell Rep.* 21, 3573–3584. doi: 10.1016/j.celrep.2017.11.093
- Fratta, P., Mizielinska, S., Nicoll, A. J., Zloh, M., Fisher, E. M., Parkinson, G., et al. (2012). C9orf72 hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes. *Sci. Rep.* 2:1016. doi: 10.1038/srep01016
- 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

Conflict of interest

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- 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
- Geng, Y., Liu, C., Cai, Q., Luo, Z., Miao, H., Shi, X., et al. (2021). Crystal structure of parallel G-quadruplex formed by the two-repeat ALS- and FTD-related GGGGCC sequence. *Nucleic Acids Res.* 49, 5881–5890. doi: 10.1093/nar/gkab302
- Gitler, A. D., and Tsuiji, H. (2016). There has been an awakening: Emerging mechanisms of C9orf72 mutations in FTD/ALS. *Brain Res.* 1647, 19–29. doi: 10.1016/j.brainres.2016.04.004
- Goodman, L. D., Prudencio, M., Kramer, N. J., Martinez-Ramirez, L. F., Srinivasan, A. R., Lan, M., et al. (2019). Toxic expanded GGGGCC repeat transcription is mediated by the PAF1 complex in C9orf72-associated FTD. *Nat. Neurosci.* 22, 863–874. doi: 10.1038/s41593-019-0396-1
- Graff-Radford, N. R., and Woodruff, B. K. (2007). Frontotemporal dementia. *Semin. Neurol.* 27, 48–57. doi: 10.1055/s-2006-956755
- Haeusler, A. R., Donnelly, C. J., and Rothstein, J. D. (2016). The expanding biology of the C9orf72 nucleotide repeat expansion in neurodegenerative disease. *Nat. Rev. Neurosci.* 17, 383–395. doi: 10.1038/nrn.2016.38
- Haeusler, A. R., Donnelly, C. J., Periz, G., Simko, E. A., Shaw, P. G., Kim, M. S., et al. (2014). C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature* 507, 195–200. doi: 10.1038/nature13124
- Halleger, M., Chakrabarti, A. M., Lee, F. C. Y., Lee, B. L., Amaliotti, A. G., Odeh, H. M., et al. (2021). TDP-43 condensation properties specify its RNA-binding and regulatory repertoire. *Cell* 168, 462–472. doi: 10.1016/j.cell.2021.07.018
- Hao, Z., Liu, L., Tao, Z., Wang, R., Ren, H., Sun, H., et al. (2019). Motor dysfunction and neurodegeneration in a C9orf72 mouse line expressing poly-PR. *Nat. Commun.* 10, 2906. doi: 10.1038/s41467-019-10956-w
- Herranz-Martin, S., Chandran, J., Lewis, K., Mulcahy, P., Higginbottom, A., Walker, C., et al. (2017). Viral delivery of C9orf72 hexanucleotide repeat expansions in mice leads to repeat-length-dependent neuropathology and behavioural deficits. *Dis. Model. Mech.* 10, 859–868. doi: 10.1242/dmm.029892
- Huppert, J. L. (2008). Four-stranded nucleic acids: Structure, function and targeting of G-quadruplexes. *Chem. Soc. Rev.* 37, 1375–1384. doi: 10.1039/b702491f
- Jain, A., and Vale, R. D. (2017). RNA phase transitions in repeat expansion disorders. *Nature* 546, 243–247. doi: 10.1038/nature22386
- Jambeau, M., Meyer, K. D., Hruska-Plochan, M., Tabet, R., Lee, C. Z., Ray-Soni, A., et al. (2022). Comprehensive evaluation of human-derived anti-poly-GA antibodies in cellular and animal models of C9orf72 disease. *Proc. Natl. Acad. Sci. U.S.A.* 119, e2123487119. doi: 10.1073/pnas.2123487119
- 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
- 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
- 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
- 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, 774–788. doi: 10.1016/j.cell.2016.10.002
- Lillo, P., and Hodges, J. R. (2009). Frontotemporal dementia and motor neuron disease: Overlapping clinic-pathological disorders. *J. Clin. Neurosci.* 16, 1131–1135. doi: 10.1016/j.jocn.2009.03.005
- 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, 789–802. doi: 10.1016/j.cell.2016.10.003
- Liu, Y., Pattamatta, A., Zu, T., Reid, T., Bardhi, O., Borchelt, D. R., et al. (2016). C9orf72 BAC mouse model with motor deficits and neurodegenerative features of ALS/FTD. *Neuron* 90, 521–534. doi: 10.1016/j.neuron.2016.04.005
- Liu, Y., Xing, H., Ernst, A. F., Liu, C., Maugee, C., Yokoi, F., et al. (2022). Hyperactivity of Purkinje cell and motor deficits in C9orf72 knockout mice. *Mol. Cell Neurosci.* 121, 103756. doi: 10.1016/j.mcn.2022.103756
- Lopez-Gonzalez, R., Lu, Y., Gendron, T. F., Karydas, A., Tran, H., Yang, D., et al. (2016). 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. doi: 10.1016/j.neuron.2016.09.015
- Lu, S., Hu, J., Arogundade, O. A., Goginashvili, A., Vazquez-Sanchez, S., Diedrich, J. K., et al. (2022). Heat-shock chaperone HSPB1 regulates cytoplasmic TDP-43 phase separation and liquid-to-gel transition. *Nat. Cell Biol.* 24, 1378–1393. doi: 10.1038/s41556-022-00988-8
- Mackenzie, I. R., Arzberger, T., Kremmer, E., Troost, D., Lorenzl, S., Mori, K., et al. (2013). Dipeptide repeat protein pathology in C9ORF72 mutation cases: Clinicopathological correlations. *Acta Neuropathol.* 126, 859–879. doi: 10.1007/s00401-013-1181-y
- Maharana, S., Wang, J., Papadopoulos, D. K., Richter, D., Pozniakovskiy, A., Poser, I., et al. (2018). RNA buffers the phase separation behavior of prion-like RNA binding proteins. *Science* 360, 918–921. doi: 10.1126/science.aar7366
- Maity, A., Winnerdy, F. R., Chen, G., and Phan, A. T. (2021). Duplexes formed by G(4)C(2) repeats contain alternate slow- and fast-flipping G.G base Pairs. *Biochemistry* 60, 1097–1107. doi: 10.1021/acs.biochem.0c00916
- Majounie, E., Renton, A. E., Mok, K., Dopper, E. G., Waite, A., Rollinson, S., et al. (2012). 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. doi: 10.1016/S1474-4422(12)70043-1
- Majumder, V., Gregory, J. M., Barria, M. A., Green, A., and Pal, S. (2018). TDP-43 as a potential biomarker for amyotrophic lateral sclerosis: A systematic review and meta-analysis. *BMC Neurol.* 18, 90. doi: 10.1186/s12883-018-1091-7
- Meijboom, K. E., Abdallah, A., Fordham, N. P., Nagase, H., Rodriguez, T., Kraus, C., et al. (2022). CRISPR/Cas9-mediated excision of ALS/FTD-causing hexanucleotide repeat expansion in C9ORF72 rescues major disease mechanisms in vivo and in vitro. *Nat. Commun.* 13, 6286. doi: 10.1038/s41467-022-33332-7
- Mordes, D. A., Morrison, B. M., Ament, X. H., Cantrell, C., Mok, J., Eggen, P., et al. (2020). Absence of survival and motor deficits in 500 repeat C9ORF72 BAC mice. *Neuron* 108, 775–783. doi: 10.1016/j.neuron.2020.08.009
- Mori, K., Arzberger, T., Grasser, F. A., Gijssels, I., May, S., Rentzsch, K., et al. (2013a). Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. *Acta Neuropathol.* 126, 881–893. doi: 10.1007/s00401-013-1189-3
- Mori, K., Lammich, S., Mackenzie, I. R., Forne, I., Zilow, S., Kretzschmar, H., et al. (2013b). 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. (2013c). The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTD/ALS. *Science* 339, 1335–1338. doi: 10.1126/science.1232927
- 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
- Nguyen, L., Laboissonniere, L. A., Guo, S., Pilotto, F., Scheidegger, O., Oestmann, A., et al. (2020a). Survival and motor phenotypes in FVB C9-500 ALS/FTD BAC transgenic mice reproduced by multiple labs. *Neuron* 108, 784–796. doi: 10.1016/j.neuron.2020.09.009
- Nguyen, L., Montrasio, F., Pattamatta, A., Tusi, S. K., Bardhi, O., Meyer, K. D., et al. (2020b). Antibody therapy targeting RAN proteins rescues C9 ALS/FTD phenotypes in C9orf72 mouse model. *Neuron* 105, 645–662. doi: 10.1016/j.neuron.2019.11.007
- Norpel, J., Cavadini, S., Schenk, A. D., Graff-Meyer, A., Hess, D., Seebacher, J., et al. (2021). Structure of the human C9orf72-SMCR8 complex reveals a multivalent protein interaction architecture. *PLoS Biol.* 19, e3001344. doi: 10.1371/journal.pbio.3001344
- O'Rourke, J. G., Bogdanik, L., Yanez, 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
- Peters, O. M., Cabrera, G. T., Tran, H., Gendron, T. F., McKeon, J. E., Metterville, J., et al. (2015). Human C9ORF72 hexanucleotide expansion reproduces RNA foci and dipeptide repeat proteins but not neurodegeneration in BAC transgenic mice. *Neuron* 88, 902–909. doi: 10.1016/j.neuron.2015.11.018
- Prpar Mihevc, S., Darovic, S., Kovanda, A., Bajc Cesnik, A., Zupunski, V., and Rogelj, B. (2017). Nuclear trafficking in amyotrophic lateral sclerosis and frontotemporal lobar degeneration. *Brain* 140, 13–26. doi: 10.1093/brain/aww197
- Rademakers, R., Neumann, M., and Mackenzie, I. R. (2012). Advances in understanding the molecular basis of frontotemporal dementia. *Nat. Rev. Neurosci.* 8, 423–434. doi: 10.1038/nrn.2012.117
- Ramaswami, M., Taylor, J. P., and Parker, R. (2013). Altered ribostasis: RNA-protein granules in degenerative disorders. *Cell* 154, 727–736. doi: 10.1016/j.cell.2013.07.038
- Reddy, K., Zamiri, B., Stanley, S. Y. R., 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 multimolecular 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
- Riemslogh, F. W., van der Toorn, E. C., Verhagen, R. F. M., Maas, A., Bosman, L. W. J., Hukema, R. K., et al. (2021). Inducible expression of human C9ORF72 36x

- G(4)C(2) hexanucleotide repeats is sufficient to cause RAN translation and rapid muscular atrophy in mice. *Dis. Model. Mech.* 14:dmm044842. doi: 10.1242/dmm.044842
- Roden, C., and Gladfelter, A. S. (2021). RNA contributions to the form and function of biomolecular condensates. *Nat. Rev. Mol. Cell Biol.* 22, 183–195. doi: 10.1038/s41580-020-0264-6
- Rowland, L. P., and Schneider, N. A. (2001). Amyotrophic lateral sclerosis. *N. Engl. J. Med.* 344, 1688–1700. doi: 10.1056/NEJM200105313442207
- Sannohe, Y., and Sugiyama, H. (2010). Overview of formation of G-quadruplex structures. *Curr. Protoc. Nucleic Acid Chem.* 12, 11–17. doi: 10.1002/0471142700.nc1702s40
- Sarma, R. H., Lee, C. H., Evans, F. E., Yathindra, N., and Sundaralingam, M. (1974). Probing the interrelation between the glycosyl torsion, sugar pucker, and the backbone conformation in C(8) substituted adenine nucleotides by 1H and 1H-(31P) fast Fourier transform nuclear magnetic resonance methods and conformational energy calculations. *J. Am. Chem. Soc.* 96, 7337–7348. doi: 10.1021/ja00830a028
- Schludi, M. H., Becker, L., Garrett, L., Gendron, T. F., Zhou, Q., Schreiber, F., et al. (2017). Spinal poly-GA inclusions in a C9orf72 mouse model trigger motor deficits and inflammation without neuron loss. *Acta Neuropathol.* 134, 241–254. doi: 10.1007/s00401-017-1711-0
- Selvaraj, B. T., Livesey, M. R., Zhao, C., Gregory, J. M., James, O. T., Cleary, E. M., et al. (2018). C9ORF72 repeat expansion causes vulnerability of motor neurons to Ca(2+)-permeable AMPA receptor-mediated excitotoxicity. *Nat. Commun.* 9:347. doi: 10.1038/s41467-017-02729-0
- Sen, D., and Gilbert, W. (1988). Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature* 334, 364–366. doi: 10.1038/334364a0
- Shen, Z., Jia, B., Xu, Y., Wessen, J., Pal, T., Chan, H. S., et al. (2023). Biological condensates form percolated networks with molecular motion properties distinctly different from dilute solutions. *eLife* 12:e81907. doi: 10.7554/eLife.81907
- Shi, Y., Lin, S., Staats, K. A., Li, Y., Chang, W. H., Hung, S. T., et al. (2018). Haploinsufficiency leads to neurodegeneration in C9ORF72 ALS/FTD human induced motor neurons. *Nat. Med.* 24, 313–325. doi: 10.1038/nm.4490
- Shin, Y., and Brangwynne, C. P. (2017). Liquid phase condensation in cell physiology and disease. *Science* 357:eaa4382. doi: 10.1126/science.aaf4382
- Sivadasan, R., Hornburg, D., Drepper, C., Frank, N., Jablonka, S., Hansel, A., et al. (2016). C9ORF72 interaction with cofilin modulates actin dynamics in motor neurons. *Nat. Neurosci.* 19, 1610–1618. doi: 10.1038/nn.4407
- Sket, P., Pohleven, J., Kovanda, A., Stalekar, M., Zupunski, V., Zalar, M., et al. (2015). Characterization of DNA G-quadruplex species forming from C9ORF72 G4C2-expanded repeats associated with amyotrophic lateral sclerosis and frontotemporal lobar degeneration. *Neurobiol. Aging* 36, 1091–1096. doi: 10.1016/j.neurobiolaging.2014.09.012
- Su, M. Y., Fromm, S. A., Remis, J., Toso, D. B., and Hurley, J. H. (2021). Structural basis for the ARF GAP activity and specificity of the C9orf72 complex. *Nat. Commun.* 12:3786. doi: 10.1038/s41467-021-24081-0
- Su, M. Y., Fromm, S. A., Zoncu, R., and Hurley, J. H. (2020). Structure of the C9orf72 ARF GAP complex that is haploinsufficient in ALS and FTD. *Nature* 585, 251–255. doi: 10.1038/s41586-020-2633-x
- Su, Z., Zhang, Y., Gendron, T. F., Bauer, P. O., Chew, J., Yang, W. Y., et al. (2014). Discovery of a biomarker and lead small molecules to target r(GGGGCC)-associated defects in c9FTD/ALS. *Neuron* 83, 1043–1050. doi: 10.1016/j.neuron.2014.07.041
- Sudria-Lopez, E., Koppers, M., de Wit, M., van der Meer, C., Westeneng, H. J., Zundel, C. A., et al. (2016). Full ablation of C9orf72 in mice causes immune system-related pathology and neoplastic events but no motor neuron defects. *Acta Neuropathol.* 132, 145–147. doi: 10.1007/s00401-016-1581-x
- 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
- Tang, D., Sheng, J., Xu, L., Zhan, X., Liu, J., Jiang, H., et al. (2020). Cryo-EM structure of C9ORF72-SMCR8-WDR41 reveals the role as a GAP for Rab8a and Rab11a. *Proc. Natl. Acad. Sci. U.S.A.* 117, 9876–9883. doi: 10.1073/pnas.2002110117
- Ugolino, J., Ji, Y. J., Conchina, K., Chu, J., Nirujogi, R. S., Pandey, A., et al. (2016). Loss of C9orf72 enhances autophagic activity via deregulated mTOR and TFE3 signaling. *PLoS Genet.* 12:e1006443. doi: 10.1371/journal.pgen.1006443
- Verdone, B. M., Cicardi, M. E., Wen, X., Sriramoji, S., Russell, K., Markandaiah, S. S., et al. (2022). A mouse model with widespread expression of the C9orf72-linked glycine-arginine dipeptide displays non-lethal ALS/FTD-like phenotypes. *Sci. Rep.* 12:5644. doi: 10.1038/s41598-022-09593-z
- Wu, X., Cai, Q., Feng, Z., and Zhang, M. (2020). Liquid-liquid phase separation in neuronal development and synaptic signaling. *Dev. Cell* 55, 18–29. doi: 10.1016/j.devcel.2020.06.012
- Xi, Z., Rainero, I., Rubino, E., Pinessi, L., Bruni, A. C., Maletta, R. G., et al. (2014). Hypermethylation of the CpG-island near the C9orf72 G(4)C(2)-repeat expansion in FTLD patients. *Hum. Mol. Genet.* 23, 5630–5637. doi: 10.1093/hmg/ddu279
- 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
- Yamakawa, M., Ito, D., Honda, T., Kubo, K., Noda, M., Nakajima, K., et al. (2015). Characterization of the dipeptide repeat protein in the molecular pathogenesis of c9FTD/ALS. *Hum. Mol. Genet.* 24, 1630–1645. doi: 10.1093/hmg/ddu576
- Yang, M., Liang, C., Swaminathan, K., Herrlinger, S., Lai, F., Shiekhhattar, R., et al. (2016). A C9ORF72/SMCR8-containing complex regulates ULK1 and plays a dual role in autophagy. *Sci. Adv.* 2:e1601167. doi: 10.1126/sciadv.1601167
- Yu, C. H., Teulade-Fichou, M. P., and Olsthoorn, R. C. (2014). Stimulation of ribosomal frameshifting by RNA G-quadruplex structures. *Nucleic Acids Res.* 42, 1887–1892. doi: 10.1093/nar/gkt1022
- Yu, H., Lu, S., Gasior, K., Singh, D., Vazquez-Sanchez, S., Tapia, O., et al. (2021). HSP70 chaperones RNA-free TDP-43 into anisotropic intranuclear liquid spherical shells. *Science* 37:eabb4309. doi: 10.1126/science.abb4309
- Zamiri, B., Mirceta, M., Bomsztyk, K., Macgregor, R. B. Jr., and Pearson, C. E. (2015). Quadruplex formation by both G-rich and C-rich DNA strands of the C9orf72 (GGGGCC)⁸(GGCCCC)⁸ repeat: Effect of CpG methylation. *Nucleic Acids Res.* 43, 10055–10064. doi: 10.1093/nar/gkv1008
- Zamiri, B., Reddy, K., Macgregor, R. B. Jr., and Pearson, C. E. (2014). 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. doi: 10.1074/jbc.C113.502336
- Zbinden, A., Perez-Berlanga, M., De Rossi, P., and Polymenidou, M. (2020). Phase separation and neurodegenerative diseases: A disturbance in the force. *Dev. Cell* 55, 45–68. doi: 10.1016/j.devcel.2020.09.014
- 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.* 19, 668–677. doi: 10.1038/nn.4272
- 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
- Zhou, B., Geng, Y., Liu, C., Miao, H., Ren, Y., Xu, N., et al. (2018). Characterizations of distinct parallel and antiparallel G-quadruplexes formed by two-repeat ALS and FTD related GGGGCC sequence. *Sci. Rep.* 8:2366. doi: 10.1038/s41598-018-20852-w
- Zhou, B., Liu, C., Geng, Y., and Zhu, G. (2015). Topology of a G-quadruplex DNA formed by C9orf72 hexanucleotide repeats associated with ALS and FTD. *Sci. Rep.* 5:16673. doi: 10.1038/srep16673
- Zhou, Q., Lehmer, C., Michaelsen, M., Mori, K., Alterauge, D., Baumjohann, D., et al. (2017). Antibodies inhibit transmission and aggregation of C9orf72 poly-GA dipeptide repeat proteins. *EMBO Mol. Med.* 9, 687–702. doi: 10.15252/emmm.201607054
- Zhou, Q., Mareljic, N., Michaelsen, M., Parhizkar, S., Heindl, S., Nuscher, B., et al. (2020). Active poly-GA vaccination prevents microglia activation and motor deficits in a C9orf72 mouse model. *EMBO Mol. Med.* 12:e10919. doi: 10.15252/emmm.201910919
- Zu, T., Guo, S., Bardhi, O., Ryskamp, D. A., Li, J., Khoramian Tusi, S., et al. (2020). Metformin inhibits RAN translation through PKR pathway and mitigates disease in C9orf72 ALS/FTD mice. *Proc. Natl. Acad. Sci. U.S.A.* 117, 18591–18599. doi: 10.1073/pnas.2005748117
- Zu, T., Liu, Y., Banez-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