



Wiley Interdisciplinary Reviews

**Article Title:**

Partners in crime: proteins implicated in RNA repeat expansion diseases

**Article Category:**

- PERSPECTIVE
- ADVANCED REVIEW
  
- PRIMER
- FOCUS ARTICLE
  
- OVERVIEW
- SOFTWARE FOCUS

## Authors:

### First author full name

Anna Baud<sup>#</sup>

Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Uniwersytetu Poznanskiego 6, Poznan, Poland. anna.baud@amu.edu.pl ORCID iD 0000-0003-3710-5722

### Second author full name

Magdalena Derbis <sup>#</sup>

Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Uniwersytetu Poznanskiego 6, Poznan, Poland.

Current address: Department of Molecular Probes and Prodrugs, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, Poznan, Poland. mderbis@ibch.poznan.pl

### Third author full name

Katarzyna Tutak <sup>#</sup>

Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Uniwersytetu Poznanskiego 6, Poznan, Poland. katarzyna.tutak@amu.edu.pl ORCID iD 0000-0002-2801-9956

### 4<sup>th</sup> author full name

Krzysztof Sobczak<sup>\*</sup>

Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Uniwersytetu Poznanskiego 6, Poznan, Poland. ksobczak@amu.edu.pl , ORCID iD 0000-0001-8352-9812

<sup>\*</sup>Corresponding author

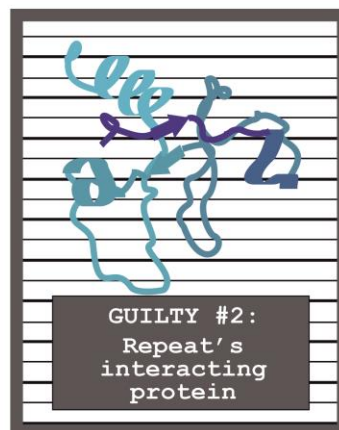
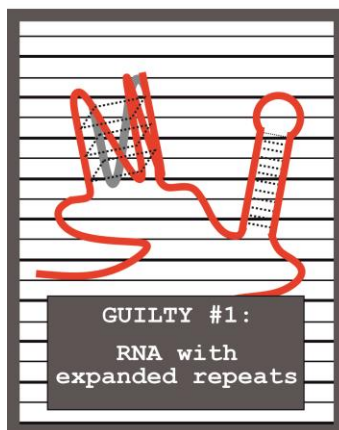
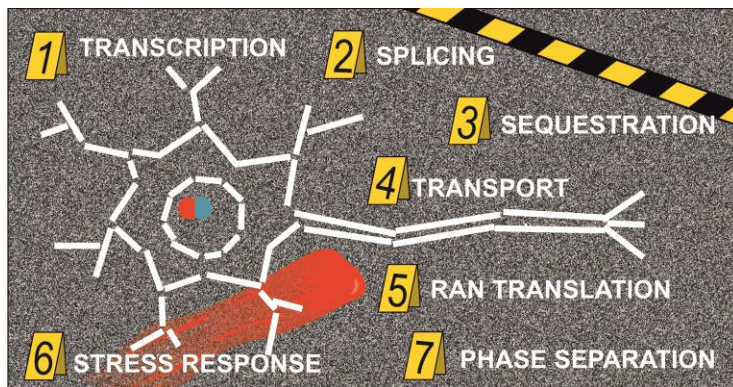
<sup>#</sup>Contributed equally

**Conflict of Interest:** All authors declare no conflict of interest

## Abstract

Short tandem repeats (STRs) are repetitive nucleotide sequences robustly distributed in the human genome. Their expansion underlies the pathogenesis of multiple neurological disorders, including Huntington's disease (HD), amyotrophic lateral sclerosis and frontotemporal dementia (ALS/FTD), fragile X-associated tremor/ataxia syndrome (FXTAS) and myotonic dystrophies (DMs), known as repeat expansion disorders (REDs). Several molecular pathomechanisms associated with toxic RNA containing expanded repeats (RNA<sup>exp</sup>) are shared among REDs and contribute to disease progression, however, detailed mechanistic insight into those processes is limited. To deepen our understanding of the interplay between toxic RNA<sup>exp</sup> molecules and multiple protein partners, in this review, we discuss the roles of selected RNA-binding proteins (RBPs) that interact with RNA<sup>exp</sup> and thus act as "partners in crime" in the progression of REDs. We gather current findings concerning RBPs involved at different stages of the RNA<sup>exp</sup> life cycle, such as transcription, splicing, transport and AUG-independent translation of expanded repeats. We argue that the activity of selected RBPs can be unique or common among REDs depending on the expanded repeat type. We also present proteins that are functionally depleted due to sequestration on RNA<sup>exp</sup> within nuclear foci and those which participate in RNA<sup>exp</sup>-dependent innate immunity activation. Moreover, we discuss the utility of selected RBPs as targets in the development of therapeutic strategies.

## Graphical/Visual Abstract and Caption



## 1. INTRODUCTION

Short tandem repeats (STRs), also termed microsatellites, are repeats of 1-8 nucleotide-long sequences and occur very frequently within different parts of genomes. Some STRs are overrepresented in exons, and depending on the sequence, their frequencies in open reading frames and 5' and 3' untranslated regions (UTRs) differ (Kozlowski et al., 2010). Moreover, trinucleotide repeats are more frequent in the coding parts of genes, whereas 5- and 6-nucleotide repeats are more frequent in noncoding parts (Malik, Kelley, et al., 2021). STRs are polymorphic in length and prone to expansion or, to a lesser extent, contraction either in cells of the same individual (somatic instability) or between generations (germline instability) (Ashizawa et al., 1992; Paulson, 2018; Trang et al., 2015; Trottier et al., 1994). The origin of STR instability remains incompletely understood, but the possible mechanisms include replication slippage (McMurray, 2010), replication fork stalling

(Gadgil et al., 2017), mismatch repair and bidirectional transcription (Castel et al., 2010).

To date, several possible models of pathogenesis caused by STR expansions have been proposed, and in different repeat expansion disorders (REDs), combinations of these pathomechanisms play crucial roles in disease development (extensively reviewed in (Malik, Kelley, et al., 2021)). At the DNA level, STRs in noncoding parts of genes may enhance the cotranscriptional formation of RNA/DNA hybrids called R-loops, which are associated with induction of the DNA damage response and gene silencing due to methylation or blockage of transcription (Reddy et al., 2011). At the RNA level, a dominant gain-of-function mechanism in which RNA bearing expanded repeats (RNA<sup>exp</sup>) sequesters RNA-binding proteins (RBPs), mostly in cell nuclei, impairing their function, has been postulated in many diseases (i.e. A. Mankodi et al., 2000; Miller et al., 2000; Sellier et al., 2010, 2013). Moreover, if RNA<sup>exp</sup> is transported to the cytoplasm, its toxicity can be exerted by the translation of the repeats that do not reside in the canonical AUG-initiated open reading frame *via* process called repeat-associated non-AUG (RAN) translation (Zu et al., 2011). Finally, a gain-of-function mechanism involving the protein products of genes bearing STRs in their coding sequences is associated with their tendency to form toxic aggregates or activate stress response pathways (Cook et al., 2020; Y. J. Zhang et al., 2018). In all of the mentioned pathomechanisms, processes involving RNA<sup>exp</sup> molecules can be considered a causative agent from transcription to translation.

**Sidebar title: Gene/RNA/protein gain- and loss-of-function in repeat expansion diseases (REDs)**

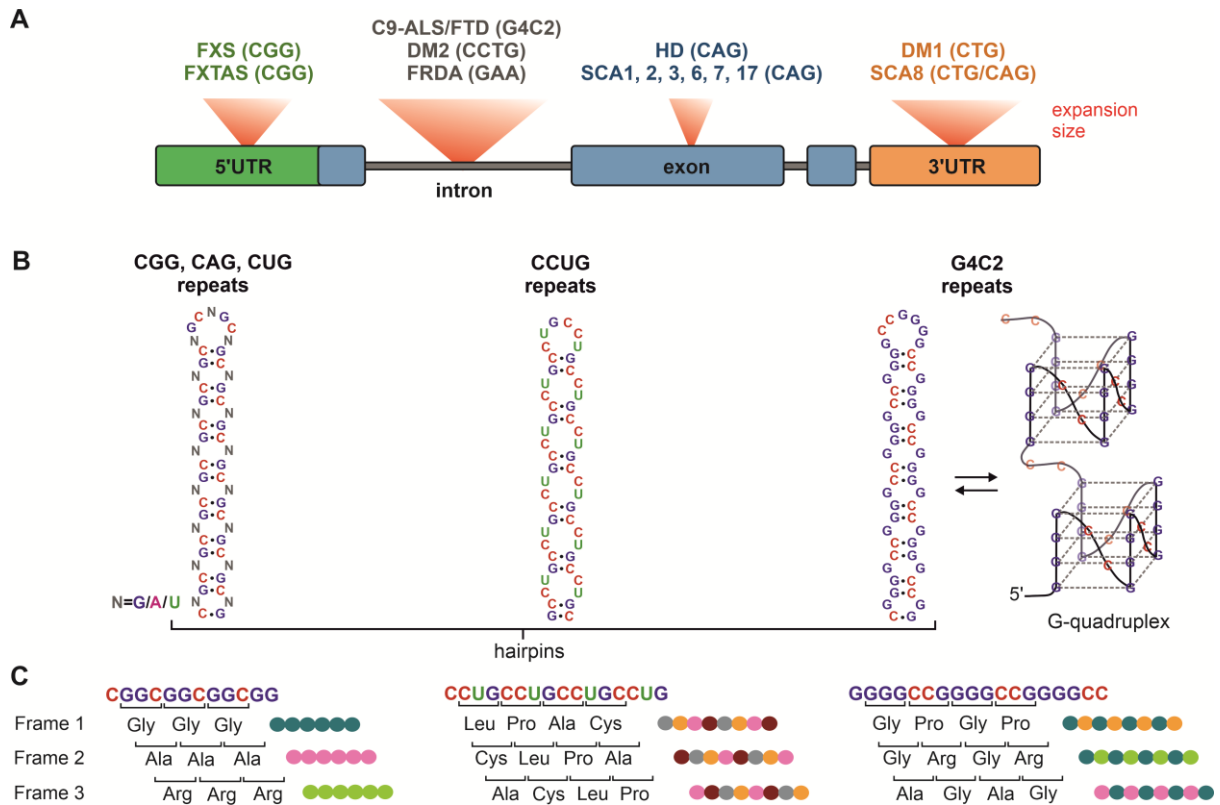
**Gene loss-of-function** repeats expansion in DNA might result in partial or complete lack of gene product. Altered gene expression can be caused by transcriptional gene silencing mediated by mutated allele hypermethylation or impaired transcription. For example, in fragile X-syndrome, heterochromatinization of *FMR1* locus silences expression of Fragile X mental retardation protein 1 (FMRP). Alternatively, in Friedreich's ataxia actively transcribed nascent RNA with GAA repeats can form DNA-RNA R-loop structures which initiates epigenetic silencing of *FXN* gene. **RNA gain-of-function** – expanded repeats within (pre-)mRNAs form stable secondary structures, accumulate in the cell nuclei, attract and trap RNA binding proteins and form ribonucleoprotein complexes called RNA foci. RNA foci are pathological hallmarks of many REDs. For example, in myotonic dystrophy type 1 (DM1),

muscleblind-like proteins (MBNLs) are sequestered to CUG repeat foci which leads to the global alternative splicing abnormalities due to MBNLs functional depletion. **Protein gain-of-function** presence of expanded repeats in RNA, if transported to cytoplasm, may result in production of mutated variant of the protein containing repeated amino acids, like in case of polyQ expansion in huntingtin. Moreover, peptides derived from RNA<sup>exp</sup> can aggregate and represent toxic properties. Insoluble protein aggregates are detected in many REDs, e.g. polyglycine in FXTAS. On the other hand, in C9-ALS/FTD intron retention can result in decreased synthesis of C9orf72 protein, which exemplifies **protein loss-of-function** mechanism due to repeat expansion.

However, it takes two to tango: in the crime of RED pathogenesis, the protein-RNA<sup>exp</sup> interaction cannot be neglected. RBPs bind specific RNA sequences and/or its specific secondary/tertiary structures through RNA-binding domains (RBDs), form ribonucleoprotein (RNP) complexes and engage in fundamental cell processes, i.e., the regulation of transcription, mRNA splicing, maturation, transport, stability, cellular localization and translation (reviewed in (Jazurek et al., 2016)). The most common RBDs contain either an RNA-recognition motif (RRM), a heterogeneous nuclear RNP (hnRNP) K homology domain (KH), zinc fingers (ZNFs), an S1 domain, a double-stranded RBD (dsRBD) or a combination of these domains (reviewed in (Lunde et al., 2007)). Eukaryotic cells contain hundreds of different RBPs with unique RNA-binding specificity, many of which have been described as involved in the development of REDs.

To date, over 50 human inherited REDs, mainly neurological or neuromuscular diseases, have been described (Figure 1A). Those typically multisystemic diseases can be inherited in an autosomal dominant manner, i.e., Huntington's disease (HD) (CAG<sup>exp</sup> in the protein-coding region) (MacDonald et al., 1993), myotonic dystrophies (DMs) (CTG<sup>exp</sup> in the 3'UTR) (Brook et al., 1992), and C9orf72-linked amyotrophic lateral sclerosis and frontotemporal dementia (C9-ALS/FTD) (G4C2<sup>exp</sup> in the intron) (DeJesus-Hernandez et al., 2011); an autosomal recessive manner, i.e., Friedreich's ataxia (FRDA) (GAA<sup>exp</sup> in the intron of the frataxin *FXN* gene) (Campuzano et al., 1996); or an X-linked dominant manner, i.e., fragile X syndrome (FXS) and fragile X-associated tremor/ataxia syndrome (FXTAS) (CGG<sup>exp</sup> in the 5'UTR of the fragile X mental retardation 1 *FMR1* gene (Tassone et al., 2000). In HD, CAG<sup>exp</sup> in the coding region of the huntingtin gene (*HTT*) results in the production of a mutant protein

containing a polyglutamine (polyQ) stretch, which is susceptible to misfolding and aggregation and thus underlies disease pathogenesis (Martí, 2016; Persichetti et al., 1995). In FXS, CGG<sup>exp</sup> exceeding 200 repeats, named full mutation, causes hypermethylation and silencing of the *FMR1* promoter (Tassone et al., 2000, 2007). The absence of the synaptic functional regulator FMR1 protein (FMRP) is linked to alterations in brain synaptic plasticity, impairing cognitive functions and resulting in intellectual disability. In FXTAS, the mutant RNA containing 55-200 CGG<sup>exp</sup> can sequester some RBPs (Sellier et al., 2010, 2013) or trigger the production of toxic polyglycine (polyG) protein (Todd et al., 2013). The pathological hallmark of FXTAS is the presence of large, ubiquitin-positive inclusions in the nuclei of neurons and astrocytes (Greco et al., 2002), while the clinical symptoms include intention tremor, cerebellar ataxia, parkinsonism and brain atrophy (Greco et al., 2006; Hagerman et al., 2001). In C9-ALS/FTD, G4C2<sup>exp</sup> in the first intron of the *C9orf72* gene leads to insufficiency of the product of the mutated gene, the C9orf72 protein (DeJesus-Hernandez et al., 2011); toxic gain-of-RNA function (Xu et al., 2013); and the production of toxic dipeptide repeat proteins (DPRs) (Ash et al., 2013; Gendron et al., 2013). Although DM1 is caused by CTG<sup>exp</sup> in the 3'UTR of the *DMPK* gene and DM2 by CCTG<sup>exp</sup> in intron 1 of the nucleic acid-binding protein-coding *CNBP* gene, these two neuromuscular disorders exhibit common pathological symptoms, i.e., skeletal muscle weakness, wasting and cognitive dysfunction (Brook et al., 1992; Liquori et al., 2001). Both CUG<sup>exp</sup> and CCUG<sup>exp</sup> are toxic and lead to nuclear sequestration of the Muscleblind-like (MBNL) proteins, impairing their physiological functions (Fardaei et al., 2001, 2002; Mankodi, 2001; Miller et al., 2000).



**Figure 1. Characteristics of expanded STRs specific for different diseases.**

**(A) Localization and size of STRs within specific gene regions.** Expanded STRs, depending on the sequence, are located in different parts of gene. Size of expansion of STRs necessary for development of individual REDs differ between diseases, however, it may be roughly specified that the longest expansions are related within introns and further within 3'UTRs, middle-size expansions within 5'UTRs and the shortest within exons, and thus with protein coding sequences. Here we show representative REDs from larger group of diseases.

**(B) Structures formed by RNA<sup>exp</sup>.** Trinucleotide CNG repeats form RNA hairpin structures, all characterized by high thermodynamic stability, the highest for CGG, next CAG and the lowest for CUG repeats. Hairpin structures are also formed by CCUG and G4C2 repeats. Moreover, G4C2 and CGG repeats are able to form G-quadruplex structure.

**(C) Protein products derived from repeat associated non-AUG (RAN) translation.** RAN translation may potentially start and produce proteins in all three reading frames. In the process of RAN translation of trinucleotide repeats the homopolymeric proteins with tracts of single amino acids are biosynthesized. CCUG tetranucleotide repeats are RAN translated to proteins containing tracts of four amino acids, the same in all reading frames. RAN translation of G4C2 hexanucleotide repeats is the source of DPRs composed of tracts of two amino acids repeats, the one of which is glycine in all reading frames.

In this review, we focus on roles of RNA<sup>exp</sup> and its protein partners in the pathogenesis of selected REDs, C9-ALS/FTD, FXTAS, HD, and DM, at different stages of the RNA<sup>exp</sup> life cycle (Table 1). Importantly, the pathomechanisms of these

diseases are shared among many other REDs. We discuss not only the proteins that play a role in nuclear processing, i.e., the transcription and splicing of RNA<sup>exp</sup>, and those that are sequestered by RNA<sup>exp</sup> in various disorders, but also proteins that play a role in RNA<sup>exp</sup> transport, non-canonical translation, stress response and phase separation of products of mutant genes. We also discuss these proteins in the context of the development of therapeutic strategies. The described proteins may be considered partners in crime that, together with RNA<sup>exp</sup> molecules, are responsible for the development and progression of REDs.

**Table 1.** Proteins implicated in pathomechanisms of repeat expansion diseases (REDs).

Process	RNA <sup>exp</sup>	Protein	Protein name	UniProt ID	Effect/Method	
Transcription	CAG	<b>SUPT4H1</b>	Transcription elongation factor SPT4	P63272 (human)	KO in mouse reduced production of mutated huntingtin	(C. R. Liu et al., 2012)
	CAG	<b>SPT4-A</b>	Transcription elongation factor SPT4-A	P63271 (mouse)	KD (ASO) in mouse selectively reduced mutated mRNA and protein	(H. M. Cheng et al., 2015)
	G4C2/ C4G2	<b>SUPT4H1</b>	Transcription elongation factor SPT4	P63272 (human)	KO in yeast and in fly reduced RNA foci and RAN translation	(Kramer et al., 2016)
	G4C2	PAF1, LEO1	RNA polymerase II-associated factor 1 homolog; RNA polymerase-associated protein LEO1	Q8N7H5 (human), Q8WVC0 (human)	KD in fly suppressed toxicity, upregulated in patients	(Goodman, Prudencio, Kramer, et al., 2019)
	G4C2	AFF2/FMR2	AF4/FMR2 family member 2	P51816 (human)	KD in fly reduced RNA <sup>exp</sup> and RAN protein level, KO in iPSNs reduced level of <i>C9orf72</i> RNA, RNA foci and RAN translation products	(Yuva-Aydemir et al., 2019)
	GGCCUG	<b>SPT4-A, SPT5</b>	Transcription elongation factor SPT4-A; Transcription elongation factor SPT5	P63271 (mouse), O55201 (mouse)	KD in Neuro2A cells reduced RNA foci and RAN translation	(Furuta et al., 2019)
Sequestration, co-localization	CAG	MBNL1	Muscleblind-like protein 1	Q9NR56 (human)	OE rescued splicing abnormalities in HeLa cells expressing CAG <sup>exp</sup>	(Mykowska et al., 2011)
	G4C2	Nucleolin	Nucleolin	P19338 (human)	nucleolar stress in C9-ALS/FTD patients	(Haeusler et al., 2014)
	G4C2	Pur alpha	Transcriptional activator protein Pur-alpha	Q00577 (human)	OE reduced neurodegeneration in fly, cell and fish models.	(Swinnen et al., 2018; Xu et al., 2013)
	G4C2	hnRNP H	Heterogeneous nuclear ribonucleoprotein H	P31943 (human)	colocalized with RNA foci in C9-ALS/FTD patients, missplicing of hnRNP H-dependent transcript	(Y. B. Lee et al., 2013)
	G4C2	hnRNP H1/F, ALYREF and SRSF2	Heterogeneous nuclear ribonucleoprotein F, THO complex subunit 4, SRSF protein kinase 2	P52597 (human), Q86V81 (human), P78362 (human)	colocalized with RNA foci in C9-ALS/FTD patients	(Cooper-Knock et al., 2014)
	C4G2	SRSF1, hnRNP A1, hnRNP H/F, ALYREF	SRSF protein kinase 1, Heterogeneous nuclear ribonucleoprotein A1, Heterogeneous nuclear ribonucleoprotein F, THO complex subunit 4	O70551 (human), P09651 (human), P52597 (human), Q86V81 (human)	colocalized with RNA foci in C9-ALS/FTD patients	(Cooper-Knock et al., 2015)
	G4C2	ADARB2	Double-stranded RNA-specific editase B2	Q9NS39 (human)	KD decreased RNA foci in iPSNs, colocalized with RNA foci in C9-ALS/FTD patients	(Donnelly et al., 2013)

	G4C2	ZPF106	Zinc finger protein 106	Q9H2Y7 (human)	OE suppressed neurotoxicity in C9-ALS/FTD fly model	(Celona et al., 2017)
	G4C2	Matrin-3	Matrin-3	P43243 (human)	OE suppressed neurotoxicity in C9-ALS/FTD fly model	(Ramesh et al., 2020)
	CUG	MBNL1	Muscleblind-like protein 1	Q9NR56 (human)	OE/ASO blockers complementary to RNA <sup>exp</sup> reversed DM1 phenotype in mouse model	(Kanadia et al., 2006; Wheeler et al., 2009)
	CUG	hnRNP H	Heterogeneous nuclear ribonucleoprotein H	P31943 (human)	colocalized with RNA foci in DM1 patients	(D. H. Kim et al., 2005)
	CCUG	MBNL1	Muscleblind-like protein 1	Q9NR56 (human)	colocalized with RNA foci in DM2 patients	(Mankodi, 2001)
	CCUG	hnRNP core protein and snRNP Sm antigen	Small nuclear ribonucleoprotein E	P62304 (human)	colocalized with RNA foci in DM2 patients	(Perdoni et al., 2009)
	CCUG	RBFOX1	RNA binding protein fox-1 homolog	Q9NWB1 (human)	colocalized with RNA foci in DM2 patients	(Sellier et al., 2018)
	CGG	FMRpolyG/PolyG	FMRpolyG	-	colocalized with RNA <sup>exp</sup> , cause neurotoxicity in primary neurons derived from FXTAS mouse	(Asamitsu et al., 2021)
	CGG	hnRNP A2/B1	Heterogeneous nuclear ribonucleoproteins A2/B1	P22626 (human)	OE suppressed neurotoxicity in FXTAS fly model	(Sofola et al., 2007)
	CGG	CELF1	CUGBP Elav-like family member 1	Q92879 (human)	OE suppressed neurotoxicity in FXTAS fly model	(Sofola et al., 2007)
	CGG	Pur alpha	Transcriptional activator protein Pur-alpha	Q00577 (human)	OE suppressed neurotoxicity in FXTAS fly model	(Jin et al., 2007)
	CGG	TRA2A	Transformer-2 protein homolog alpha	Q13595 (human)	colocalized with RNA <sup>exp</sup> foci in FXTAS cellular model and with inclusions in FXTAS mouse model and patients	(Cid-Samper et al., 2018)
	CGG	DROSHA-DGCR8	Ribonuclease 3 and Microprocessor complex subunit DGCR8	Q9NRR4 (human), Q8WYQ5 (human)	OE suppressed neurotoxicity in FXTAS cellular model, miRNA processing is impaired in FXTAS patients	(Sellier et al., 2013)
	CGG	SAM68	KH domain-containing, RNA-binding, signal transduction-associated protein 1	Q07666 (human)	colocalized with RNA <sup>exp</sup> foci in FXTAS cellular model, SAM68-dependent alternative splicing is impaired in FXTAS patients	(Sellier et al., 2010)
<b>Nucleocytoplasmic transport</b>	CAG	U2AF65	Splicing factor U2AF 65 kDa subunit	P26368 (human)	KD in fly enhanced RNA <sup>exp</sup> toxicity and accumulation, reduction in HD mouse caused RNA <sup>exp</sup> nuclear accumulation	(Tsoi et al., 2011)
	G4C2	RanGAP1	Ran GTPase-activating protein 1	P46060 (human)	OE suppressed RNA <sup>exp</sup> toxicity in C9-ALS/FTD fly, OE rescued impairment of	(K. Zhang et al., 2015)

					NCT in C9-ALS iPSNs	
	G4C2/ CGG	SRSF1	Serine/arginine-rich splicing factor 1	O70551 (human)	KD reduced nuclear export of mutated G4C2 <sup>exp</sup> /CGG <sup>exp</sup> transcripts in cells and neurotoxicity in C9-ALS/FTD iPSNs and in FXTAS & C9-ALS/FTD fly models	(Hautbergue et al., 2017; Malik, Tseng, et al., 2021)
	G4C2	POM 121	Nuclear envelope pore membrane protein POM121	Q8TEM1(human)	OE rescued expression of nuclear pore components and NCT impairment in C9-ALS/FTD iPSNs	(Coyne et al., 2020)
	CUG	hnRNP H	Heterogeneous nuclear ribonucleoprotein H	P31943 (human)	KD reduced nuclear retention of RNA <sup>exp</sup> in cells	(D. H. Kim et al., 2005)
	CUG	Staufen1	Double-stranded RNA-binding protein Staufen homolog 1	O95793(human)	OE reduced nuclear retention of RNA <sup>exp</sup> in cells	(Ravel-Chapuis et al., 2012)
<b>RAN translation</b>	G4C2/ CGG	eIF4A	Eukaryotic initiation factor 4A-I	P60842 (human)	inhibition reduced RAN translation <i>in vitro</i>	(Green et al., 2017; Kearse et al., 2016)
	CGG/ G4C2	eIF4B and eIF4H	Eukaryotic translation initiation factor 4B and Eukaryotic translation initiation factor 4H	P23588 and Q15056 (human)	KD reduced RAN translation in C9-ALS/FTD fly model	(Goodman, Prudencio, Srinivasan, et al., 2019; Linsalata et al., 2019)
	CGG	DDX3X	ATP-dependent RNA helicase DDX3X	O00571(human)	KD reduced RAN translation in FXTAS fly model	(Linsalata et al., 2019)
	G4C2	DDX3X	ATP-dependent RNA helicase DDX3X	O00571(human)	KD enhanced RAN translation in C9-ALS/FTD fly model and in C9-ALS/FTD iPSNs	(W. Cheng et al., 2019)
	G4C2/CGG	DHX36	ATP-dependent DNA/RNA helicase DHX36	Q9H2U1(human)	KD reduced RAN translation in C9-ALS/FTD iPSNs	(H. Liu et al., 2021; Tseng et al., 2021)
	CGG	eIF1 and eIF5	Eukaryotic translation initiation factor 1 and Eukaryotic translation initiation factor 5	P41567 and P55010 (human)	modulated RAN translation <i>in vitro</i>	(Linsalata et al., 2019)
	CAG/G4C2	eIF3F	Eukaryotic translation initiation factor 3 subunit F	O00303 (human)	KD reduced RAN translation <i>in vitro</i>	(Ayhan et al., 2018)
	CGG	5MP	eIF5-mimic protein (also known as BZW2)	E9PFD4 (human)	OE reduced RAN translation in FXTAS fly model	(Singh et al., 2021)
	G4C2	eIF2D	Eukaryotic translation initiation factor 2D	P41214(human)	KD reduced RAN translation in <i>C. elegans</i> C9-ALS/FTD model	(Sonobe et al., 2021)
	G4C2/ CAG	RPS25	40S ribosomal protein S25	P62851 (human)	KD reduced RAN translation in C9-ALS/FTD fly model and in C9-ALS/FTD iPSNs	(Yamada et al., 2019)
<b>Stress response</b>	G4C2/ CAG/ CCTG/	PKR	Interferon-induced, double-stranded RNA-activated protein kinase	P19525 (human)	inhibition of PKR reduced RAN translation in C9-ALS/FTD mouse model	(Zu et al., 2020)

	CAGG					
	G4C2/ CGG	eIF2 $\alpha$	Eukaryotic translation initiation factor 2 subunit 1	P05198 (human)	enhanced RAN translation when phosphorylated in stress conditions <i>in vitro</i>	(W. Cheng et al., 2018; Green et al., 2017)
	G4C2/ CAGG/ CCTG	eIF2A	Eukaryotic translation initiation factor 2A	Q9BY44(human)	KO reduced RAN translation <i>in vitro</i>	(Sonobe et al., 2018; Tusi et al., 2021)
	G4C2	PERK	Eukaryotic translation initiation factor 2-alpha kinase 3	Q9NZJ5 (human)	Elevated activity in response to accumulation of RAN peptides increases RAN translation	(Zu et al., 2020)
	CUG/CAG	Dicer	Endoribonuclease Dicer	Q9UPY3 (human)	Cleaves RNA repeat regions into ~21nt fragments	(Krol et al., 2007; Yu et al., 2011; Lawlor et al., 2011)
	CAG/CUG	ADAR1	Double-stranded RNA-specific adenosine deaminase	P55265 (human)	Co-expression of ADAR1 with (CAG/CUG) <sub>100</sub> dsRNA rescued repeat-related pathology in <i>Drosophila</i>	(Van Eyk, 2019)
	CAG/CUG	TLR	Toll-like receptors		KD of TLRs in <i>Drosophila</i> decreased the toxicity of CAG/CUG <sub>-100</sub> repeats	(Samaraweera, 2013)
<b>MLO formation (Phase separation)</b>	CAG	SRSF2	Serine/arginine-rich splicing factor 2	Q01130 (human)	marker of nuclear speckles, colocalized with RNA <sup>exp</sup> foci	(Jain & Vale, 2017)
	G4C2	SRSF2	Serine/arginine-rich splicing factor 2	Q01130(human)	marker of nuclear speckles, colocalized with RNA <sup>exp</sup> foci	(Jain & Vale, 2017)
	G4C2	G3BP1, Caprin1, USP10, eIF3b, ELAVL1, TIAR	Ras GTPase-activating protein-binding protein 1, Caprin-1, Ubiquitin carboxyl-terminal hydrolase 10, Eukaryotic translation initiation factor 3 subunit B, ELAV-like protein 1, Nucleolysin TIAR	Q13283 (human), Q14444 (human), Q14694 (human), P55884 (human), Q15717 (human) Q01085 (human)	markers of stress granules, condensed <i>in vitro</i> with RNA <sup>exp</sup> and lysates from cell lines and mouse brain	(Fay et al., 2017)
	G4C2	FMRP	Synaptic functional regulator FMR1	Q06787 (human)	marker of transport granules, colocalized with RNA <sup>exp</sup> , FMRP-dependent translation regulation was impaired in C9-ALS/FTD iPSNs	(Burguete et al., 2015)

KD knockdown; KO knockout; OE overexpression; ASO antisense oligonucleotides; iPSNs induced pluripotent stem cells-derived neurons; NCT nucleocytoplasmic transport; MLO membraneless organelles.

## 2. TOXIC RNA<sup>EXP</sup> MOLECULES

In various REDs, distinct toxic RNA<sup>exp</sup> molecules are generated. Interestingly, for the development of each RED, a specific “repeat load” is needed. This “repeat load” not only consists of the length of STRs within RNA<sup>exp</sup> and cellular RNA<sup>exp</sup> levels but also the specific location, for example, certain cellular compartments and specific backgrounds within cell types. Thus, some RNA<sup>exp</sup> molecules are pathogenic when the length of inherited expanded repeats exceeds a few dozen of copies, while for others, when the number of repeats exceeds hundreds or even thousands of copies (Paulson, 2018). STR length is also related to pathomechanisms in which RNA<sup>exp</sup> molecules are involved; for example, RBP sequestration plays a larger role in REDs caused by very long STRs, such as DMs (up to 10,000 CTG or CCTG repeats) (Malik, Kelley, et al., 2021; Rohilla & Gagnon, 2017). The toxicity of RNA<sup>exp</sup> may further increase with the elongation of STRs during lifespan of patients, manifesting as increased disease severity or a younger age of onset (Paulson, 2018; Swami et al., 2009).

RNA<sup>exp</sup> molecules were shown to form different types of structures with different thermodynamic stabilities, depending on the repeated sequence motif: unstructured single-stranded RNAs (e.g., AAG repeats), semistable hairpins, fairly stable hairpins or very stable G-quadruplexes (Figure 1B). The most common RNA<sup>exp</sup> molecules in REDs, CNG repeats, follow the thermodynamic stability order CGG>CAG>CUG>CCG (Sobczak et al., 2010). At physiological KCl concentration, pH and temperature, G4C2<sup>exp</sup> and CGG<sup>exp</sup> form a stable parallel G-quadruplex (Asamitsu et al., 2021; Haeusler et al., 2014). These higher-order structures show extremely high thermodynamic stability and are directly linked to the pathogenic potential of RNA<sup>exp</sup>, as they are recognized by different RBPs. Importantly, the structure of RNA<sup>exp</sup> may be influenced by the presence of nucleotide interruptions in the sequences of repeats. In general, these protective elements against STR instability at the genomic level reduce RNA<sup>exp</sup> toxicity. Interruptions including CGG, CTC, GGC or CAG triplets in CTG<sup>exp</sup> occur in ~3-5% of DM1 patients, reduce somatic instability and may result in atypical or milder symptoms, later ages of onset and progression (Brida et al., 2010; Cumming et al., 2018; Wenninger et al., 2021).

Similarly, AGG interruptions within CGG<sup>exp</sup> in *FMR1*, increase genetic stability of alleles with 45-90 repeats (Nolin et al., 2015; Villate et al., 2020; Yrigollen et al., 2014). The CAA interruptions in CAG<sup>exp</sup> in *SCA2* are linked with delayed disease onset. On the other hand numerous CCG/CGG interruptions in CTG/CAG<sup>exp</sup> of *ATXN8* related to Spinocerebellar ataxia type 8 (*SCA8*), increase thermodynamic stability of RNA hairpin, , change amino acid composition of RAN proteins enhancing their toxicity and thus are associated with increased disease penetrance (Perez et al., 2021).

Despite differences concerning their sequence, the length of their STRs, structure, etc., various RNA<sup>exp</sup> molecules are engaged in similar pathogenic mechanisms, such as RBP sequestration, impairment of nucleocytoplasmic transport (NCT) or RAN translation. This is related to the natural multivalency of RNA<sup>exp</sup> molecules, which makes them susceptible to interaction with each other and with multiple RBPs. It further leads to the formation of RNA foci, structures composed of RNA<sup>exp</sup> and proteins, mainly in the nucleus, that vary in size, frequency and composition but are common between different REDs (reviewed in (N. Zhang & Ashizawa, 2017)).

#### **Sidebar title: Repeat load**

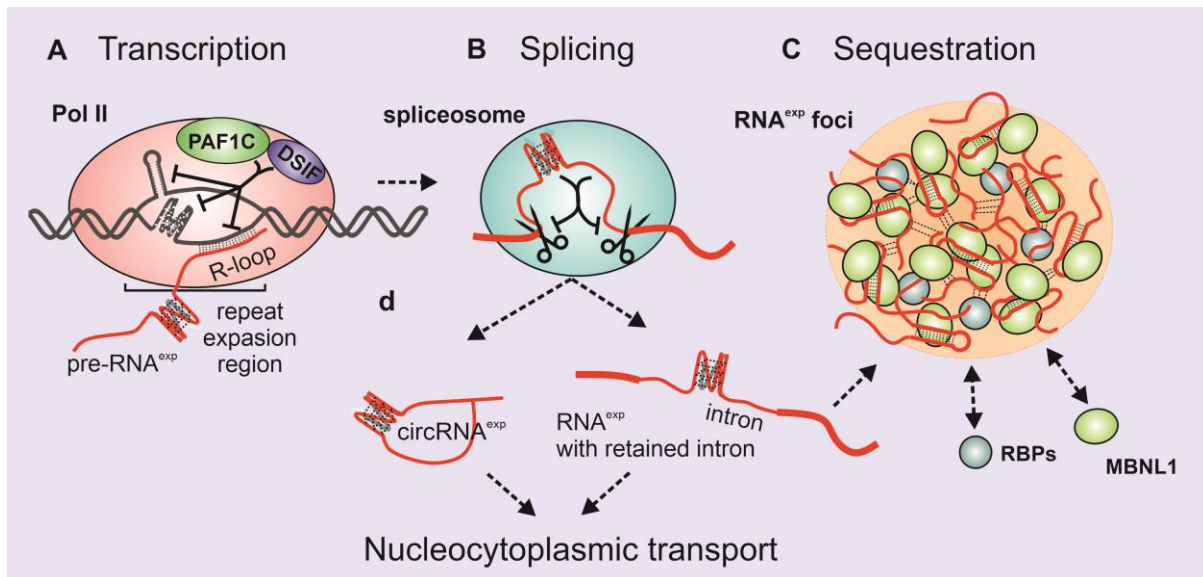
Specific repeat load, necessary for development of individual RED, consists of a few components. First, the length of expanded repeats, which may further increase between generations or in different cells within the same organism, due to germline and more robust somatic instability. Second, the biogenesis and cellular stability of expanded repeat-bearing transcripts. Third, localization of RNA<sup>exp</sup> in a defined cell type and cellular compartment where it interacts with specific partners. Finally, the sequence itself, which, if enriched in nucleotide interruptions in the repeat tract, change the potential of RNA<sup>exp</sup> molecule for relevant interactions with RBPs or toxicity of RAN protein products related to their amino acid composition.

### **3. TRANSCRIPTION AND SPLICING OF RNA<sup>EXP</sup>**

Transcription of genes containing expanded repeats plays an important role in the pathogenesis of some REDs, as RNA<sup>exp</sup> can exert toxicity by sequestration of many RBPs as a function of RNA<sup>exp</sup> load. The different noncanonical secondary structures formed by STR expansions need to be resolved during transcription, impeding transcriptional efficiency (Goodman & Bonini, 2020).

Slowed or stalled transcription of STRs can lead to the formation of R-loops in the gene locus. These nucleic acid structures are RNA/DNA hybrids formed during transcription when nascent RNA hybridizes to the DNA template strand behind elongating RNA polymerase II (Pol II) (Thomas et al., 1976). In FXTAS, the GC-rich 5'UTR of *FMR1* is susceptible to cotranscriptional R-loop formation upon reannealing of the nascent *FMR1* pre-mRNA to the complementary DNA strand because a G-rich RNA:C-rich DNA heteroduplex is thermodynamically more stable than the corresponding DNA:DNA duplex (Loomis et al., 2014; Reddy et al., 2011). R-loops over expanded repeats may form a structural block, directly interfering with Pol II transcription elongation and influencing transcription efficiency (Belotserkovskii et al., 2017; Crossley et al., 2019; Groh et al., 2014). Excessive R-loop formation in an individual locus can result in double-stranded DNA brakes and activation of the DNA damage response, triggering a series of signalling events that may be pathogenic (Diab et al., 2018; Loomis et al., 2014). In line with this mechanism, the phosphorylated histone variant  $\gamma$ H2AX, a DNA damage-response molecule, was detected in FXTAS patients (Hoem et al., 2011; Iwahashi et al., 2006). Moreover, methylation caused by the formation of atypical DNA/DNA or DNA/RNA structures in the expanded repeats at a given locus was shown to significantly modulate the expression of mutant genes. For example, in FRDA, R-loops can trigger heterochromatinization, which results in the *FXN* gene silencing (Li et al., 2016).

Expanded CTG (in DM1), G4C2 (in C9-ALS/FTD) and CGG (in FXS) repeats were shown to dysregulate the transcription process *via* inhibition or impairment of Pol II initiation or elongation (Brouwer et al., 2013; Colak et al., 2014; Haeusler et al., 2014). However, at least two protein complexes, the **DSIF** complex and **PAF1** complex (PAF1C), were shown to promote Pol II transcription at repeat expansion sites (Figure 2A).



**Figure 2. Nuclear processing and accumulation of RNA<sup>exp</sup> molecules.**

**(A) Transcription.** DSIF and PAF1 complexes promote transcription of repeat expansion regions through inhibition of formation of DNA secondary structures and R-loops (described for CAG<sup>exp</sup> and G4C2<sup>exp</sup>).

**(B) Splicing.** Majority of pre-mRNAs with expanded repeats undergo correct splicing, however, in some part of mature transcripts intron retention takes place (described for CCUG<sup>exp</sup> and G4C2<sup>exp</sup>). Moreover, G4C2<sup>exp</sup>-containing spliced intron is stabilized in a circular form. Bold line, exon; thin line, intron.

**(C) Sequestration.** RNA<sup>exp</sup> molecules accumulate in nuclei where they bind multiple RBPs and sequester some of them and form RNA<sup>exp</sup> foci (described for CAG<sup>exp</sup>, CUG<sup>exp</sup>, CCUG<sup>exp</sup>, CGG<sup>exp</sup>, G4C2<sup>exp</sup>). **(A, B, C)** arrow with dotted line, change in place and/or in time; solid lines show induction or inhibition of certain processes.

The DSIF complex, composed of two highly conserved proteins, SUPT4H1 and **SPT5H** (or the yeast orthologs Spt4 and Spt5), is a transcription elongation factor that regulates Pol II processivity by reducing the efficiency of its dissociation from template DNA (Wada et al., 1998). During transcription, Spt5 interacts with the Pol II coiled-coil domain and encircles the RNA/DNA hybrid, which may prevent the dissociation of Pol II from the template (Klein et al., 2011; Martinez-Rucobo et al., 2011). Spt4 interacts indirectly with Pol II *via* Spt5, and the DSIF complex interacts with the DNA template outside of the transcription bubble (Klein et al., 2011; Martinez-Rucobo et al., 2011). Spt4 possesses a ZNF domain that probably plays a role in modulating interactions with DNA and stabilizes RNA polymerase/template complexes, preventing Pol II from pausing (Crickard et al., 2016; Wenzel et al., 2008).

SUPT4H1 was shown to be required for the transcription of expanded CAG (C. R. Liu et al., 2012) and G4C2 (Kramer et al., 2016) repeats in HD and C9-ALS/FTD, respectively. Deletion of SPT4-A in mouse striatal neurons expressing *Htt* containing long (>100) CAG repeats (a mouse model of HD) resulted in reduced synthesis of mutant huntingtin containing long polyQ, thus diminishing its aggregation and toxicity (C. R. Liu et al., 2012). *In vivo* studies showed that downregulation of SUPTH1 by delivering antisense oligonucleotides (ASOs) that activate RNaseH-mediated target RNA degradation into the brains of HD model mice reduced the levels of mRNA and protein from the mutant but not the normal *Htt* allele (H. M. Cheng et al., 2015). Deletion of Spt4 in C9-ALS/FTD yeast, *C. elegans* and *Drosophila* models reduced the expression of G4C2<sup>exp</sup> and C4G2<sup>exp</sup> transcripts, blocked accumulation of these transcripts into RNA foci and decreased the levels of RAN-translated polyglycine-proline (polyGP; one of three dipeptide repeats (DPRs) produced from G4C2<sup>exp</sup>) (Figure 1C). These findings were also confirmed in fibroblasts derived from C9-ALS patients (Kramer et al., 2016).

The PAF1C is composed of the highly conserved **PAF1**, **LEO1**, **CDC73**, **CTR9** and **RTF1** proteins and plays a role in the initiation, promoter proximal pausing, elongation and RNA processing/termination stages of transcription (Goodman & Bonini, 2020). Recently, an RNA interference (RNAi)-based screen in a *Drosophila* model of C9-ALS/FTD revealed that the *Drosophila* orthologs of PAF1 and LEO1 (dPaf1 and dLeo1, respectively) are selectively involved in transcription of the G4C2<sup>exp</sup>-containing allele, while downregulation of other PAF1C components affected transcription of both long and short G4C2 tracts. The downregulation of dPaf1 and dLeo1 suppressed the toxicity of repeat expansion in fly tissues (Goodman, Prudencio, Kramer, et al., 2019). Notably, the RNA levels of hPAF1 and hLEO1 were upregulated in post-mortem cortical tissue from C9-ALS patients, supporting the link between the PAF1C and G4C2<sup>exp</sup> in C9-ALS/FTD (Goodman, Prudencio, Kramer, et al., 2019).

Recently, the **AFF2/FMR2** protein, a component of superelongation complex (SEC)-like 2, was found to selectively regulate transcription of the *C9orf72* allele containing long G4C2 tracts in C9-ALS induced pluripotent stem cell (iPSC)-derived neurons (iPSNs) (Yuva-Aydemir et al., 2019). Together, the emerging roles of SUPT4H1 (H.

M. Cheng et al., 2015; Furuta et al., 2019; Kramer et al., 2016; C. R. Liu et al., 2012), the PAF1C (Goodman, Prudencio, Kramer, et al., 2019) and AFF2/FMR2 (Yuva-Aydemir et al., 2019) in the transcription of RNA with expanded STRs suggest that the factors implicated in transcriptional elongation are potential therapeutic targets for REDs.

In eukaryotes, splicing plays a major role in cotranscriptional gene expression, and almost 95% of protein-encoding genes undergo alternative splicing (AS). Previously, it was unclear how expanded repeats present in introns exert toxicity at the RNA and protein levels. While the majority of pre-mRNA molecules with expanded repeats undergo proper splicing and maturation, GC-rich expansion leads to intron retention (IR) (Ł. J. Sznajder et al., 2018; Wang et al., 2021) (Figure 2B). This interesting AS event was observed in some of the genes with repeat expansion, and the resulting RNA molecules contained an unprocessed sequence, which failed to be excised from pre-mRNA and was protected from the 5'- and 3'-ends. IR may be a consequence of spliceosome stalling or abnormal association of RBPs with *cis*-regulatory sequences, which, in the case of expanded repeats, can confer structural arrangements (Taylor & Sobczak, 2020). Although IR can be considered a physiological mechanism (Wong et al., 2013), it may also play a relevant role in REDs with intronic GC-rich sequences. For instance, *C9orf72* mRNA with an enlarged 5'UTR retaining the G4C2<sup>exp</sup> intron can accumulate in the nucleus and was detected in the frontal cortex in heterozygous expansion carriers (Mori et al., 2013; Ł. J. Sznajder et al., 2018). Similarly, CCUG<sup>exp</sup> associated with DM2 was shown to induce the retention of host very long intron 1 and elevated levels of mutant *CNBP* mRNA. Retention of introns containing CCUG<sup>exp</sup> has been detected in many DM2 tissues, including skeletal muscle and the frontal cortex of the brain, and lymphoblastoid cell lines (Ł. J. Sznajder et al., 2018). It was also shown that generally, GC-rich sequences in DNA, due to their secondary structures, can slow the RNA Pol II elongation rate and cause RNA Pol II pausing over retained introns (Velooso et al., 2014). Nevertheless, the trans-acting regulators of IR related to microsatellite repeat disorders remain to be elucidated and require further exploration.

In summary, the transcription and splicing of RNA<sup>exp</sup> depend on the STR sequence and location of the repeats within the gene. Noncanonical secondary structures, including G-quadruplexes and R-loops formed by repeat expansions, impact the transcription rate and/or IR, probably by altering accessibility to trans-acting factors (Table 1), which ultimately has a significant effect on RNA<sup>exp</sup> load and toxicity.

#### 4. SEQUESTRATION OF PROTEINS ON RNA<sup>EXP</sup> MOLECULES

As mentioned above, many REDs have been linked to pathogenic RNA gain-of-function models, also termed RNA toxicity, in which mutant RNA<sup>exp</sup> molecules accumulate in the nucleus, forms aggregates and sequesters specific RBPs within nuclear foci (Figure 2C, Table 1). This in turn leads to functional depletion of these proteins and impairment of their physiological functions (N. Zhang & Ashizawa, 2017). Depending on not only the type and load of expanded repeats but also on the tissue type and availability of RBPs, RNA<sup>exp</sup> foci can present distinct morphologies and abundance (reviewed in (Wojciechowska & Krzyzosiak, 2011)). It should be noted that some proteins are indirectly sequestered in a process mediated by the interaction of RNA<sup>exp</sup> molecules with other protein partners (Yang & Hu, 2016).

##### DMs

CUG<sup>exp</sup> foci were reported for the first time in muscle and fibroblast biopsies from DM1 patients (Taneja et al., 1995), and they have since been observed in smooth muscle (Cardani et al., 2008), the heart (Mankodi et al., 2005), and the brain (Jiang et al., 2004) as well. Foci formation depends on the length of the CUG<sup>exp</sup> repeat; in biopsies of the *vastus lateralis* muscle, sporadic inclusions were observed within the 70-100 CUG<sup>exp</sup> repeat range (Mankodi, 2001). CUG<sup>exp</sup> repeats fold into very stable RNA hairpin structures, attract **MBNL1** and sequester it away from its normal RNA-binding sites, leading to its functional insufficiency in cells (Miller et al., 2000). MBNL1 is an RBP that recognizes multiple GCs flanked by pyrimidines, and under physiological conditions, the major role of MBNL is the regulation of tissue-specific AS and polyadenylation (Konieczny et al., 2014). Increased activity of MBNL1 and 2 during tissue development shifts the pattern of target mRNA AS from foetal- to adult-specific, while its downregulation in DM1 has the opposite effect (Konieczny et al., 2014). Missplicing of multiple mRNAs is exacerbated in the development of this progressive disease (a consequence of somatic expansion). Abnormalities in the AS

of muscle-specific chloride channel (*CLCN1*) (Kino et al., 2009), insulin receptor (*INSR*) (Ho et al., 2004), bridging integrator 1 (*BIN1*) (Fugier et al., 2011) and calcium channel voltage-dependent L type alpha 1s subunit (*CACNA1S*) cause specific DM symptoms, including myotonia, insulin resistance and muscle weakness, respectively. MBNL1 was also shown to interact with other RBPs (e.g., hnRNP H, H2, H3, F, A2/B1, K, and L; the probable ATP-dependent RNA helicases DDX5 and DDX17; and ATP-dependent RNA helicase A (DHX9)) (Paul et al., 2011). Although only a fraction of these interactors colocalize with CUG<sup>exp</sup> foci, long CUG repeats alter the stoichiometry of MBNL1-protein complexes. Increased levels of **hnRNP H**, **H2**, **H3**, and **F** and **DDX5** dysregulate the AS of many DM1-specific RNA targets in a manner similar to MBNL1 depletion, showing that the interruption of functional interactions between MBNL1 and other RBPs may contribute to DM1 pathogenesis (Paul et al., 2011).

Apart from MBNL1 sequestration, other factors also play crucial roles in DM1 pathogenesis. In DM1 myoblasts, skeletal muscle and heart tissues, the steady-state level of CUG-binding protein (**CELF1** aka **CUGBP1**) is augmented, leading to translation defects and misregulation of the AS of some CELF1 mRNA targets. Although the possible sequestration of CELF1 on mutant RNA containing CUG repeats was suggested two decades ago (Timchenko et al., 2001), other studies argued that CELF1 does not colocalize with CUG<sup>exp</sup> foci (Fardaei et al., 2001; Jiang et al., 2004; Mankodi et al., 2005) and that its increased level is a result of hyperphosphorylation by protein kinase C, which leads to an increase in protein half-life and activity (Kuyumcu-Martinez et al., 2007).

Additionally, hnRNP H was shown to bind CUG<sup>exp</sup> *in vivo* and colocalize to a limited extent with RNA foci in DM1 patient-derived myoblasts (D. H. Kim et al., 2005). The level of this protein increases ~2-fold, probably due to the signalling events downstream of CUG<sup>exp</sup> (Paul et al., 2006). Interestingly, hnRNP H was also shown to interact with MBNL1 in normal myoblasts in an RNA-independent manner, and in DM1 myoblasts, elevated levels of MBNL1 increased the colocalization of hnRNP H with RNA foci, suggesting that hnRNP H is recruited to intranuclear DM1 foci by MBNL1 (Paul et al., 2006).

DM2 presents a clinical phenotype similar to that of DM1, even though its symptoms are generally milder, and its progression is less severe (Meola & Cardani, 2017). Ribonuclear foci containing CCUG<sup>exp</sup> were shown to efficiently sequester MBNL proteins (Mankodi, 2001). Moreover, the hnRNP core protein and snRNP Sm antigen were found to colocalize with some MBNL1-CCUG<sup>exp</sup> foci (Perdoni et al., 2009). Recently, *in vitro* studies showed that RNA-binding protein fox-1 homolog 1 (**RBFOX1**) directly binds the CCUG<sup>exp</sup>. Although this regulator of mRNA metabolism colocalized with CCUG<sup>exp</sup> foci in muscle biopsies derived from DM2 patients, its splicing regulatory functions were not impaired (Sellier et al., 2018). Interestingly, RBFOX1 seems to compete with MBNL1 for CCUG repeat-binding sites, which results in reduced sequestration of MBNL proteins by CCUG<sup>exp</sup>. This phenomenon, together with the intronic localization of expanded repeats (described above), may explain the fewer splicing alterations in DM2 than in DM1 (Sellier et al., 2018). Additionally, the misregulated AS patterns in DM1 and DM2 are similar, but the extent of these changes is tissue-specific (Nakamori et al., 2013). Indeed, the muscle weakness associated with DM1 affects distal muscles, while proximal muscles are affected in DM2.

Taken together, the imbalance of multiple interacting proteins may cause aberrant splicing patterns and contribute to DM pathogenesis. The sequestration rate depends on the load of CUG<sup>exp</sup> or CCUG<sup>exp</sup> (repeat length and expression level of RNA with expanded repeats). Due to unequal somatic instability, the efficiency of protein sequestration in distinct tissues or even myofibers differs. Moreover, MBNL proteins are mobile in RNA foci, and when saturated, they can diffuse freely between RNA foci and the nucleoplasm, which may underlie MBNL sequestration (L. J. Sznajder et al., 2016). Multiple therapeutic strategies involving limiting the sequestration or degradation of CUG<sup>exp</sup> were previously developed (reviewed in (Konieczny et al., 2014)).

## **FXTAS**

The pathological hallmark of FXTAS is the presence of solitary, large (2-5 µm), ubiquitin-positive inclusions in the nuclei of neurons, astrocytes (Greco et al., 2002) and Purkinje cells (Ariza et al., 2016). The number of inclusions is correlated with CGG repeat length, suggesting that toxic RNA gain-of-function plays an important

role in FXTAS pathogenesis. In FXTAS, the repeat length is limited to approximately 200 repeats; longer expansions cause the development of a completely different neurodevelopmental disease, a FXS fragile X syndrome, in which complete silencing of *FMR1* occurs (Heitz et al., 1992). *FMR1* mRNA containing CGG<sup>exp</sup> can sequester one or more RBPs, thus impairing their physiological function. Another possible mechanism in FXTAS is that repeat-associated non-AUG (RAN) translation leads to the production of toxic proteins containing monoamino acid tracts, mainly polyG, (Figure 1C, described in the chapter on RAN translation), which colocalize with ubiquitin in FXTAS in different regions of brain sections (Todd et al., 2013) and, to a lesser extent, in other tissues (Buijssen et al., 2014). Recently, the direct interaction of CGG<sup>exp</sup> and **polyG** was shown to promote polyG aggregation and lead to neuronal dysfunction (Asamitsu et al., 2021). PolyG colocalizes with Hsp70 (Bonapace et al., 2019; Jin et al., 2003; Todd et al., 2013) and lamina-associated polypeptide (LAP) 2 $\beta$ , eventually leading to abnormal nuclear envelope structure and cell death (Sellier et al., 2017). Together, these data suggest that multiple mechanisms may play a role in FXTAS and that inclusion formation can be triggered by both toxic RNA and RAN translation products.

The composition of inclusions has been investigated in a few studies. Inclusions purified from FXTAS patients were shown to be composed of multiple proteins, including lamin and hnRNP A2 (Iwahashi et al., 2006). Recent studies have identified over two hundred proteins enriched within these inclusions in comparison to whole nuclei in FXTAS (Ma et al., 2019). No predominant protein was observed, but over half of the identified proteins were involved in RNA binding, protein turnover and DNA damage repair. The most enriched proteins included small ubiquitin-related modifier 2 (SUMO 2), ubiquitin and p62/sequestosome-1. Although the mechanism of inclusion formation is still not well understood, these results suggest that these inclusions may be a consequence of protein aggregation that exceeds the threshold of proteasomal degradation (Ma et al., 2019) and that aggregates may contain RNA<sup>exp</sup>.

In accordance with the toxic RNA gain-of-function hypothesis, several proteins were found to bind CGG<sup>exp</sup>. Among them, the cytoplasmic form of **hnRNP A2/B1** interacts directly with CGG<sup>exp</sup> (Jin et al., 2007; Sofola et al., 2007) and mediates the binding of

**CELF1** (Sofola et al., 2007). Importantly, overexpression of hnRNP A2/B1 or CELF1 suppressed CGG<sup>exp</sup> toxicity in *Drosophila*, suggesting their involvement in FXTAS pathogenesis (Jin et al., 2007; Sofola et al., 2007). In *Drosophila*, the overexpression of ALS-associated TAR DNA-binding protein 43 (**TDP-43**) partially alleviated neurodegeneration, probably by preventing sequestration of the hnRNP A2/B1 protein homologs Hrb87F and Hrb98De on CGG<sup>exp</sup> and thus restoring their functions (He et al., 2014). Additionally, the transcriptional activator protein Pur-alpha (**Pur α**) was found to bind CGG<sup>exp</sup> *in vitro* and was also observed in the inclusions of post-mortem brain tissues from patients with FXTAS (Jin et al., 2007). The splicing regulator transformer-2 protein homolog alpha (**TRA2A**) binds CGG<sup>exp</sup> *in vitro* and colocalizes with nuclear CGG<sup>exp</sup> foci and polyG in FXTAS inclusions. Moreover, in cells expressing (CGG)<sub>60</sub>, TRA2A-dependent splicing of genes linked to mental retardation (i.e., *ACTB*) or intellectual disabilities (i.e., *DOCK3*) was impaired (Cid-Samper et al., 2018).

CGG<sup>exp</sup> also sequesters the double-stranded RBP **DGCR8** and its partner, ribonuclease type 3 (**DROSHA**). The enzymatic complex microprocessor, in which these two proteins play a crucial role, is involved in the processing of primary precursors of microRNA (pri-miRNA), and when impaired in FXTAS, mature miRNA levels are reduced in neuronal cells, leading to neurodegeneration (Sellier et al., 2013). Moreover, colocalization studies showed that the DROSHA-DGCR8 complex interacts with the splicing regulator KH domain-containing, RNA-binding, signal transduction-associated protein 1 (KHDRBS1, alternative name **SAM68**), which leads to its aggregation in CGG<sup>exp</sup> foci and ultimately to the abnormal AS of several target mRNAs (Sellier et al., 2013). Importantly, overexpression of DGCR8, but not SAM68, rescued neuronal cell death induced by the expression of CGG<sup>exp</sup> in cultured mouse cortical neurons, suggesting that the sequestration of DGCR8 may play a role in the neuronal degeneration observed in FXTAS and that SAM68 does not directly bind CGG<sup>exp</sup> (Sellier et al., 2010, 2013).

## **C9-ALS/FTD**

The presence of nuclear sense G4C2 and antisense C4G2 RNA foci transcribed from G4C2<sup>exp</sup> in *C9orf72*, observed in the post-mortem cerebral cortex and spinal cord tissue of C9-ALS/FTD patients, significantly augments the complexity of

proposed RNA-mediated toxicity in this disease (DeJesus-Hernandez et al., 2011). Nuclear foci containing the sense transcript are more abundant in cerebellar granule neurons, while antisense RNA-containing foci are more common in Purkinje cells and motor neurons (Cooper-Knock et al., 2015). Multiple RBPs, including Pur  $\alpha$  (Xu et al., 2013), Nucleolin (Haeusler et al., 2014) and various hnRNPs (Cooper-Knock et al., 2015; Haeusler et al., 2014; Y. B. Lee et al., 2013; Mori, Lammich, et al., 2013), double-stranded RNA-specific editase B2 (ADARB2) (Donnelly et al., 2013), serine/arginine-rich splicing factor 1 (SRSF1) (Cooper-Knock et al., 2014), THO complex subunit 4 (ALYREF) (Cooper-Knock et al., 2014), ZFP106 (Celona et al., 2017), Matrin-3 (Ramesh et al., 2020) and essential paraspeckle proteins (Česník et al., 2019), were shown to interact/colocalize with G4C2<sup>exp</sup> or C4G2<sup>exp</sup> foci, yet the colocalization of a candidate protein with RNA foci does not always mirror its impact on the disease phenotype. More detailed studies to assess the role of RNA-mediated depletion/modification of the activity of these RBPs in C9-ALS/FTD should be performed. Below, we briefly describe some of the candidate proteins and their significance regarding disease phenotype.

*In vitro* studies showed that **Nucleolin**, **hnRNP U**, **hnRNP F**, and 60S ribosomal protein L7 (**RPL7**) bind (G4C2)<sub>4</sub> RNA, while **hnRNP K** preferentially binds antisense (C4G2)<sub>4</sub> RNA (Haeusler et al., 2014). Variable colocalization of Nucleolin with G4C2<sup>exp</sup> foci was further confirmed in motor cortex tissue and Purkinje neurons, suggesting that Nucleolin sequestration can impair Nucleolin homeostasis and result in nucleolar stress (Cooper-Knock et al., 2015; Haeusler et al., 2014). In line with this, an increase in nucleolar volume was observed in frontal cortex neurons containing RNA foci, showing that nucleolar abnormalities are linked to RNA toxicity (Mizielinska et al., 2017).

The interaction between G4C2<sup>exp</sup> and **Pur  $\alpha$**  was initially shown *in vitro* and further confirmed in the mouse Neuro2A cell line and a *Drosophila* model carrying G4C2<sup>exp</sup>; Pur  $\alpha$  inclusions were also observed in ALS/FTD patients with G4C2<sup>exp</sup>. Importantly, the overexpression of Pur  $\alpha$  in both cell and *Drosophila* models alleviated the neurodegeneration mediated by G4C2<sup>exp</sup> (Xu et al., 2013). The colocalization of Pur  $\alpha$  with G4C2 RNA foci was also shown in a zebrafish model of C9-ALS/FTD, in which

its overexpression prevented axonal abnormalities induced by toxic RNA (Swinnen et al., 2018).

**hnRNP H** colocalizes with G4C2<sup>exp</sup> foci, changing its activity, as inclusion of TARBP2 intron 7, normally orchestrated by hnRNP H, was dysregulated in patient brain tissues. Lee *et al.* suggested that sequestration of this protein may enhance the nuclear retention and aggregation of G4C2<sup>exp</sup> RNA, resulting in impaired RNA processing and toxicity (Y. B. Lee et al., 2013). Additionally, hnRNP H1/F, **ALYREF** and **SRSF2** were shown to colocalize in sense RNA foci in cerebellar granule cells and motor neurons derived from C9-ALS/FTD patients (Cooper-Knock et al., 2014). Later, the same group reported colocalization of SRSF1, hnRNP A1, hnRNP H/F, ALYREF and hnRNP K with antisense foci in cerebellar Purkinje neurons derived from C9-ALS/FTD patients and confirmed the direct interaction of these proteins, with the exception of hnRNP K, with synthetic (C4G2)<sub>5</sub> repeats by UV cross-linking studies (Cooper-Knock et al., 2015).

The **ADARB2** protein, an RNA editing regulator, was also shown to colocalize with G4C2<sup>exp</sup> foci in patient tissues (Donnelly et al., 2013). The siRNA-mediated knockdown of *ADARB2* decreased the number of RNA foci in C9-ALS/FTD iPSCs, and the cells devoid of ADARB2 were more susceptible to the toxic effects of an excitatory neurotransmitter (glutamate) and showed increased cell death. The use of antisense nucleotides (ASOs) targeting G4C2<sup>exp</sup> reduced the nuclear retention of ADARB2 and rescued the glutamate susceptibility phenotype (Donnelly et al., 2013).

The ZFN protein **ZFP106** was recently reported to bind G4C2 repeats and other RBPs, including TDP-43 and Fused in sarcoma (FUS), suggesting its potential role in C9-ALS/FTD pathogenesis (Celona et al., 2017). In line with this, *Zfp106*-knockout mice are characterized by motor neuron and muscle fiber degeneration and muscle atrophy, and this phenotype can be reversed by overexpression of ZFP106 in motor neurons. Additionally, overexpression of ZFP106 in a *Drosophila* C9-ALS/FTD model suppressed neurotoxicity, but did not reduce the expression of G4C2<sup>exp</sup> RNA, suggesting that this protein may participate in the toxic RNA gain-of-function mechanism of C9-ALS/FTD (Celona et al., 2017).

Colocalization of the punctate and diffuse forms of **Matrin-3** with G4C2<sup>exp</sup> foci was reported in C9-ALS/FTD iPSNs and post-mortem C9-ALS/FTD motor cortex sections and validated by RNA pull-down assays (Ramesh et al., 2020). The expression of Matrin-3 in G4C2 *Drosophila* models improved eye degeneration and suppressed motor deficits but did not change G4C2 mRNA levels. Interestingly, truncation of the Matrin-3 RBD did not fully inhibit binding to G4C2 RNA, suggesting that this interaction is sometimes direct but also mediated to some extent through other protein partners (Ramesh et al., 2020).

Similar to that in FXTAS, protein sequestration in C9-ALS/FTD is not limited to RNA<sup>exp</sup> foci. RNA<sup>exp</sup> can be translated bidirectionally, resulting in production of the DPRs poly-GP, poly-GA, poly-GR, poly-PR and poly-PA (Figure 1C) (Ash et al., 2013; Gendron et al., 2013; Mori, et al., 2013; Zu et al., 2013). The positively charged, hydrophilic poly-GR repeats were found to form inclusions and colocalize with TDP-43 in an RNA-independent manner in motor neurons derived from C9-ALS/FTD patients, which suggests a role for polyGR in TDP-43 proteinopathy, a hallmark of the majority of ALS and FTD cases (Saber et al., 2018). NCT proteins and nucleoporins also partially colocalized with poly-GR aggregates, suggesting that sequestration of these proteins can contribute to the cytoplasmic accumulation of TDP-43 (Cook et al., 2020). Proteomic analysis of polyGA aggregates in HEK293 cells expressing polyGA-GFP showed enrichment in the UNC119 protein, which is involved in the trafficking of lipidated cargo proteins (May et al., 2014).

In summary, various proteins can be sequestered on RNA<sup>exp</sup>, and as a result, their physiological functions can be impaired in certain REDs (Table 1). Notably, some proteins do not interact with RNA<sup>exp</sup> directly but rather co-aggregate with other protein partners during foci formation. To distinguish between toxic protein sequestration and colocalization, mechanistic studies showing impairment of candidate protein functions should be performed. Moreover, the role of RNA-gain of function model and protein sequestration in the pathogenesis of some REDs should be confirmed, as our current understanding is often based on either *in vitro* studies or simple disease models, with the exception of DM1. Confirmation of pathogenic

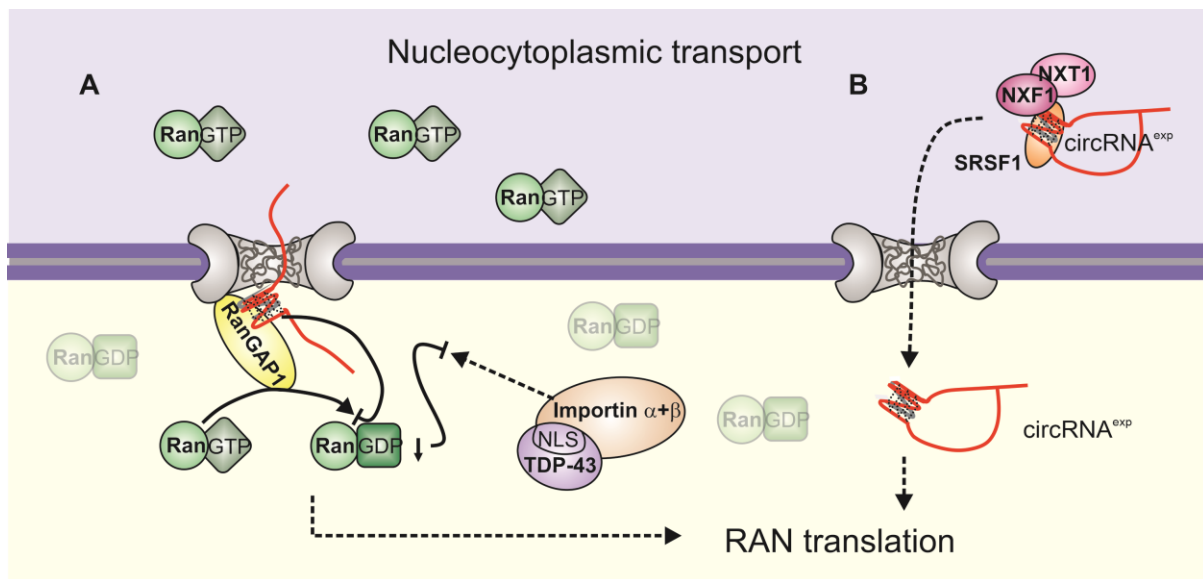
protein sequestration in patient-derived samples would be a valuable source of knowledge on this subject.

#### **Sidebar title: sequestration versus co-localisation**

**Protein sequestration** is selective recruitment and immobilization of specific proteins on RNA molecules in membraneless compartments. Physiological sequestration is used in cellular processes to temporally withdraw proteins from available pool (e.g. nucleolar sequestration of MDM2 leads to stabilization of p53 in the nucleus, which leads to the cell growth arrest or apoptosis). Toxic sequestration is a pathogenic mechanism, where detained proteins are titrated away from their biological targets, leading to irreversible paucity in specific compartments and inhibition of their physiological roles (e.g. sequestration of MBNL1 on CUG<sup>exp</sup> causes altered splicing of different pre-mRNA targets, a hallmark of DM1 pathogenesis). Toxic sequestration is often targeted in RED therapies, by using therapeutic agents which bind to toxic RNA and thus release sequestered proteins. **Colocalization** is a spatial overlap of two probes, e.g. protein and RNA, which can be detected by fluorescence microscopy. It should be distinguished from sequestration, as colocalization of protein and RNA does not necessarily lead to the impairment of protein activity.

## **5. NUCLEOCYTOPLASMIC TRANSPORT**

Impairment of nucleocytoplasmic transport (NCT) is a common pathology in neurodegenerative diseases, including those caused by STR expansion, and a growing body of evidence indicates the great importance of this process in the development of neurodegeneration (extensively reviewed in (H. J. Kim & Taylor, 2017)). A number of reports have suggested the contribution of toxic proteins such as polyQ proteins or RAN translation proteins to the disruption of NCT (Boeynaems et al., 2016; Cook et al., 2020; Grima et al., 2017; Hayes et al., 2020; Jovičić et al., n.d.; K. Y. Shi et al., 2017; Y. J. Zhang et al., 2016). However, some of the articles presented below support the involvement of toxic RNA<sup>exp</sup> due to the direct binding of proteins that participate in NCT (Figure 3A and B, Table 1).



**Figure 3. Involvement of RNA<sup>exp</sup> in nucleocytoplasmic transport (NCT).**

**(A) Impairment of NCT by RNA<sup>exp</sup>.** Gradient of RanGDP/RanGTP proteins between nucleus and cytoplasm, supported by RanGAP1, enables proper NCT. Binding of G4C2<sup>exp</sup> to RanGAP1 leads to impaired import of nuclear proteins, exemplified by TDP-43

**(B) Export of RNA<sup>exp</sup>.** Nuclear export adaptor SRSF1 binds to RNA with G4C2<sup>exp</sup> and C4G2<sup>exp</sup> and supports its transport to cytoplasm through NXF1-dependent pathway. NXF1 and its cofactor NXT1 participate also in export of circular RNA<sup>exp</sup> (circRNA<sup>exp</sup>) derived from G4C2<sup>exp</sup>-bearing intron lariats.

**(A, B)** arrow with dotted line, change in place and/or in time; solid lines show induction or inhibition of certain processes.

### C9-ALS/FTD

RNA with expanded G4C2 repeats exerts toxicity in both the nucleus and cytoplasm; therefore, abnormalities in NCT may significantly contribute to the pathomechanism of C9-ALS/FTD. Genetic screens of a *Drosophila* model of this disease revealed that the pathology caused by G4C2 is related to the impairment of NCT and that Ran GTPase-activating protein (**RanGAP**, RanGAP1 in humans) is a potent suppressor of G4C2 repeat-associated neurodegeneration (Freibaum et al., 2015; K. Zhang et al., 2015) (Figure 3A). RanGAP1 participates in the conversion of cytoplasmic RanGTP to RanGDP, a process required to maintain the Ran protein gradient between the nucleus and cytoplasm, which is necessary for the correct operation of active transport through the nuclear pore complex (Bischoff et al., 1994; Görlich et al., 1996). Binding of RanGAP to the G-quadruplex structure formed by RNA with G4C2 repeats led to disruption of the Ran protein gradient in C9-ALS iPSNs and

inhibition of protein import to the nucleus in both a *Drosophila* model and iPSCs (K. Zhang et al., 2015). This is illustrated by nuclear depletion of **TDP-43** over the course of ALS/FTD (Neumann et al., 2006; K. Zhang et al., 2015). On the other hand, a recent report showed that an increased concentration of TDP-43 in the cytoplasm may lead to the formation of liquid droplets, further causing mislocalization of RanGAP1 and other proteins engaged in NCT (Chou et al., 2018; Gasset-Rosa et al., 2019). G4C2<sup>exp</sup> repeat-related neurodegeneration and impairment of NCT were abolished by overexpression of RanGAP or the application of molecules that inhibit its interaction with the repeats (K. Zhang et al., 2015).

Among the proteins that directly bind G4C2 repeat RNA nuclear export adaptors are serine/arginine-rich splicing factors (**SRSFs**) (Cooper-Knock et al., 2014) (Figure 3B). Further studies showed that knockdown of **SRSF1** significantly reduced repeat-related neurodegeneration in a fly model of C9-ALS/FTD (Hautbergue et al., 2017). In addition to its splicing regulatory functions, SRSF1 participates in the transport of mRNAs *via* a nuclear RNA export factor 1 (NXF1)-dependent pathway (Y. Huang et al., 2003; Müller-McNicoll et al., 2016). The proposed model assumes that SRSF1 directly binds RNA with expanded G4C2 or antisense C4G2 repeats and supports its export to the cytoplasm driven by NXF1, where it undergoes RAN translation to toxic DPRs (Hautbergue et al., 2017). Importantly, this process specifically concerns expansion-bearing transcripts, as SRSF1 deficiency reduced only the nuclear export of *C9orf72* mRNA that retained an intron with expanded repeats but did not influence the nuclear export of wild-type *C9orf72* mRNA in C9-ALS iPSCs (Hautbergue et al., 2017). Recently, the NXF1-NXT1 pathway was also shown to regulate the nuclear export of circular RNA containing G4C2<sup>exp</sup> produced due to defective debranching of the spliced intron of *C9orf72* mRNA (Wang et al., 2021). Here, the presence of G-rich repeats influences the stability of the circular form and its export to the cytoplasm. Interestingly, such circular RNA<sup>exp</sup> undergoes RAN translation and supports DPR production (Wang et al., 2021).

Although direct binding was not shown, a recent study suggests that RNA with G4C2<sup>exp</sup> contributes to the reduction of nuclear envelope pore membrane protein POM 121 (**POM121**) (Coyne et al., 2020). The observed effect was RNA<sup>exp</sup>-specific and independent of DPR translation or loss of the *C9orf72* protein (Coyne et al.,

2020). Deficiency of POM121 resulted in further decreases in a set of nucleoporins, leading to a significant change in nuclear pore complex composition, mislocalization of RanGTPase and reduced survival of C9-ALS iPSNs (Coyne et al., 2020).

### **PolyQ diseases**

Similar to the interplay between SRSF1 and G4C2 observed in ALS, RNAs with expanded CAG repeats interact with another nuclear export adapter, splicing factor U2AF 65 kDa subunit (**U2AF65**) (Tsoi et al., 2011). U2AF65 directly binds RNA with expanded but not normal CAG repeats and then NXF1, facilitating the export of RNA<sup>exp</sup> to the cytoplasm (Tsoi et al., 2011). Knockdown of the *Drosophila* ortholog of U2AF65 enhanced the nuclear accumulation of RNA<sup>exp</sup> and RNA<sup>exp</sup>-mediated toxicity (Tsoi et al., 2011). In a mouse model of polyQ disease expressing RNA with CAG<sup>exp</sup>, symptomatic mice showed a significant reduction in U2AF65 in comparison to U2AF65 levels in asymptomatic mice, which confirmed the contribution of this protein to CAG<sup>exp</sup>-related pathology (Tsoi et al., 2011). Interestingly, subsequent work suggested that nuclear retention of CAG<sup>exp</sup>-bearing RNA results from its binding to MBNL1, leading to impairment of the interaction with U2AF65 and nuclear export by NXF1 (Sun et al., 2015).

### **FXTAS**

Some RBPs that contribute to NCT may also affect this process in FXTAS due to their binding to CGG<sup>exp</sup> RNA. A recent study presented the direct interaction of **SRSF1** with these repeats (Malik, Tseng, et al., 2021). In line with the previously shown effects of the SRSF1/G4C2 interaction (Hautbergue et al., 2017), depletion of SRSF1 activity in a cellular model of FXTAS decreased RAN translation and led to an increased abundance of CGG<sup>exp</sup>-bearing transcripts in the nucleus (Malik, Tseng, et al., 2021). SRSF1 knockdown rescued repeat-related toxicity in both C9-ALS/FTD and FXTAS *Drosophila* models (Hautbergue et al., 2017; Malik, Tseng, et al., 2021). Interestingly, articles reporting the pathological impact of RNA with G4C2 (Coyne et al., 2020) and CAG (Tsoi et al., 2011) repeats on NCT proteins showed that CGG-related toxicity did not induce mislocalization of POM121 and was not modulated by U2AF65.

## DMs

*DMPK* transcripts containing very long expanded CUG repeats are almost fully retained in the nucleus (Brook et al., 1992; Davis et al., 1997). Nuclear export of such mRNAs is also modulated by RBPs. **hnRNP H** was shown to hinder this process (D. H. Kim et al., 2005). In contrast, another CUG<sup>exp</sup>-binding protein, the double-stranded RBP Staufen homolog 1 (**Staufen1**), was shown to rescue NCT of RNA<sup>exp</sup> in cellular models of DM1 (Ravel-Chapuis et al., 2012). In DM2, overexpression of MBNL1 increased RNA foci formation and reduced the production of RAN proteins, which suggests its impact on the nuclear retention of CCUG<sup>exp</sup> (Zu et al., 2017).

In summary, two faces of NCT pathology relate to the interplay between RNA<sup>exp</sup> and RBPs. The first concerns functional impairment of RBPs engaged in different steps of NCT due to their binding to RNA<sup>exp</sup>, as exemplified by RanGAP1 and POM121. The second involves the contribution of RBPs to toxicity exerted by different RNA<sup>exp</sup> molecules in either the nucleus (e.g., CUG<sup>exp</sup>) or cytoplasm (e.g., G4C2<sup>exp</sup>) by enabling their retention or transport to the cytoplasm (e.g., hnRNP H or SRSF1, respectively).

## 6. RAN TRANSLATION

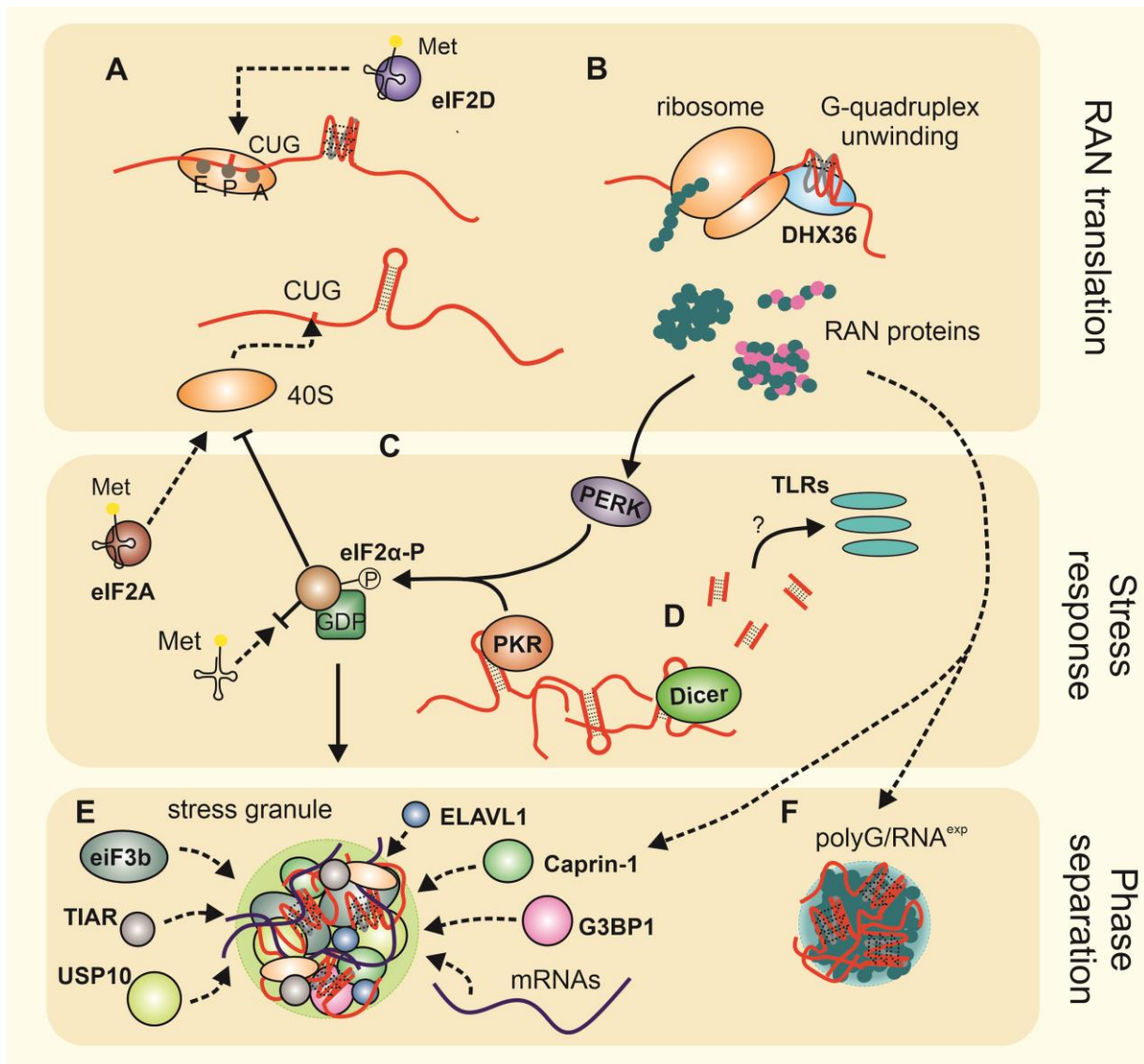
Repeat associated non-AUG (RAN) translation is a protein synthesis mechanism that, in contrast to canonical translation, does not require an AUG start codon for initiation (Zu et al., 2011). RAN translation was reported for many diseases associated with repeat expansions, such as HD (Bañez-Coronel et al., 2015), C9-ALS/FTD (Ash et al., 2013; Mori et al., 2013; Zu et al., 2013), FXTAS (Todd et al., 2013), DM2 (Zu et al., 2017), and others (Banez-Coronel & Ranum, 2019; Zu et al., 2011). RAN proteins may be produced from sense or antisense transcripts and contain homopolymeric tracts comprised of a single amino acid, such as polyglycine (polyG) or polyalanine (polyA) tracts in the case of trinucleotide repeats (Todd et al., 2013), dipeptide repeats (DPRs), such as polyGA in case of G4C2<sup>exp</sup> in C9-ALS/FTD (Ash et al., 2013; Mori et al., 2013; Zu et al., 2013) or tetrapeptide repeats in DM2: polyLPAC and polyQAGR translated from CCUG<sup>exp</sup> and CAGG<sup>exp</sup>, respectively (Zu et al., 2017) (Figure 1C). RAN products have toxic properties; they mostly tend to aggregate, creating nuclear or cytoplasmic inclusions (Mori et al.,

2013; Sellier et al., 2017), and sequester other proteins, resulting in their functional depletion (Sellier et al., 2010, 2013). Insights describing the involvement of RAN translation products in the pathogenesis of particular REDs are beyond the scope of this review and were presented elsewhere (Banez-Coronel & Ranum, 2019). It is worth noting that so far deleterious properties of RAN peptides were mostly observed in experiments based on artificial cellular or animal models (e.g. Sellier et al., 2017; Sonobe et al., 2021; Todd et al., 2013). Indeed, RAN proteins were identified in several patient's samples (Ash et al., 2013; Bañez-Coronel et al., 2015; Mori et al., 2013; Sellier et al., 2017; Todd et al., 2013; Zu et al., 2013, 2017), however, our knowledge about toxicity of those endogenously expressed peptides is limited (Freibaum & Taylor, 2017). For example, current findings concerning presence of polyG in FXTAS patients implicate that RAN translation of CGG<sup>exp</sup> may occur at relatively low level and it is not well-defined to what extent this process contributes to the pathogenesis of this disease (Ma et al., 2019; Salcedo-Arellano et al., 2020). Additionally, RAN translation can also occur when transcripts contain repeats of a normal range, indicating the possible regulatory functions of this process under physiological conditions (C. M. Rodriguez et al., 2020).

Despite the fact that RAN translation was first described nearly a decade ago, mechanistic insights into this process remain elusive (Zu et al., 2011) (Figure 4A-C). In the majority of cases, initiation of RAN translation depends on a canonical cap-dependent scanning model; however, it was shown that AUG-independent translation initiates at less favoured so-called near-cognate codons (such as CUG, GUG, ACG) located upstream or within expanded, structured repeats (Green et al., 2017; Kearse et al., 2016; Tabet et al., 2018). RAN translation of DPRs from G4C2<sup>exp</sup> was shown to initiate with Met-tRNA at CUG start codon in cap-dependent manner (Green et al., 2017; Tabet et al., 2018). The mechanism of translation initiation from alternative start codons has not been fully elucidated; however, one of the proposed models of CGG RAN translation in FXTAS involves a mechanism in which the 43S Preinitiation Complex (PIC) scans through the 5'UTR of *FMR1* mRNA and encounters steric hindered such as hairpins and/or G-quadruplexes formed by expanded repeats, which leads to ribosomal stalling and the initiation of translation at noncanonical codons, resulting in the production of RAN polypeptides (Kearse et al., 2016). Resolving structured RNA seems to be crucial for RAN translation, as it was

shown that inhibition of the canonical DEAD box helicase **eIF4A**, which is needed for ribosome scanning, abolished both CGG<sup>exp</sup>- and G4C2<sup>exp</sup>-dependent RAN translation (Green et al., 2017; Kearse et al., 2016). It was demonstrated that the costimulatory factors of eIF4A, **eIF4B** and **eIF4H**, also exhibit RAN translation modulatory properties (Linsalata et al., 2019). In addition, in a C9-ALS/FTD fly model, the *Drosophila* homologs of eIF4B and eIF4H were shown to modulate the production of polyGR, and eIF4H was found to be significantly downregulated in cerebellar tissue and fibroblasts obtained from patients harbouring G4C2<sup>exp</sup> (Goodman, Prudencio, Srinivasan, et al., 2019). Recently, *Drosophila* Belle (a homolog of the human DEAD-box helicase **DDX3X**) was found to selectively affect the production of RAN peptides (Linsalata et al., 2019). The authors reported that DDX3X binds the 5'UTR of *FRM1* independent of its CGG repeats, excluding the possibility that repeat expansion determines the interaction between mRNA and the protein (Linsalata et al., 2019). Knockdown of Belle/DDX3X in an FXTAS fly model resulted in the decreased production of RAN dipeptides and ameliorated retinal degeneration (as expression of CGG<sup>exp</sup> was directed to the retina using the Gmr-GAL4 driver, which manifests as a severe so-called rough-eye phenotype), confirming the regulatory role of DDX3X *in vivo*. In contrast, DDX3X depletion upregulated G4C2<sup>exp</sup>-specific RAN translation (W. Cheng et al., 2019). This study showed that DDX3X binds G4C2<sup>exp</sup> repeats and relaxes their secondary structure, thus impeding initiation of the RAN translation process (W. Cheng et al., 2019). This is not surprising, as it was previously shown that the yeast homolog of DDX3X, Ded1p, controls initiation from near-cognate codons by binding and unwinding RNA secondary structures upstream of alternative codons within 5'UTRs (Guenther et al., 2018). Recently, another DEAH-box RNA helicase, **DHX36**, was identified as a new modulator of G4C2/CGG RAN translation (H. Liu et al., 2021; Tseng et al., 2021) (Figure 4B). The authors postulate that similar to DDX3X, DHX36 unfolds RNA secondary structures formed within G4C2 repeats and thus modulates noncanonical protein synthesis (H. Liu et al., 2021; Tseng et al., 2021). All of the presented evidence indicates that resolving the RNA<sup>exp</sup> structure by RBPs seems to be a crucial element of RAN translation regulation. Additionally, start codon fidelity seems to substantially contribute to the regulation of RAN translation, as a recently performed screening elucidated two other translation initiation factors, **eIF1** and **eIF5**, that influence the initiation of RAN protein synthesis *in vitro* (Linsalata et al., 2019),

which is in line with previously published reports concerning the function of these proteins, which is contribution to start codon selection (Ivanov et al., 2010; Kozel et al., 2016; Loughran et al., 2012; Tang et al., 2017). Recent *in vitro* studies demonstrated that knockdown of **eIF3F** selectively downregulates synthesis of RAN peptides derived from expanded CAG repeats (related to SCA8) as well as G4C2 repeats (Ayhan et al., 2018). The eIF3F acts as a non-core component of eIF3 complex, which is known to regulate canonical and internal ribosome entry site (IRES)-dependent translation but its role in RAN translation regulation is yet to be defined (Cate, 2017; Hashem et al., 2013). Further studies demonstrating the role of eIF3 in both canonical and RAN translation initiation on CGG<sup>exp</sup> indicated that human oncoprotein **eIF5-mimic (5MP)** displaces eIF5 through eIF3c subunit within the PIC and thus affects both modes of translation *via* increasing the accuracy of translation initiation (Singh et al., 2021; Tang et al., 2017). In addition, it was shown that overexpression of *Drosophila* 5MP homolog in FXTAS flies reduced RAN peptides-mediated toxicity (Singh et al., 2021). Recently, Sonobe and colleagues (2021) developed a new *C. elegans* model of C9-ALS/FTD. The study showed that loss-of-function mutations of **eIF2D** lead to reduced production of polyGA (and to lesser extent polyGP), positively affecting locomotor function and lifespan of transgenic worms (Sonobe et al., 2021). Authors presented that eIF2D is required for DPRs production and perhaps acts at initiation CUG codon for polyGA and delivers the Met-tRNA to the P-site of the 40S ribosomal subunit (Boivin et al., 2020; Dmitriev et al., 2010; Green et al., 2017; Sonobe et al., 2021) (Figure 4A).



**Figure 4. Processes involving RBPs and RNA<sup>exp</sup> in cytoplasm.**

**(A) The role of eIF2D in RAN translation initiation.** Non-canonical translation initiation factor eIF2D deliver Met-tRNA to the P-site of 40S ribosomal subunit at CUG codon contributing to RAN translation initiation (described for G4C2<sup>exp</sup>). A (aminoacyl) site, P (peptidyl) site, E (exit) site in the ribosome.

**(B) Elongation of RAN translation.** DHX36 helicase unwinds G-quadruplexes formed by RNA<sup>exp</sup> and thus facilitates ribosome processivity and production of toxic homopolymeric proteins or dipeptide repeat proteins (described for G4C2<sup>exp</sup> and CGG<sup>exp</sup>).

**(C) RAN translation initiation upon stress.** Stress related with the presence of double-stranded RNA (dsRNA) formed by RNA<sup>exp</sup> and RAN proteins activates PKR and PERK kinases, respectively, which catalyze phosphorylation of eIF2 $\alpha$ . This in turn, inhibits eIF2 $\alpha$ -P binding to Met-tRNA and formation of preinitiation complex (PIC) with 40S ribosome subunit. Under the stress, eIF2A may take over role of phosphorylated eIF2 $\alpha$ -P, bind Met-tRNA and participate in translation initiation at near-cognate start codons (described for G4C2<sup>exp</sup> and CCUG<sup>exp</sup>)

**(D) Stress response caused by RNA<sup>exp</sup>.** RNAi pathway component, ribonuclease Dicer, cleaves dsRNA formed by RNA<sup>exp</sup> hairpin or bidirectionally transcribed

CUG<sup>exp</sup>/CAG<sup>exp</sup> duplex into 21-mer fragments. Such RNA fragments may next activate TLRs and trigger innate immune response.

**(E) Stress granules formation.** Stress stimuli such as RNA<sup>exp</sup> and RAN proteins lead to phosphorylation of eIF2 $\alpha$ , followed by global translation suppression and formation of stress granules (SG) (described for CGG<sup>exp</sup> and G4C2<sup>exp</sup>). Moreover, G4C2<sup>exp</sup> may serve as a core component of SG and promote formation of membraneless organelles composed of different mRNAs and SG protein markers. Chronic stress may entail SG transition towards more solid-like structures.

**(F) Phase separation of RNA<sup>exp</sup> and RAN protein .** Homopolymeric RAN protein, polyG, binds to its own RNA<sup>exp</sup> what promotes its phase transition from liquid droplets towards gel-like aggregates (described for CGG<sup>exp</sup>).

**(A, B, C)** arrow with dotted line, change in place and/or in time; solid lines show induction or inhibition of certain processes.

A genetic screening performed to identify RAN translation-specific modifiers in a yeast model identified 40S ribosomal protein S25-A (**RPS25A**), which selectively inhibited RAN translation from G4C2<sup>exp</sup> (Yamada et al., 2019). Treatment with ASOs targeting RPS25 in C9-ALS iPSCs reduced DPRs production and significantly increased cell survival. Moreover, upon RNAi-mediated silencing in C9-ALS/FTD *Drosophila*, the efficiency of polyGP synthesis significantly decreased, and the lifespan of the adult male flies was extended. Additionally, in human cell lines, RPS25 silencing led to a decrease in RAN translation from another repeats, CAG<sup>exp</sup> associated with the *HTT* and *ATXN2* genes (related to HD and Spinocerebellar ataxia type 2 (SCA2), respectively). The exact function of RPS25 in translation remains to be determined; however, RPS25 was shown to play a crucial role in both cellular and viral noncanonical translation events, such as ribosomal shunting and IRES-mediated translation (Fuchs et al., 2015; Hertz et al., 2013; Landry et al., 2009; Y. Shi et al., 2016). As RPS25 exhibits specificity toward stable RNA secondary structures, structured G4C2 repeats might attract this protein (Nishiyama et al., 2007). Interestingly, RPS25 is not required for cap-dependent translation, which might explain why this factor plays an important role in G4C2 RAN translation of the *C9orf72* transcript, which can be translated in IRES-like cap-independent manner (W. Cheng et al., 2018; Wang et al., 2021).

Altogether, the presented reports imply that RBPs modulate RAN translation by contributing to multiple translation steps, such as PIC scanning, start codon fidelity, unfolding structured RNA<sup>exp</sup> and nuclear retention/NCT of mutated transcripts. Recent findings provide multiple lines of evidence that particular RBPs can modulate

RAN translation in various ways, sometimes leading to contrary effects depending on repeat type, length and RNA sequence context (Table 1).

#### Sidebar title: RAN translation

**Repeat-associated non-AUG (RAN) translation** a non-canonical protein synthesis initiated upstream or within expanded repeats in so-called near cognate start codons such as ACG, CUG, GUG. RAN translated proteins may consist of single amino acids such as polyglycine (FXTAS) or repeated dipeptides (DPRs) (C9-ALS/FTD). These RAN peptides tend to aggregate, interact with both RNAs and other proteins and thus form toxic intracellular inclusions. Mechanistic insights of RAN translation are yet to be established, however, it is known that such processes as ribosomal scanning, start codon fidelity, RNA secondary structure unwinding and cellular stress contribute to RAN translation regulation. RAN translation might be also perceived as combination of **RNA/protein gain-of-function** phenomena as structured RNA<sup>exp</sup> acts as template for production of toxic RAN peptides, affects efficiency of RAN translation and may contribute to the formation of aggregates with RAN products.

## 7. IMMUNE RESPONSE TO RNA<sup>EXP</sup> AND PRODUCTS OF RAN TRANSLATION

Growing evidence suggests that RNA<sup>exp</sup> or double-stranded RNA (dsRNA), e.g. products of bidirectional transcription in *locus* of expanded repeats, or RAN proteins itself can induce innate immune response. They can act as danger-associated molecular patterns (DAMP), and are recognized by pattern recognition receptors (PRRs), including Toll-like receptors (TLR), or double-stranded RNA-activated protein kinase (PKR) in various types of brain cells. These events might correlate with pathological innate immune system activation, induction of stress response and inflammation. The presence of structured RNAs containing different types of repeats, e.g. CUG<sup>exp</sup>, CAG<sup>exp</sup> or G4C2<sup>exp</sup> and RAN proteins, e.g. polyG or polyGP, was postulated to activate **PKR**, and PKR-like endoplasmic reticulum kinase (**PERK**) pathway, respectively (Fig. 4C) (W. Cheng et al., 2018; Green et al., 2017; Tian et al., 2000; Zu et al., 2020). PKR activation leads to the phosphorylation of **eIF2 $\alpha$** , resulting in global translation shutdown, but simultaneously selective upregulation of RAN translation. Expression of expanded CGG and G4C2 repeats in primary rat cortical neurons triggered eIF2A-phosphorylation-dependent stress granule formation. Activation of integrated stress response enhanced RAN translation and halt global translation, contributing to neurodegeneration in a feed-forward loop

(Green et al., 2017). The cytoplasmic dsRNA derived from G4C2 expansions was detected in brains of C9-ALS/FTD patients, and when expressed in mouse model it triggered type-I interferon signalling and neuronal death (S. Rodriguez et al., 2021). Several strategies to inhibit eIF2 $\alpha$  phosphorylation and thus decrease the accumulation of RAN proteins were successfully applied in both *in vitro* and *in vivo* models (W. Cheng et al., 2018; Green et al., 2017; Westergard et al., 2019; Zu et al., 2020). When eIF2 $\alpha$  activity is abolished, eIF2 alternative factor, **eIF2A**, can participate in translation regulation (Komar & Merrick, 2020). Upon eIF2A knockout, synthesis of polyGA from G4C2<sup>exp</sup> was also shown to be decreased (Sonobe et al., 2018). In addition, eIF2A can also increase the production of RAN proteins translated from both CCUG and CAGG repeats in cellular DM2 models (Tusi et al., 2021).

Interestingly, although the CGG<sup>exp</sup> repeats forms hairpins/G-quadruplex, it does not activate the PKR *in vitro* or *in vivo* and is inefficiently cleaved by a major component of RNAi machinery **Dicer** *in vitro* (Handa et al., 2003). On the other hand, in patient-derived cells with mutant *DMPK* or mutant *HTT*, CUG<sup>exp</sup> and CAG<sup>exp</sup> are cleaved by ribonuclease Dicer into short ~21nt RNA fragments (Krol et al., 2007). In *Drosophila* model, transiently expressed transcripts containing CUG<sup>exp</sup> and CAG<sup>exp</sup> were cleaved by Dicer-2 and loaded on Ago2-RISC complex (Yu et al., 2011). This finding was confirmed in another fly model, where dsRNA of CAG/CUG<sub>-100</sub> repeats was produced in endogenous bidirectional transcription, and was shown to be processed by Dicer-2 into CAG<sub>7</sub> 21-mers (Lawlor et al., 2011). Primary activity of Dicer is the cleavage dsRNAs in RNAi pathway; however, its role in antiviral activity was also reported recently (Montavon et al., 2021). Dicer requires several protein partners to cleave dsRNA in RNAi pathway, and recent studies revealed that other key components of RNAi pathway, namely R2D2 and loquacious, are not required for CAG/CUG<sub>-100</sub> dsRNA toxicity in *Drosophila*, in the contrary to Dicer and Argonaute (van Eyk et al., 2019). This suggests that Dicer and Argonaute are implicated in antiviral RNA<sup>exp</sup>-mediated cell death pathway rather than in RNAi pathway.

Adenosine deaminase of RNA-1 (**ADAR1**) is another important player in antiviral immunity, as it edits dsRNA in order to distinguish viral double-stranded nucleic acid ("non-self") from endogenous dsRNA (i.e. Alu repeats). Depending on the number of repeats, RNA<sup>exp</sup> can be conferred as "non-self" and thus triggers autoinflammatory

response. Co-expression of human hADAR1c or hADAR1i with CAG/CUG<sub>100</sub> in *Drosophila* rescued the inflammatory phenotype (van Eyk et al., 2019).

Whole-transcriptomic analysis of *Drosophila* expressing CAG/CUG<sub>-100</sub> repeats revealed alterations in several pathways of innate immunity (Samaraweera et al., 2013). Upon knockdown of **TLRs** in *Drosophila*, the toxicity of CAG/CUG<sub>-100</sub> repeats decreased, whereas opposite effect was observed upon knockdown of autophagy-related genes. This suggests that activation of *Toll* pathway in the presence of dsRNA contributes to neurodegeneration, and that autophagy is required for clearance of pathogenic dsRNA (Samaraweera et al., 2013).

Together, many reports show the link between innate immunity, toxic dsRNA and RAN proteins. These aberrant molecules may induce cellular stress in a feed-forward loop, leading to global protein translation shutdown, but increase in RAN translation, thus contributing to disease progression. Nevertheless, further studies are needed to better understand involvement of immune system in REDs.

## **8. PHASE SEPARATION**

In recent years, efforts to elucidate the underlying causes of neurodegenerative diseases have turned to the phenomenon of phase separation and its contribution to the formation of protein aggregates and RNA foci (comprehensively discussed in excellent reviews (Alberti & Dormann, 2019; Banani et al., 2017; Hyman et al., 2014; Nedelsky & Taylor, 2019; Polymenidou, 2018; Wolozin & Ivanov, 2019)). Since RNA molecules and RBPs are involved in phase separation, alterations in the physiology of this process seem to be particularly relevant for REDs. Briefly, liquid-liquid phase separation (LLPS) is the process in which two fluids demix and one form droplets within the other, as in a mixture of oil and vinegar. In living cells, macromolecules separate from their surroundings to perform specific functions and create assemblies called membraneless organelles (MLOs). Among MLOs in the nucleus are the nucleolus, paraspeckles and Cajal bodies, whereas MLOs in the cytoplasm are exemplified by processing bodies (P-bodies), transport granules or stress granules (SGs). An important feature of these structures is their flexible responses to changing conditions. The transition of liquid droplets to more solid-like foci influences

the dynamics of their assembly and disassembly. This mechanism, widely investigated in SGs, is now considered a source of the toxic foci/aggregates observed in neurodegenerative disorders, including REDs.

Proteins predisposed to undergo LLPS are characterized by multivalency and thus the ability to interact with multiple other proteins or RNA molecules or the presence of intrinsically disordered regions, which are fragments with low sequence complexity that enable noncovalent interactions with other macromolecules (reviewed in (Shin & Brangwynne, 2017)). Many RBPs possess these features and thus tend to undergo LLPS. The contribution of some RBPs and their mutant forms to aberrant phase separation in neurodegenerative disorders has been widely investigated (Bakthavachalu et al., 2018; Bolognesi et al., 2019; Pakravan et al., 2021; Patel et al., 2015) .

Similar to proteins, RNA molecules can undergo phase separation. It was shown that different RNA<sup>exp</sup> molecules (RNAs bearing CUG, CAG and G4C2 repeats), which are multivalent by nature, form droplets in isolation *in vitro* (Jain & Vale, 2017). This phenomenon occurs when the number of repeats exceeds a threshold value, similar to this, which is pathogenic in particular diseases (Jain & Vale, 2017). Consistent with this observations, RNA<sup>exp</sup> molecules accumulate in nuclear foci in different REDs. Interestingly, RNA<sup>exp</sup> droplets *in vitro* are characterized as a gel because of their solid-like features (Jain & Vale, 2017).

In the cellular environment, the formation of MLOs depends on the interactions of proteins and RNAs (Langdon et al., 2018; Maharana et al., 2018; Van Treeck et al., 2018). As stated above, a number of RBPs, including those associated with LLPS, are bound to different RNA<sup>exp</sup> (e.g., TDP-43, FUS, hnRNPs). However, surprisingly, the interplay between RNA<sup>exp</sup> and RBPs in the context of aberrant phase separation in REDs has barely been recognized.

The aforementioned work (Jain & Vale, 2017) concerning the phase separation of RNA<sup>exp</sup> *in vitro* also examined RNA<sup>exp</sup> droplets in living cells. Induction of the expression of RNA with CAG<sup>exp</sup>, but not with normal CAG repeat number resulted in the formation of liquid-like foci that colocalized with serine/arginine-rich splicing factor 2 (SRSF2 aka SC-35), a marker of nuclear speckles, MLOs involved in

splicing regulation (Mintz et al., 1999). Thus, the interaction of CAG<sup>exp</sup> RNA with nuclear speckle components leads to their retention in the nucleus (Jain & Vale, 2017). RNA with G4C2 repeats, but not antisense RNA with C4G2 repeats, formed nuclear foci in a length-dependent manner *in vitro* and in living cells (Jain & Vale, 2017). The observed difference may be associated with the secondary structures of those molecules, as the presence of G-quadruplexes promotes phase separation (Fay et al., 2017; Jain & Vale, 2017; Y. Zhang et al., 2019). Moreover, foci driven by the expression of G4C2<sup>exp</sup> RNA are less dynamic than CAG<sup>exp</sup> RNA foci but are similarly retained in nuclei and colocalize with nuclear speckles (Jain & Vale, 2017; Y. B. Lee et al., 2013).

RNA bearing G4C2 repeats was also shown to drive the formation of SGs in a length-dependent manner both *in vitro* and *in cellulo* (Fay et al., 2017; Rossi et al., 2015) (Figure 4E). RBPs specific for these MLOs were identified in foci assembled *in vitro* with the use of G4C2 RNA and lysates derived from different cell lines or the mouse brain (Fay et al., 2017). Condensed SG markers included Ras GTPase-activating protein-binding protein 1 (**G3BP1**), **Caprin-1**, ubiquitin carboxyl-terminal hydrolase 10 (**USP10**), eukaryotic translation initiation factor 3 subunit B (**eIF3b**), ELAV-like protein 1 (ELAVL1) and nucleolysin **TIAR** (Fay et al., 2017). This process was promoted by the G-quadruplex structure of G4C2 repeats (Fay et al., 2017). RNA<sup>exp</sup> bearing CAG and G4C2 repeats may also be components of mRNA transport granules in neurons, as they undergo active transport in neurites (Burguete et al., 2015). G4C2<sup>exp</sup> RNA molecules interfere with the functions of these MLOs, possibly due to interaction with their RBP components, such as **FMRP**, leading to defects in neuritic branching (Burguete et al., 2015).

Interesting data concerning the interaction of RBPs and RNA<sup>exp</sup> in the context of aberrant LLPS come from studies of FXTAS-related CGG<sup>exp</sup> RNA repeats. As for G4C2<sup>exp</sup>, overexpression of CGG<sup>exp</sup> induced the formation of SGs positive for the markers G3BP1 and FMRP (Green et al., 2017). Computational analyses revealed that *FMR1* mRNA carrying CGG<sup>exp</sup> is a candidate scaffolding RNA component of RNP MLOs (RNP granules) exemplified by SG or P-bodies (Cid-Samper et al., 2018). Such scaffolding RNAs are characterized by the presence of secondary structure within UTRs and by a large number of interactions with MLO-related RBPs,

among other characteristics (Cid-Samper et al., 2018). Indeed, the results of computational analysis, which were further validated using *in vitro* assays, showed that CGG<sup>exp</sup> molecules interact with a considerable number of RBPs (Cid-Samper et al., 2018). A significant portion of the interactors are RBPs, which are predicted to form MLOs (Cid-Samper et al., 2018). Considering its strong CGG<sup>exp</sup>-binding and granule-forming propensities, the splicing regulator **TRA2A** was determined to be a potential player in CGG<sup>exp</sup>-related toxicity (Cid-Samper et al., 2018). Its colocalization with polyG-positive nuclear inclusions was then reported in the brains of an FXTAS mouse model and patients (Cid-Samper et al., 2018). Surprisingly, polyG, the product of RAN translation of CGG<sup>exp</sup> embedded in *FMR1* mRNA, was recently described as a protein that binds its own encoding mRNA, preferentially the expanded CGG repeats (Asamitsu et al., 2021) (Figure 4F). It was shown that both RNA with CGG<sup>exp</sup> and polyG undergo phase separation *in vitro* (Asamitsu et al., 2021). Furthermore, the addition of CGG<sup>exp</sup> RNA to polyG enhanced the phase transition, leading to the formation of gel-like aggregates *in vitro* and possibly also in cells (Asamitsu et al., 2021). Taking the presented reports into consideration, one may conclude that the scaffolding properties of CGG<sup>exp</sup>-bearing RNA and possibly other RNA<sup>exp</sup> molecules lead to the formation of droplets containing RBPs and RAN translation products, which then undergo a transition to solid aggregates, the hallmark of neurodegenerative disorders. It is worth noting that DPRs, products of G4C2 repeat RAN translation, were shown to contribute to aberrant phase transition and disruption of MLO dynamics and functions (Boeynaems et al., 2017; K. H. Lee et al., 2016).

In summary, the appearance of such unusual molecules such as RNA<sup>exp</sup> that on the one hand can form foci by themselves and on the other hand are an attractive platform for interaction with RBPs in the cell most likely changes cell flexibility in tuning the processes of MLO assembly and disassembly (Table 1). The presence of different types of inclusions in REDs may be a consequence of this impairment, followed by the liquid-to-solid transition of MLOs.

## **9. TARGETING RBP<sub>s</sub> AS POTENTIAL THERAPEUTIC STRATEGIES**

For the majority of REDs, targeting RNA<sup>exp</sup> molecules by oligonucleotide therapeutics or small compounds stands as a significant strategy directed to combat

repeat-associated toxicity (Scoles & Pulst, 2018). In particular, ASOs were shown to successfully target mutated transcripts harbouring expanded repeats and alleviate RNA<sup>exp</sup>-mediated toxicity related to HD (Lane et al., 2018; Tabrizi et al., 2019), FXTAS (Derbis et al., 2021), C9-ALS/FTD (Donnelly et al., 2013; Lagier-Tourenne et al., 2013; Sareen et al., 2013), DM1 (Jauvin et al., 2017; Wheeler et al., 2009; Wojtkowiak-Szlachcic et al., 2015) and other REDs *in vivo* models (Scoles & Pulst, 2018). Such an approach seems to be very promising, as some ASO-based drugs, such as nusinersen and eteplirsen, which target specific RBP-binding site, have already been approved by the Food and Drug Administration (FDA) and contribute to the treatment of neurodegenerative diseases (Goodkey et al., 2018; Lim et al., 2017). Another approach to tackle RNA<sup>exp</sup>-related toxicity is the use of small molecules targeting the structural motifs of different RNA<sup>exp</sup> (Angelbello et al., 2019, 2021; Khan et al., 2019); current findings concerning this approach are detailed elsewhere (Angelbello et al., 2020; Crunkhorn, 2021), and here, we would like to focus on strategies targeting RBPs that contribute to the pathogenesis of REDs.

One therapeutic strategy for DM1 is to restore the activity of MBNL proteins, which are sequestered on CUG<sup>exp</sup> or CCUG<sup>exp</sup>, to its physiological level, which seems to be a successful approach to correct AS defects and improve overall phenotype (Chamberlain & Ranum, 2012; Kanadia et al., 2006; Wheeler et al., 2009). Flow cytometry-based genetic screening allowed the identification of two histone deacetylase (HDAC) inhibitors, ISOX and vorinostat, which increased MBNL1 levels and resulted in the partial correction of alternative splicing in DM1 patient-derived fibroblasts (F. Zhang et al., 2017). A recent screening revealed that the inhibition of cyclooxygenase 1 (COX-1) by nonsteroidal anti-inflammatory drugs (NSAIDs) caused demethylation of the MeR2 enhancer and restored *MBNL1* mRNA levels in myogenic cells (K. Huang et al., 2020). The another NSAID, phenylbutazone (PBZ) was shown to suppress methylation of an enhancer region in *Mbnl1* intron 1 and successfully upregulate MBNL1 levels in cells and a DM1 mouse model (G. Chen et al., 2016). Moreover, PBZ treatment reduced the recruitment of MBNL1 to CUG<sup>exp</sup> foci *in cellulo*, ameliorated splicing, mitigated muscle pathology and improved scores on mouse physical testing (G. Chen et al., 2016). Such results prove the potential value of NSAIDs as potent therapeutic candidates for DM1, especially because those drugs are already a part of the treatment regimens for other

neurodegenerative diseases (G. Chen et al., 2016; H. Chen et al., 2005; Hirohata et al., 2005; Uaesoontrachoon et al., 2014). Another strategy to increase MBNL1 levels is the use of antagomiRs to silence miR-23b and miR-218, which regulate MBNL1 levels in muscle (Cerro-Herreros et al., 2018). MBNL1 derepression by antagomiRs rescued missplicing of muscle transcripts and improved functional phenotypes in DM1 mice (Cerro-Herreros et al., 2018, 2020).

Targeting pathologically increased autophagy in DM1 is another therapeutic strategy, as treatment with an anti-autophagic drug, chloroquine, in human-derived myoblasts, *Drosophila* model and DM1 mouse elevated the level of MBNL1, corrected splicing, reduced muscle degeneration, extended the lifespan of the flies, and finally improved the functionality and histopathology of murine muscle tissues (Bargiela et al., 2015, 2019). A correlation between low miR-7 levels and increased autophagy was observed in biopsies from DM1 patients; thus, antagomiR-7 treatment was applied, which indeed improved myotube diameter and the fusion capacity in differentiating DM1 myoblasts; however, the effect turned out to be independent of MBNL1, suggesting that miR-7 acts downstream or alongside MBNL1 in the pathogenesis of DM1 (Sabater-Arcis et al., 2020).

One of the pathological phenotypic hallmarks of C9-ALS/FTD is the presence of an excessive amount of TDP-43 in SGs in the patient brain, which causes toxicity in neurons (Barmada et al., 2010). As stress conditions elicit the phosphorylation of eIF2 $\alpha$  regulated by PERK, the small molecule GSK2606414, an inhibitor of PERK, was shown to successfully mitigate TDP-43-dependent toxicity, improve mobility in C9-ALS/FTD flies and increase the survival rate of neurons (H. J. Kim et al., 2014; Wek et al., 2006). Small-molecule screening focused on the modulation of SG formation revealed that molecules with planar moieties, such as mitoxantrone, can decrease the accumulation of TDP-43 and other RBPs in SG (Fang et al., 2019). Another small molecule, nTRD22, was shown to bind the N-terminus of TDP-43, affecting RNA-binding properties and the degradation of TDP-43 *in vitro* and *in vivo* (Mollasalehi et al., 2020).

A study conducted in a C9-ALS/FTD mouse model indicated that inhibiting the stress-triggered **PKR** pathway by metformin treatment reduced the level of RAN translation. Metformin prevented the excessive accumulation of RAN peptides and

improved the mouse phenotype (Zu et al., 2020). Metformin is already approved by the FDA and serves as a well-established treatment for type 2 diabetes (Sanchez-Rangel & Inzucchi, 2017) and is part of the treatment regimens for DM1 (Bassez et al., 2018), FXS (Dy et al., 2018) and HD (Hervás et al., 2017). Recent findings demonstrate that targeting another kinase, SRSF protein kinase 1 (**SRPK1**), the protein responsible for the phosphorylation of SRSF1 by the compounds: SRPIN340 and SPHINX31, significantly suppressed CGG<sup>exp</sup> or G4C2<sup>exp</sup> RAN translation *in vitro*, and depletion of the *Drosophila* SRPK1 homolog reduced repeat-associated toxicity and improved the rough-eye phenotype in flies expressing both types of repeats (Malik, Tseng, et al., 2021).

In summary, the presented findings underlie the value of MBNL1 rescue and regulation of autophagy in DM1 therapy by selected drugs, paving the way for further clinical validation. Targeting kinases related to stress response pathways seems to be encouraging, as the presence of structured RNA<sup>exp</sup> molecules corresponds with elevated cellular stress, which is common among many REDs. However, although some already approved drugs might be applied to combat neurodegenerative diseases, new approaches to enrich existing therapies are still needed.

## Conclusion

RNA<sup>exp</sup> acts as a very attractive platform that supports interactions with a number of RBPs; thus, wherever and whenever RNA<sup>exp</sup> molecules appear, they distract proteins from performing their functions. To exert toxicity and drive the development of REDs, RNA<sup>exp</sup> molecules require their “partners in crime” at various stages of their life cycle.

At the transcriptional level, Pol II is required for the transcription of GC-rich repeat sequences. On the one hand, RNA<sup>exp</sup> impairs Pol II initiation and elongation, but in turn, the DSIF complex and PAF1C promote Pol II transcription at RNA<sup>exp</sup> sites and may be potential therapeutic targets of REDs.

NCT is impaired in REDs and leads to imbalance in homeostasis between the nucleus and cytoplasm. Some evidence indicates that RNA<sup>exp</sup> molecules bind RBPs

involved in NCT, disturbing their functions. On the other hand, RNA<sup>exp</sup>/RBPs interplay may influence the cellular localization of RNA<sup>exp</sup> molecules and help them exert toxicity in specific compartments. Thus, it is important to consider potential therapeutic strategies targeting RNA<sup>exp</sup> or interacting RBPs in the context of pathogenic effects triggered by these molecules in different cellular compartments.

Multiple proteins can be sequestered on RNA<sup>exp</sup> *via* direct interaction or coaggregation with protein partners during nuclear foci formation, leading to deterioration of their physiological functions. This impairment can often be observed as symptoms of a disease, i.e., in DM1, in which the sequestration of MBNL proteins on CUG<sup>exp</sup> leads to aberrant AS and contributes to the disease phenotype in a CUG<sup>exp</sup> load-dependent manner. Importantly, foci formation by RNA<sup>exp</sup> and sequestered proteins probably changes cell flexibility and impairs LLPS. More studies are needed to decipher the relationship between RNA<sup>exp</sup> and RBPs, which are MLO components, and their contribution to the liquid-to-solid transition leading to inclusion formation. Moreover, due to recent findings concerning the influence of CGG<sup>exp</sup> on the phase transition of its RAN translation product, polyG, a new and exciting research field related to the convergence of RNA gain-of-function and RAN translation/protein gain-of-function mechanisms has opened up.

Toxic RAN peptides derived from non-AUG-initiated translation significantly contribute to the pathogenesis of many REDs. Recent findings provide multiple lines of evidence that RBPs play diverse roles in modulating RAN translation starting from ribosome positioning, the unwinding of structured RNA<sup>exp</sup> and responding to stress conditions. The significance of these RBPs makes them promising therapeutic targets, especially because their modulatory properties have been confirmed by *in vivo* experiments.

Over the last decade, many studies have been devoted to rescue proteins trapped in RNA foci and the modulation of proteins involved in the stress response. Emerging studies suggest that RNA<sup>exp</sup> and RAN proteins can be involved in the activation of innate immune system. The inflammation can account for pathogenesis in many REDs and be a possible target for therapeutic intervention. Much attention has been given to approved drugs and testing them in the context of potential REDs therapy; however, recent discoveries have pointed out a number of potential targets, requiring

medical advancements. Additionally, more research on biomarkers development is desired to enable accurate diagnosis and monitoring a disease progression and effectiveness of given treatment.

### List of abbreviations

**3'UTR / 5'UTR** 3' / 5' untranslated regions

**AS** alternative splicing

**ASO** antisense oligonucleotide

**C9-ALS/FTD** *C9orf72*-linked amyotrophic lateral sclerosis and frontotemporal dementia

**CAG<sup>exp</sup>** expanded CAG repeats

**CCUG<sup>exp</sup>** expanded CCUG repeats

**CGG<sup>exp</sup>** expanded CGG repeats

**CTG<sup>exp</sup>** expanded CTG repeats

**CUG<sup>exp</sup>** expanded CUG repeats

**DM1** myotonic dystrophy type 1

**DM2** myotonic dystrophy type 2

**DPRs** dipeptide-repeat proteins

**FDA** Food and Drug Administration

**FRDA** Friedreich's ataxia

**FXN** frataxin

**FXS** fragile X syndrome

**FXTAS** fragile X-associated tremor/ataxia syndrome

**G4C2<sup>exp</sup>** expanded GGGGCC repeats

**HD** Huntington's disease

**hnRNPs** heterogeneous nuclear ribonucleoproteins

**HTT** huntingtin

**iPSCs** induced pluripotent stem cells

**iPSNs** induced pluripotent stem cells-derived neurons

**IR** intro retention

**KH** K homology domain

**LLPS** liquid-liquid phase separation

**MBNLs** muscleblind-like proteins

**MLO** membraneless organelles

**NCT** nucleocytoplasmic transport  
**REDS** repeats expansion disorders  
**NSAIDs** nonsteroidal anti-inflammatory drugs  
**PAF1C** PAF1 complex  
**P-bodies** processing bodies  
**PBZ** phenylbutazone  
**PERK** protein kinase R-like ER kinase  
**PIC** 43S Preinitiation Complex  
**PKR** Protein Kinase R  
**Pol II** RNA polymerase II  
**polyG** polyglycine  
**polyGA** polyglycine-alanine  
**polyGP** polyglycine-proline  
**polyQ** polyglutamine  
**RAN translation** repeat associated non-AUG translation  
**RBD** RNA binding domain  
**RBPs** RNA binding proteins  
**RNA<sup>exp</sup>** RNA containing expanded repeats  
**RNPs** ribonucleoprotein complexes  
**RRM** RNA recognition motif  
**SCA8** spinocerebellar ataxia type 8  
**SG** stress granules  
**STRs** short tandem repeats

### **Funding Information**

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 101003385 (to AB), Polish National Science Centre No. 2019/35/D/NZ2/02158 (to AB) and No. 2020/38/A/NZ3/00498 (to KS) and No. 2021/40/C/NZ3/00323 (to MD). Funding for proof-reading and open access charge: Initiative of Excellence—Research University No. 05/IDUB/2019/94 at Adam Mickiewicz University, Poznan, Poland.

## References

- Alberti, S., & Dormann, D. (2019). Liquid-Liquid Phase Separation in Disease. In *Annual Review of Genetics* (Vol. 53, pp. 171–194). Annual Reviews Inc. <https://doi.org/10.1146/annurev-genet-112618-043527>
- Angelbello, A. J., Benhamou, R. I., Rzuczek, S. G., Choudhary, S., Tang, Z., Chen, J. L., Roy, M., Wang, K. W., Yildirim, I., Jun, A. S., Thornton, C. A., & Disney, M. D. (2021). A Small Molecule that Binds an RNA Repeat Expansion Stimulates Its Decay via the Exosome Complex. *Cell Chemical Biology*, 28(1), 34-45.e6. <https://doi.org/10.1016/j.chembiol.2020.10.007>
- Angelbello, A. J., Chen, J. L., & Disney, M. D. (2020). Small molecule targeting of RNA structures in neurological disorders. In *Annals of the New York Academy of Sciences* (Vol. 1471, Issue 1, pp. 57–71). Blackwell Publishing Inc. <https://doi.org/10.1111/nyas.14051>
- Angelbello, A. J., Rzuczek, S. G., Mckee, K. K., Chen, J. L., Olafson, H., Cameron, M. D., Moss, W. N., Wang, E. T., & Disney, M. D. (2019). Precise small-molecule cleavage of an r(CUG) repeat expansion in a myotonic dystrophy mouse model. *Proceedings of the National Academy of Sciences of the United States of America*, 116(16), 7799–7804. <https://doi.org/10.1073/pnas.1901484116>
- Ariza, J., Rogers, H., Monterrubio, A., Reyes-Miranda, A., Hagerman, P. J., & Martínez-Cerdeño, V. (2016). A Majority of FXTAS Cases Present with Intranuclear Inclusions Within Purkinje Cells. *Cerebellum*, 15(5), 546–551. <https://doi.org/10.1007/s12311-016-0776-y>
- Asamitsu, S., Yabuki, Y., Ikenoshita, S., Kawakubo, K., Kawasaki, M., Usuki, S., Nakayama, Y., Adachi, K., Kugoh, H., Ishii, K., Matsuura, T., Nanba, E., Sugiyama, H., Fukunaga, K., & Shioda, N. (2021). CGG repeat RNA G-quadruplexes interact with FMRpolyG to cause neuronal dysfunction in fragile X-related tremor/ataxia syndrome. *Science Advances*, 7(3). <https://doi.org/10.1126/sciadv.abd9440>
- Ash, P. E. A., Bieniek, K. F., Gendron, T. F., Caulfield, T., Lin, W. L., DeJesus-Hernandez, M., Van Blitterswijk, M. M., Jansen-West, K., Paul, J. W., Rademakers, R., Boylan, K. B., Dickson, D. W., & Petrucelli, L. (2013). Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron*, 77(4), 639–646. <https://doi.org/10.1016/J.NEURON.2013.02.004>
- Ashizawa, T., Dubel, J. R., Dunne, P. W., Fu, Y. H., Pizzuti, A., Caskey, C. T., Boerwinkle, E.,

- Perryman, M. B., Epstein, H. F., & Hejtmancik, J. F. (1992). Anticipation in myotonic dystrophy: II. complex relationships between clinical findings and structure of the gct repeat. *Neurology*, 42(10), 1877–1883. <https://doi.org/10.1212/wnl.42.10.1877>
- Ayhan, F., Perez, B. A., Shorrock, H. K., Zu, T., Banez-Coronel, M., Reid, T., Furuya, H., Clark, H. B., Troncoso, J. C., Ross, C. A., Subramony, S., Ashizawa, T., Wang, E. T., Yachnis, A. T., & Ranum, L. P. (2018). SCA 8 RAN polySer protein preferentially accumulates in white matter regions and is regulated by eIF 3F . *The EMBO Journal*, 37(19), 1–15. <https://doi.org/10.15252/emj.201899023>
- Bakthavachalu, B., Huelsmeier, J., Sudhakaran, I. P., Hillebrand, J., Singh, A., Petrauskas, A., Thiagarajan, D., Sankaranarayanan, M., Mizoue, L., Anderson, E. N., Pandey, U. B., Ross, E., VijayRaghavan, K., Parker, R., & Ramaswami, M. (2018). RNP-Granule Assembly via Ataxin-2 Disordered Domains Is Required for Long-Term Memory and Neurodegeneration. *Neuron*, 98(4), 754-766.e4. <https://doi.org/10.1016/J.NEURON.2018.04.032>
- Banani, S. F., Lee, H. O., Hyman, A. A., & Rosen, M. K. (2017). Biomolecular condensates: Organizers of cellular biochemistry. In *Nature Reviews Molecular Cell Biology* (Vol. 18, Issue 5, pp. 285–298). Nature Publishing Group. <https://doi.org/10.1038/nrm.2017.7>
- Bañez-Coronel, M., Ayhan, F., Tarabochia, A. D., Zu, T., Perez, B. A., Tusi, S. K., Pletnikova, O., Borchelt, D. R., Ross, C. A., Margolis, R. L., Yachnis, A. T., Troncoso, J. C., & Ranum, L. P. W. (2015). RAN Translation in Huntington Disease. *Neuron*, 88(4), 667–677. <https://doi.org/10.1016/j.neuron.2015.10.038>
- Banez-Coronel, M., & Ranum, L. P. W. (2019). Repeat-associated non-AUG (RAN) translation: insights from pathology. In *Laboratory Investigation* (Vol. 99, Issue 7, pp. 929–942). Nature Publishing Group. <https://doi.org/10.1038/s41374-019-0241-x>
- Bargiela, A., Cerro-Herreros, E., Fernandez-Costa, J. M., Vilchez, J. J., Llamusi, B., & Artero, R. (2015). Increased autophagy and apoptosis contribute to muscle atrophy in a myotonic dystrophy type 1 Drosophila model. *DMM Disease Models and Mechanisms*, 8(7), 679–690. <https://doi.org/10.1242/dmm.018127>
- Bargiela, A., Sabater-Arcis, M., Espinosa-Espinosa, J., Zulaica, M., De Munain, A. L., & Artero, R. (2019). Increased Muscleblind levels by chloroquine treatment improve myotonic dystrophy type 1 phenotypes in in vitro and in vivo models. *Proceedings of the National Academy of Sciences of*

- the United States of America*, 116(50), 25203–25213. <https://doi.org/10.1073/pnas.1820297116>
- Barmada, S. J., Skibinski, G., Korb, E., Rao, E. J., Wu, J. Y., & Finkbeiner, S. (2010). Cytoplasmic mislocalization of TDP-43 is toxic to neurons and enhanced by a mutation associated with familial amyotrophic lateral sclerosis. *Journal of Neuroscience*, 30(2), 639–649. <https://doi.org/10.1523/JNEUROSCI.4988-09.2010>
- Bassez, G., Audureau, E., Hogrel, J. Y., Arrouasse, R., Baghdoyan, S., Bhugaloo, H., Gourlay-Chu, M. L., Le Corvoisier, P., & Peschanski, M. (2018). Improved mobility with metformin in patients with myotonic dystrophy type 1: A randomized controlled trial. *Brain*, 141(10), 2855–2865. <https://doi.org/10.1093/brain/awy231>
- Belotserkovskii, B. P., Shin, J. H. S., & Hanawalt, P. C. (2017). Strong transcription blockage mediated by R-loop formation within a G-rich homopurine-homopyrimidine sequence localized in the vicinity of the promoter. *Nucleic Acids Research*, 45(11), 6589–6599. <https://doi.org/10.1093/nar/gkx403>
- Bischoff, F. R., Klebe, C., Kretschmer, J., Wittinghofer, A., & Ponstingl, H. (1994). RanGAP1 induces GTPase activity of nuclear Ras-related Ran. *Proceedings of the National Academy of Sciences of the United States of America*, 91(7), 2587. <https://doi.org/10.1073/PNAS.91.7.2587>
- Boeynaems, S., Bogaert, E., Kovacs, D., Konijnenberg, A., Timmerman, E., Volkov, A., Guharoy, M., De Decker, M., Jaspers, T., Ryan, V. H., Janke, A. M., Baatsen, P., Vercruyssen, T., Kolaitis, R. M., Daelemans, D., Taylor, J. P., Kedersha, N., Anderson, P., Impens, F., ... Van Den Bosch, L. (2017). Phase Separation of C9orf72 Dipeptide Repeats Perturbs Stress Granule Dynamics. *Molecular Cell*, 65(6), 1044-1055.e5. <https://doi.org/10.1016/j.molcel.2017.02.013>
- Boeynaems, S., Bogaert, E., Michiels, E., Gijssels, I., Sieben, A., Jovičić, A., De Baets, G., Scheveneels, W., Steyaert, J., Cuijt, I., Verstrepen, K. J., Callaerts, P., Rousseau, F., Schymkowitz, J., Cruts, M., Van Broeckhoven, C., Van Damme, P., Gitler, A. D., Robberecht, W., & Van Den Bosch, L. (2016). Drosophila screen connects nuclear transport genes to DPR pathology in c9ALS/FTD. *Scientific Reports*, 6. <https://doi.org/10.1038/srep20877>
- Boivin, M., Pfister, V., Gaucherot, A., Ruffenach, F., Negroni, L., Sellier, C., & Charlet-Berguerand, N. (2020). Reduced autophagy upon C9ORF72 loss synergizes with dipeptide repeat protein toxicity in G4C2 repeat expansion disorders. *The EMBO Journal*, 39(4). <https://doi.org/10.15252/embj.2018100574>

- Bolognesi, B., Faure, A. J., Seuma, M., Schmiedel, J. M., Tartaglia, G. G., & Lehner, B. (2019). The mutational landscape of a prion-like domain. *Nature Communications* 2019 10:1, 10(1), 1–12. <https://doi.org/10.1038/s41467-019-12101-z>
- Bonapace, G., Gullace, R., Concolino, D., Iannello, G., Procopio, R., Gagliardi, M., Arabia, G., Barbagallo, G., Lupo, A., Manfredini, L. I., Annesi, G., & Quattrone, A. (2019). Intracellular FMRpolyG-Hsp70 complex in fibroblast cells from a patient affected by fragile X tremor ataxia syndrome. *Heliyon*, 5(6), e01954. <https://doi.org/10.1016/j.heliyon.2019.e01954>
- Braida, C., Stefanatos, R. K. A., Adam, B., Mahajan, N., Smeets, H. J. M., Niel, F., Goizet, C., Arveiler, B., Koenig, M., Lagier-Tourenne, C., Mandel, J. L., Faber, C. G., de Die-Smulders, C. E. M., Spaans, F., & Monckton, D. G. (2010). Variant CCG and GGC repeats within the CTG expansion dramatically modify mutational dynamics and likely contribute toward unusual symptoms in some myotonic dystrophy type 1 patients. *Human Molecular Genetics*, 19(8), 1399–1412. <https://doi.org/10.1093/hmg/ddq015>
- Brook, J. D., McCurrach, M. E., Harley, H. G., Buckler, A. J., Church, D., Aburatani, H., Hunter, K., Stanton, V. P., Thirion, J. P., Hudson, T., Sohn, R., Zemelmann, B., Snell, R. G., Rundle, S. A., Crow, S., Davies, J., Shelbourne, P., Buxton, J., Jones, C., ... Housman, D. E. (1992). Molecular basis of myotonic dystrophy: Expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell*, 68(4), 799–808. [https://doi.org/10.1016/0092-8674\(92\)90154-5](https://doi.org/10.1016/0092-8674(92)90154-5)
- Brouwer, J. R., Huguet, A., Nicole, A., Munnich, A., & Gourdon, G. (2013). Transcriptionally repressive chromatin remodelling and CpG methylation in the presence of expanded CTG-repeats at the DM1 locus. *Journal of Nucleic Acids*, 2013. <https://doi.org/10.1155/2013/567435>
- Buijsen, R. A. M., Sellier, C., Severijnen, L. A. W. F. M., Oulad-Abdelghani, M., Verhagen, R. F. M., Berman, R. F., Charlet-Berguerand, N., Willemsen, R., & Hukema, R. K. (2014). FMRpolyG-positive inclusions in CNS and non-CNS organs of a fragile X premutation carrier with fragile X-associated tremor/ataxia syndrome. *Acta Neuropathologica Communications*, 2(1), 1–5. <https://doi.org/10.1186/s40478-014-0162-2>
- Burguete, A. S., Almeida, S., Gao, F. B., Kalb, R., Akins, M. R., & Bonini, N. M. (2015). GGGGCC microsatellite RNA is neuritically localized, induces branching defects, and perturbs transport granule function. *eLife*, 4(DECEMBER2015). <https://doi.org/10.7554/eLife.08881>

- Campuzano, V., Montermini, L., Molto, M. D., Pianese, L., Cossee, M., Cavalcanti, F., Monros, E., Duclos, F., Monticelli, A., Zara, F., Cafiizares, J., Koutnikova, H., Bidichandani, S. I., Gellera, C., Brice, A., Trouillas, P., Michele, G. De, Filla, A., Frutos, R. De, ... Pandolfol, M. (1996). Friedreich's Ataxia: Autosomal Recessive Disease Caused by an Intronic GAA Triplet Repeat Expansion. *Science*, 271, 1423–1427.
- Cardani, R., Mancinelli, E., Saino, G., Bonavina, L., & Meola, G. (2008). A putative role of ribonuclear inclusions and MBNL1 in the impairment of gallbladder smooth muscle contractility with cholelithiasis in myotonic dystrophy type 1. *Neuromuscular Disorders*, 18(8), 641–645. <https://doi.org/10.1016/j.nmd.2008.06.366>
- Castel, A. L., Cleary, J. D., & Pearson, C. E. (2010). Repeat instability as the basis for human diseases and as a potential target for therapy. *Nature Reviews Molecular Cell Biology*, 11(3), 165–170. <https://doi.org/10.1038/nrm2854>
- Cate, J. H. D. (2017). Human eIF3: From 'blobology' to biological insight. In *Philosophical Transactions of the Royal Society B: Biological Sciences* (Vol. 372, Issue 1716). Royal Society. <https://doi.org/10.1098/rstb.2016.0176>
- Celona, B., Von Dollen, J., Vatsavayai, S. C., Kashima, R., Johnson, J. R., Tang, A. A., Hata, A., Miller, B. L., Huang, E. J., Krogan, N. J., Seeley, W. W., & Black, B. L. (2017). Suppression of c9orf72 RNA repeat-induced neurotoxicity by the ALS-associated RNA-binding protein Zfp106. *ELife*, 6, 1–17. <https://doi.org/10.7554/eLife.19032>
- Cerro-Herreros, E., González-Martínez, I., Moreno-Cervera, N., Overby, S., Pérez-Alonso, M., Llamusi, B., & Artero, R. (2020). Therapeutic Potential of AntagomiR-23b for Treating Myotonic Dystrophy. *Molecular Therapy - Nucleic Acids*, 21, 837–849. <https://doi.org/10.1016/j.omtn.2020.07.021>
- Cerro-Herreros, E., Sabater-Arcis, M., Fernandez-Costa, J. M., Moreno, N., Perez-Alonso, M., Llamusi, B., & Artero, R. (2018). MiR-23b and miR-218 silencing increase Muscleblind-like expression and alleviate myotonic dystrophy phenotypes in mammalian models. *Nature Communications*, 9(1). <https://doi.org/10.1038/s41467-018-04892-4>
- Česnik, A. B., Darovic, S., Mihevc, S. P., Štalekar, M., Malnar, M., Motaln, H., Lee, Y. B., Mazej, J., Pohleven, J., Grosch, M., Modic, M., Fonovič, M., Turk, B., Drukker, M., Shaw, C. E., & Rogelj, B. (2019). Nuclear RNA foci from C9ORF72 expansion mutation form paraspeckle-like bodies.

*Journal of Cell Science*, 132(5). <https://doi.org/10.1242/jcs.224303>

- Chamberlain, C. M., & Ranum, L. P. W. (2012). Mouse model of muscleblind-like 1 overexpression: Skeletal muscle effects and therapeutic promise. *Human Molecular Genetics*, 21(21), 4645–4654. <https://doi.org/10.1093/hmg/dds306>
- Chen, G., Masuda, A., Konishi, H., Ohkawara, B., Ito, M., Kinoshita, M., Kiyama, H., Matsuura, T., & Ohno, K. (2016). Phenylbutazone induces expression of MBNL1 and suppresses formation of MBNL1-CUG RNA foci in a mouse model of myotonic dystrophy. *Scientific Reports*, 6. <https://doi.org/10.1038/srep25317>
- Chen, H., Jacobs, E., Schwarzschild, M. A., McCullough, M. L., Calle, E. E., Thun, M. J., & Ascherio, A. (2005). Nonsteroidal antiinflammatory drug use and the risk for Parkinson's disease. *Annals of Neurology*, 58(6), 963–967. <https://doi.org/10.1002/ana.20682>
- Cheng, H. M., Chern, Y., Chen, I. H., Liu, C. R., Li, S. H., Chun, S. J., Rigo, F., Bennett, C. F., Deng, N., Feng, Y., Lin, C. S., Yan, Y. T., Cohen, S. N., & Cheng, T. H. (2015). Effects on Murine Behavior and Lifespan of Selectively Decreasing Expression of Mutant Huntingtin Allele by Supt4h Knockdown. *PLoS Genetics*, 11(3), 1–17. <https://doi.org/10.1371/journal.pgen.1005043>
- Cheng, W., Wang, S., Mestre, A. A., Fu, C., Makarem, A., Xian, F., Hayes, L. R., Lopez-Gonzalez, R., Drenner, K., Jiang, J., Cleveland, D. W., & Sun, S. (2018). C9ORF72 GGGGCC repeat-associated non-AUG translation is upregulated by stress through eIF2 $\alpha$  phosphorylation. *Nature Communications*, 9(1). <https://doi.org/10.1038/s41467-017-02495-z>
- Cheng, W., Wang, S., Zhang, Z., Morgens, D. W., Hayes, L. R., Lee, S., Portz, B., Xie, Y., Nguyen, B. V., Haney, M. S., Yan, S., Dong, D., Coyne, A. N., Yang, J., Xian, F., Cleveland, D. W., Qiu, Z., Rothstein, J. D., Shorter, J., ... Sun, S. (2019). CRISPR-Cas9 screens identify the RNA helicase DDX3X as a repressor of C9ORF72 (GGGGCC) $_n$  repeat-associated non-AUG translation. *Neuron*, 104(5), 885. <https://doi.org/10.1016/J.NEURON.2019.09.003>
- Chou, C. C., Zhang, Y., Umoh, M. E., Vaughan, S. W., Lorenzini, I., Liu, F., Sayegh, M., Donlin-Asp, P. G., Chen, Y. H., Duong, D. M., Seyfried, N. T., Powers, M. A., Kukar, T., Hales, C. M., Gearing, M., Cairns, N. J., Boylan, K. B., Dickson, D. W., Rademakers, R., ... Rossoll, W. (2018). TDP-43 pathology disrupts nuclear pore complexes and nucleocytoplasmic transport in ALS/FTD. *Nature Neuroscience*, 21(2), 228–239. <https://doi.org/10.1038/s41593-017-0047-3>
- Cid-Samper, F., Gelabert-Baldrich, M., Lang, B., Lorenzo-Gotor, N., Ponti, R. D., Severijnen, L. A. W.

- F. M., Bolognesi, B., Gelpi, E., Hukema, R. K., Botta-Orfila, T., & Tartaglia, G. G. (2018). An Integrative Study of Protein-RNA Condensates Identifies Scaffolding RNAs and Reveals Players in Fragile X-Associated Tremor/Ataxia Syndrome. *Cell Reports*, 25(12), 3422-3434.e7. <https://doi.org/10.1016/j.celrep.2018.11.076>
- Colak, D., Zaninovic, N., Cohen, M. S., Rosenwaks, Z., Yang, W. Y., Gerhardt, J., Disney, M. D., & Jaffrey, S. R. (2014). Promoter-bound trinucleotide repeat mRNA drives epigenetic silencing in fragile X syndrome. *Science*, 343(6174), 1002–1005. <https://doi.org/10.1126/science.1245831>
- Cook, C. N., Wu, Y., Odeh, H. M., Gendron, T. F., Jansen-West, K., del Rosso, G., Yue, M., Jiang, P., Gomes, E., Tong, J., Daughrity, L. M., Avendano, N. M., Castanedes-Casey, M., Shao, W., Oskarsson, B., Tomassy, G. S., McCampbell, A., Rigo, F., Dickson, D. W., ... Petrucelli, L. (2020). C9orf72 poly(GR) aggregation induces TDP-43 proteinopathy. *Science Translational Medicine*, 12(559), 1–12. <https://doi.org/10.1126/SCITRANSLMED.ABB3774>
- Cooper-Knock, J., Higginbottom, A., Stopford, M. J., Highley, J. R., Ince, P. G., Wharton, S. B., Pickering-Brown, S., Kirby, J., Hautbergue, G. M., & Shaw, P. J. (2015). Antisense RNA foci in the motor neurons of C9ORF72-ALS patients are associated with TDP-43 proteinopathy. *Acta Neuropathologica*, 130(1), 63–75. <https://doi.org/10.1007/s00401-015-1429-9>
- Cooper-Knock, J., Walsh, M. J., Higginbottom, A., Highley, J. R., Dickman, M. J., Edbauer, D., Ince, P. G., Wharton, S. B., Wilson, S. A., Kirby, J., Hautbergue, G. M., & Shaw, P. J. (2014). Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. *Brain*, 137(7), 2040–2051. <https://doi.org/10.1093/brain/awu120>
- Coyne, A. N., Zaepfel, B. L., Hayes, L., Fitchman, B., Salzberg, Y., Luo, E. C., Bowen, K., Trost, H., Aigner, S., Rigo, F., Yeo, G. W., Harel, A., Svendsen, C. N., Sareen, D., & Rothstein, J. D. (2020). G4C2 Repeat RNA Initiates a POM121-Mediated Reduction in Specific Nucleoporins in C9orf72 ALS/FTD. *Neuron*, 107(6), 1124-1140.e11. <https://doi.org/10.1016/j.neuron.2020.06.027>
- Crickard, J. B., Fu, J., & Reese, J. C. (2016). Biochemical analysis of yeast suppressor of Ty 4/5 (Spt4/5) reveals the importance of nucleic acid interactions in the prevention of RNA polymerase II arrest. *Journal of Biological Chemistry*, 291(19), 9853–9870. <https://doi.org/10.1074/jbc.M116.716001>
- Crossley, M. P., Bocek, M., & Cimprich, K. A. (2019). R-Loops as Cellular Regulators and Genomic

Threats. In *Molecular Cell* (Vol. 73, Issue 3, pp. 398–411). Cell Press.

<https://doi.org/10.1016/j.molcel.2019.01.024>

Crunkhorn, S. (2021). Small molecule targets toxic RNA repeats. In *Nature reviews. Drug discovery* (Vol. 20, Issue 1, p. 20). NLM (Medline). <https://doi.org/10.1038/d41573-020-00203-z>

Cumming, S. A., Hamilton, M. J., Robb, Y., Gregory, H., McWilliam, C., Cooper, A., Adam, B., McGhie, J., Hamilton, G., Herzyk, P., Tschannen, M. R., Worthey, E., Petty, R., Ballantyne, B., Warner, J., Farrugia, M. E., Longman, C., & Monckton, D. G. (2018). De novo repeat interruptions are associated with reduced somatic instability and mild or absent clinical features in myotonic dystrophy type 1. *European Journal of Human Genetics*, 26(11), 1635–1647. <https://doi.org/10.1038/s41431-018-0156-9>

Davis, B. M., Mccurrach, M. E., Taneja, K. L., Singer, R. H., & Housman, D. E. (1997). Expansion of a CUG trinucleotide repeat in the 3' untranslated region of myotonic dystrophy protein kinase transcripts results in nuclear retention of transcripts. *Proceedings of the National Academy of Sciences of the United States of America*, 94(14), 7388–7393. <https://doi.org/10.1073/pnas.94.14.7388>

DeJesus-Hernandez, M., Mackenzie, I. R., Boeve, B. F., Boxer, A. L., Baker, M., Rutherford, N. J., Nicholson, A. M., Finch, N. C. A., Flynn, H., Adamson, J., Kouri, N., Wojtas, A., Sengdy, P., Hsiung, G. Y. R., Karydas, A., Seeley, W. W., Josephs, K. A., Coppola, G., Geschwind, D. H., ... Rademakers, R. (2011). Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron*, 72(2), 245–256. <https://doi.org/10.1016/j.neuron.2011.09.011>

Derbis, M., Kul, E., Niewiadomska, D., Sekrecki, M., Piasecka, A., Taylor, K., Hukema, R. K., Stork, O., & Sobczak, K. (2021). Short antisense oligonucleotides alleviate the pleiotropic toxicity of RNA harboring expanded CGG repeats. *Nature Communications*, 12(1), 1265. <https://doi.org/10.1038/s41467-021-21021-w>

Diab, M. A., Mor-Shaked, H., Cohen, E., Cohen-Hadad, Y., Ram, O., Epsztejn-Litman, S., & Eiges, R. (2018). The g-rich repeats in FMR1 and C9orf72 loci are hotspots for local unpairing of DNA. *Genetics*, 210(4), 1239–1252. <https://doi.org/10.1534/genetics.118.301672>

Dmitriev, S. E., Terenin, I. M., Andreev, D. E., Ivanov, P. A., Dunaevsky, J. E., Merrick, W. C., & Shatsky, I. N. (2010). GTP-independent tRNA delivery to the ribosomal P-site by a novel

eukaryotic translation factor. *Journal of Biological Chemistry*, 285(35), 26779–26787.

<https://doi.org/10.1074/jbc.M110.119693>

- Donnelly, C. J., Zhang, P. W., Pham, J. T., Heusler, A. R., Mistry, N. A., Vidensky, S., Daley, E. L., Poth, E. M., Hoover, B., Fines, D. M., Maragakis, N., Tienari, P. J., Petrucelli, L., Traynor, B. J., Wang, J., Rigo, F., Bennett, C. F., Blackshaw, S., Sattler, R., & Rothstein, J. D. (2013). RNA Toxicity from the ALS/FTD C9ORF72 Expansion Is Mitigated by Antisense Intervention. *Neuron*, 80(2), 415–428. <https://doi.org/10.1016/j.neuron.2013.10.015>
- Dy, A. B. C., Tassone, F., Eldeeb, M., Salcedo-Arellano, M. J., Tartaglia, N., & Hagerman, R. (2018). Metformin as targeted treatment in fragile X syndrome. *Clinical Genetics*, 93(2), 216–222. <https://doi.org/10.1111/cge.13039>
- Fang, M. Y., Markmiller, S., Vu, A. Q., Javaherian, A., Dowdle, W. E., Jolivet, P., Bushway, P. J., Castello, N. A., Baral, A., Chan, M. Y., Linsley, J. W., Linsley, D., Mercola, M., Finkbeiner, S., Lecuyer, E., Lewcock, J. W., & Yeo, G. W. (2019). Small-Molecule Modulation of TDP-43 Recruitment to Stress Granules Prevents Persistent TDP-43 Accumulation in ALS/FTD. *Neuron*, 103(5), 802-819.e11. <https://doi.org/10.1016/j.neuron.2019.05.048>
- Fardaei, M., Larkin, K., Brook, J. D., & Hamshere, M. G. (2001). In vivo co-localisation of MBNL protein with DMPK expanded-repeat transcripts. *Nucleic Acids Research*, 29(13), 2766–2771. <https://doi.org/10.1093/nar/29.13.2766>
- Fardaei, M., Rogers, M. T., Thorpe, H. M., Larkin, K., Hamshere, M. G., Harper, P. S., & Brook, J. D. (2002). Three proteins, MBNL, MBLL and MBXL, co-localize in vivo with nuclear foci of expanded-repeat transcripts in DM1 and DM2 cells. *Human Molecular Genetics*, 11(7), 805–814. <https://doi.org/10.1093/hmg/11.7.805>
- Fay, M. M., Anderson, P. J., & Ivanov, P. (2017). ALS/FTD-Associated C9ORF72 Repeat RNA Promotes Phase Transitions In Vitro and in Cells. *Cell Reports*, 21(12), 3573–3584. <https://doi.org/10.1016/j.celrep.2017.11.093>
- Freibaum, B. D., Lu, Y., Lopez-Gonzalez, R., Kim, N. C., Almeida, S., Lee, K.-H., Badders, N., Valentine, M., Miller, B. L., Wong, P. C., Petrucelli, L., Kim, H. J., Gao, F.-B., & Taylor, J. P. (2015). GGGGCC repeat expansion in C9ORF72 compromises nucleocytoplasmic transport HHS Public Access. *Nature*, 525(7567), 129–133. <https://doi.org/10.1038/nature14974>
- Freibaum, B. D., & Taylor, J. P. (2017). The Role of Dipeptide Repeats in C9ORF72-Related ALS-

- FTD. *Frontiers in Molecular Neuroscience*, 10. <https://doi.org/10.3389/FNMOL.2017.00035>
- Fuchs, G., Petrov, A. N., Marceau, C. D., Popov, L. M., Chen, J., OLeary, S. E., Wang, R., Carette, J. E., Sarnow, P., & Puglisi, J. D. (2015). Kinetic pathway of 40S ribosomal subunit recruitment to hepatitis C virus internal ribosome entry site. *Proceedings of the National Academy of Sciences of the United States of America*, 112(2), 319–325. <https://doi.org/10.1073/pnas.1421328111>
- Fugier, C., Klein, A. F., Hammer, C., Vassilopoulos, S., Ivarsson, Y., Toussaint, A., Tosch, V., Vignaud, A., Ferry, A., Messaddeq, N., Kokunai, Y., Tsuburaya, R., De La Grange, P., Dembele, D., Francois, V., Precigout, G., Boulade-Ladame, C., Hummel, M. C., De Munain, A. L., ... Charlet-Berguerand, N. (2011). Misregulated alternative splicing of BIN1 is associated with T tubule alterations and muscle weakness in myotonic dystrophy. *Nature Medicine*, 17(6), 720–725. <https://doi.org/10.1038/nm.2374>
- Furuta, N., Tsukagoshi, S., Hirayanagi, K., & Ikeda, Y. (2019). Suppression of the yeast elongation factor Spt4 ortholog reduces expanded SCA36 GGCCUG repeat aggregation and cytotoxicity. *Brain Research*, 1711(December 2018), 29–40. <https://doi.org/10.1016/j.brainres.2018.12.045>
- Gadgil, R., Barthelemy, J., Lewis, T., & Leffak, M. (2017). Replication stalling and DNA microsatellite instability. *Biophysical Chemistry*, 225, 38–48. <https://doi.org/10.1016/j.bpc.2016.11.007>
- Gasset-Rosa, F., Lu, S., Yu, H., Chen, C., Melamed, Z., Guo, L., Shorter, J., Da Cruz, S., & Cleveland, D. W. (2019). Cytoplasmic TDP-43 De-mixing Independent of Stress Granules Drives Inhibition of Nuclear Import, Loss of Nuclear TDP-43, and Cell Death. *Neuron*, 102(2), 339–357.e7. <https://doi.org/10.1016/j.neuron.2019.02.038>
- Gendron, T. F., Bieniek, K. F., Zhang, Y. J., Jansen-West, K., Ash, P. E. A., Caulfield, T., Daugherty, L., Dunmore, J. H., Castanedes-Casey, M., Chew, J., Cosio, D. M., Van Blitterswijk, M., Lee, W. C., Rademakers, R., Boylan, K. B., Dickson, D. W., & Petrucelli, L. (2013). Antisense transcripts of the expanded C9ORF72 hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS. *Acta Neuropathologica*, 126(6), 829–844. <https://doi.org/10.1007/s00401-013-1192-8>
- Goodkey, K., Aslesh, T., Maruyama, R., & Yokota, T. (2018). Nusinersen in the Treatment of Spinal Muscular Atrophy. In *Methods in Molecular Biology* (Vol. 1828, pp. 69–76). Humana Press Inc. [https://doi.org/10.1007/978-1-4939-8651-4\\_4](https://doi.org/10.1007/978-1-4939-8651-4_4)
- Goodman, L. D., & Bonini, N. M. (2020). New Roles for Canonical Transcription Factors in Repeat

- Expansion Diseases. *Trends in Genetics*, 36(2), 81–92. <https://doi.org/10.1016/j.tig.2019.11.003>
- Goodman, L. D., Prudencio, M., Kramer, N. J., Martinez-Ramirez, L. F., Srinivasan, A. R., Lan, M., Parisi, M. J., Zhu, Y., Chew, J., Cook, C. N., Berson, A., Gitler, A. D., Petrucelli, L., & Bonini, N. M. (2019). Toxic expanded GGGGCC repeat transcription is mediated by the PAF1 complex in C9orf72-associated FTD. *Nature Neuroscience*, 22(6), 863–874. <https://doi.org/10.1038/s41593-019-0396-1>
- Goodman, L. D., Prudencio, M., Srinivasan, A. R., Rifai, O. M., Lee, V. M. Y., Petrucelli, L., & Bonini, N. M. (2019). eIF4B and eIF4H mediate GR production from expanded G4C2 in a Drosophila model for C9orf72-associated ALS. *Acta Neuropathologica Communications*, 7(1), 62. <https://doi.org/10.1186/s40478-019-0711-9>
- Görlich, D., Panté, N., Kutay, U., Aebi, U., & Bischoff, F. R. (1996). Identification of different roles for RanGDP and RanGTP in nuclear protein import. *The EMBO Journal*, 15(20), 5584. <https://doi.org/10.1002/j.1460-2075.1996.tb00943.x>
- Greco, C. M., Berman, R. F., Martin, R. M., Tassone, F., Schwartz, P. H., Chang, A., Trapp, B. D., Iwahashi, C., Brunberg, J., Grigsby, J., Hessler, D., Becker, E. J., Papazian, J., Leehey, M. A., Hagerman, R. J., & Hagerman, P. J. (2006). Neuropathology of fragile X-associated tremor/ataxia syndrome (FXTAS). *Brain*, 129(1), 243–255. <https://doi.org/10.1093/brain/awh683>
- Greco, C. M., Hagerman, R. J., Tassone, F., Chudley, A. E., Del Bigio, M. R., Jacquemont, S., Leehey, M., & Hagerman, P. J. (2002). Neuronal intranuclear inclusions in a new cerebellar tremor/ataxia syndrome among fragile X carriers. *Brain*, 125(8), 1760–1771. <https://doi.org/10.1093/brain/awf184>
- Green, K. M., Glineburg, M. R., Kears, M. G., Flores, B. N., Linsalata, A. E., Fedak, S. J., Goldstrohm, A. C., Barmada, S. J., & Todd, P. K. (2017). RAN translation at C9orf72-associated repeat expansions is selectively enhanced by the integrated stress response. *Nature Communications*, 8(1). <https://doi.org/10.1038/s41467-017-02200-0>
- Grima, J. C., Daigle, J. G., Arbez, N., Cunningham, K. C., Zhang, K., Ochaba, J., Geater, C., Morozko, E., Stocksdales, J., Glatzer, J. C., Pham, J. T., Ahmed, I., Peng, Q., Wadhwa, H., Pletnikova, O., Troncoso, J. C., Duan, W., Snyder, S. H., Ranum, L. P. W., ... Rothstein, J. D. (2017). Mutant Huntingtin Disrupts the Nuclear Pore Complex. *Neuron*, 94(1), 93-107.e6. <https://doi.org/10.1016/j.neuron.2017.03.023>

- Groh, M., Lufino, M. M. P., Wade-Martins, R., & Gromak, N. (2014). R-loops Associated with Triplet Repeat Expansions Promote Gene Silencing in Friedreich Ataxia and Fragile X Syndrome. *PLoS Genetics*, *10*(5). <https://doi.org/10.1371/journal.pgen.1004318>
- Guenther, U. P., Weinberg, D. E., Zubradt, M. M., Tedeschi, F. A., Stawicki, B. N., Zagore, L. L., Brar, G. A., Licatalosi, D. D., Bartel, D. P., Weissman, J. S., & Jankowsky, E. (2018). The helicase Ded1p controls use of near-cognate translation initiation codons in 5' UTRs. *Nature*, *559*(7712), 130–134. <https://doi.org/10.1038/s41586-018-0258-0>
- H, L., YN, L., T, P., G, P., MT, B., AR, F.-D., JD, R., LR, H., S, M., & J, W. (2021). A Helicase Unwinds Hexanucleotide Repeat RNA G-Quadruplexes and Facilitates Repeat-Associated Non-AUG Translation. *Journal of the American Chemical Society*, *143*(19), 7368–7379. <https://doi.org/10.1021/JACS.1C00131>
- Haeusler, A. R., Donnelly, C. J., Periz, G., Simko, E. A. J., Shaw, P. G., Kim, M. S., Maragakis, N. J., Troncoso, J. C., Pandey, A., Sattler, R., Rothstein, J. D., & Wang, J. (2014). C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature*, *507*(7491), 195–200. <https://doi.org/10.1038/nature13124>
- Hagerman, R. J., Leehey, M., Heinrichs, W., Tassone, F., Wilson, R., Hills, J., Grigsby, J., Gage, B., & Hagerman, P. J. (2001). Intention tremor, parkinsonism, and generalized brain atrophy in male carriers of fragile X. *Neurology*, *57*(1), 127–130. <https://doi.org/10.1212/WNL.57.1.127>
- Handa, V., Saha, T., & Usdin, K. (2003). The fragile X syndrome repeats form RNA hairpins that do not activate the interferon-inducible protein kinase, PKR, but are cut by Dicer. *Nucleic Acids Research*, *31*(21), 6243–6248. <https://doi.org/10.1093/nar/gkg818>
- Hashem, Y., Des Georges, A., Dhote, V., Langlois, R., Liao, H. Y., Grassucci, R. A., Pestova, T. V., Hellen, C. U. T., & Frank, J. (2013). Hepatitis-C-virus-like internal ribosome entry sites displace eIF3 to gain access to the 40S subunit. *Nature*, *503*(7477), 539–543. <https://doi.org/10.1038/nature12658>
- Hautbergue, G. M., Castelli, L. M., Ferraiuolo, L., Sanchez-Martinez, A., Cooper-Knock, J., Higginbottom, A., Lin, Y. H., Bauer, C. S., Dodd, J. E., Myszczyńska, M. A., Alam, S. M., Garneret, P., Chandran, J. S., Karyka, E., Stopford, M. J., Smith, E. F., Kirby, J., Meyer, K., Kaspar, B. K., ... Shaw, P. J. (2017). SRSF1-dependent nuclear export inhibition of C9ORF72 repeat transcripts prevents neurodegeneration and associated motor deficits. *Nature*

*Communications*, 8(May), 1–18. <https://doi.org/10.1038/ncomms16063>

Hayes, L. R., Duan, L., Bowen, K., Kalab, P., & Rothstein, J. D. (2020). C9orf72 arginine-rich dipeptide repeat proteins disrupt karyopherin-mediated nuclear import. *ELife*, 9.

<https://doi.org/10.7554/eLife.51685>

He, F., Krans, A., Freibaum, B. D., Paul Taylor, J., & Todd, P. K. (2014). TDP-43 suppresses CGG repeat-induced neurotoxicity through interactions with HnRNP A2/B1. *Human Molecular Genetics*, 23(19), 5036–5051. <https://doi.org/10.1093/hmg/ddu216>

Heitz, D., Devys, D., Imbert, G., Kretz, C., & Mandel, J. L. (1992). Inheritance of the fragile X syndrome: Size of the fragile X premutation is a major determinant of the transition to full mutation. *Journal of Medical Genetics*, 29(11), 794–801. <https://doi.org/10.1136/jmg.29.11.794>

Hertz, M. I., Landry, D. M., Willis, A. E., Luo, G., & Thompson, S. R. (2013). Ribosomal Protein S25 Dependency Reveals a Common Mechanism for Diverse Internal Ribosome Entry Sites and Ribosome Shunting. *Molecular and Cellular Biology*, 33(5), 1016–1026.

<https://doi.org/10.1128/mcb.00879-12>

Hervás, D., Fornés-Ferrer, V., Gómez-Escribano, A. P., Sequedo, M. D., Peiró, C., Millán, J. M., & Vázquez-Manrique, R. P. (2017). Metformin intake associates with better cognitive function in patients with Huntington's disease. *PLoS ONE*, 12(6).

<https://doi.org/10.1371/journal.pone.0179283>

Hirohata, M., Ono, K., Naiki, H., & Yamada, M. (2005). Non-steroidal anti-inflammatory drugs have anti-amyloidogenic effects for Alzheimer's  $\beta$ -amyloid fibrils in vitro. *Neuropharmacology*, 49(7), 1088–1099. <https://doi.org/10.1016/j.neuropharm.2005.07.004>

Ho, T. H., Charlet-B, N., Poulos, M. G., Singh, G., Swanson, M. S., & Cooper, T. A. (2004).

Muscleblind proteins regulate alternative splicing. *EMBO Journal*, 23(15), 3103–3112.

<https://doi.org/10.1038/sj.emboj.7600300>

Hoem, G., Raske, C. R., Garcia-Arocena, D., Tassone, F., Sanchez, E., Ludwig, A. L., Iwahashi, C. K., Kumar, M., Yang, J. E., & Hagerman, P. J. (2011). CGG-repeat length threshold for FMR1 RNA pathogenesis in a cellular model for FXTAS. *Human Molecular Genetics*, 20(11), 2161–2170. <https://doi.org/10.1093/hmg/ddr101>

Huang, K., Masuda, A., Chen, G., Bushra, S., Kamon, M., Araki, T., Kinoshita, M., Ohkawara, B., Ito, M., & Ohno, K. (2020). Inhibition of cyclooxygenase-1 by nonsteroidal anti-inflammatory drugs

- demethylates MeR2 enhancer and promotes Mbnl1 transcription in myogenic cells. *Scientific Reports*, 10(1). <https://doi.org/10.1038/s41598-020-59517-y>
- Huang, Y., Gattoni, R., Stévenin, J., & Steitz, J. A. (2003). SR Splicing Factors Serve as Adapter Proteins for TAP-Dependent mRNA Export. *Molecular Cell*, 11(3), 837–843. [https://doi.org/10.1016/S1097-2765\(03\)00089-3](https://doi.org/10.1016/S1097-2765(03)00089-3)
- Hyman, A. A., Weber, C. A., & Jülicher, F. (2014). Liquid-liquid phase separation in biology. *Annual Review of Cell and Developmental Biology*, 30, 39–58. <https://doi.org/10.1146/annurev-cellbio-100913-013325>
- Ivanov, I. P., Loughran, G., Sachs, M. S., & Atkins, J. F. (2010). Initiation context modulates autoregulation of eukaryotic translation initiation factor 1 (eIF1). *Proceedings of the National Academy of Sciences of the United States of America*, 107(42), 18056–18060. <https://doi.org/10.1073/pnas.1009269107>
- Iwahashi, C. K., Yasui, D. H., An, H. J., Greco, C. M., Tassone, F., Nannen, K., Babineau, B., Lebrilla, C. B., Hagerman, R. J., & Hagerman, P. J. (2006). Protein composition of the intranuclear inclusions of FXTAS. *Brain*, 129(1), 256–271. <https://doi.org/10.1093/brain/awh650>
- Jain, A., & Vale, R. D. (2017). RNA phase transitions in repeat expansion disorders. *Nature*, 546(7657), 243–247. <https://doi.org/10.1038/nature22386>
- Jauvin, D., Chrétien, J., Pandey, S. K., Martineau, L., Revillod, L., Bassez, G., Lachon, A., McLeod, A. R., Gourdon, G., Wheeler, T. M., Thornton, C. A., Bennett, C. F., & Puymirat, J. (2017). Targeting DMPK with Antisense Oligonucleotide Improves Muscle Strength in Myotonic Dystrophy Type 1 Mice. *Molecular Therapy - Nucleic Acids*, 7, 465–474. <https://doi.org/10.1016/j.omtn.2017.05.007>
- Jazurek, M., Ciesiolka, A., Starega-Roslan, J., Bilinska, K., & Krzyzosiak, W. J. (2016). Identifying proteins that bind to specific RNAs - Focus on simple repeat expansion diseases. *Nucleic Acids Research*, 44(19), 9050–9070. <https://doi.org/10.1093/nar/gkw803>
- Jiang, H., Mankodi, A., Swanson, M. S., Moxley, R. T., & 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. *Human Molecular Genetics*, 13(24), 3079–3088. <https://doi.org/10.1093/hmg/ddh327>
- Jin, P., Duan, R., Qurashi, A., Qin, Y., Tian, D., Rosser, T. C., Liu, H., Feng, Y., & Warren, S. T.

- (2007). Pur  $\alpha$  Binds to rCGG Repeats and Modulates Repeat-Mediated Neurodegeneration in a *Drosophila* Model of Fragile X Tremor/Ataxia Syndrome. *Neuron*, 55(4), 556–564.  
<https://doi.org/10.1016/j.neuron.2007.07.020>
- Jin, P., Zarnescu, D. C., Zhang, F., Pearson, C. E., Lucchesi, J. C., Moses, K., & Warren, S. T. (2003). RNA-mediated neurodegeneration caused by the fragile X premutation rCGG repeats in *Drosophila*. *Neuron*, 39, 739–747.
- Jovičić, A., Mertens, J., Boeynaems, S., Bogaert, E., Chai, N., Yamada, S. B., Iij, J. W. P., Sun, S., Herdy, J. R., Bieri, G., Kramer, N. J., Gage, F. H., Robberecht, W., & Gitler, A. D. (n.d.). *Modifiers of C9orf72 DPR toxicity implicate nucleocytoplasmic transport impairments in c9FTD/ALS*. <https://doi.org/10.1038/nn.4085>
- Kanadia, R. N., Shin, J., Yuan, Y., Beattie, S. G., Wheeler, T. M., Thornton, C. A., & Swanson, M. S. (2006). Reversal of RNA missplicing and myotonia after muscleblind overexpression in a mouse poly(CUG) model for myotonic dystrophy. *Proceedings of the National Academy of Sciences of the United States of America*, 103(31), 11748–11753. <https://doi.org/10.1073/pnas.0604970103>
- Kearse, M. G., Green, K. M., Krans, A., Rodriguez, C. M., Linsalata, A. E., Goldstrohm, A. C., & Todd, P. K. (2016). CGG Repeat-Associated Non-AUG Translation Utilizes a Cap-Dependent Scanning Mechanism of Initiation to Produce Toxic Proteins. *Molecular Cell*, 62(2), 314–322.  
<https://doi.org/10.1016/j.molcel.2016.02.034>
- Khan, E., Mishra, S. K., Mishra, R., Mishra, A., & Kumar, A. (2019). Discovery of a potent small molecule inhibiting Huntington's disease (HD) pathogenesis via targeting CAG repeats RNA and Poly Q protein. *Scientific Reports*, 9(1). <https://doi.org/10.1038/s41598-019-53410-z>
- Kim, D. H., Langlois, M. A., Lee, K. B., Riggs, A. D., Puymirat, J., & Rossi, J. J. (2005). HnRNP H inhibits nuclear export of mRNA containing expanded CUG repeats and a distal branch point sequence. *Nucleic Acids Research*, 33(12), 3866–3874. <https://doi.org/10.1093/nar/gki698>
- Kim, H. J., Raphael, A. R., Ladow, E. S., Mcgurk, L., Weber, R. A., Trojanowski, J. Q., Lee, V. M. Y., Finkbeiner, S., Gitler, A. D., & Bonini, N. M. (2014). Therapeutic modulation of eIF2 $\alpha$  phosphorylation rescues TDP-43 toxicity in amyotrophic lateral sclerosis disease models. *Nature Genetics*, 46(2), 152–160. <https://doi.org/10.1038/ng.2853>
- Kim, H. J., & Taylor, J. P. (2017). Lost in Transportation: Nucleocytoplasmic Transport Defects in ALS and Other Neurodegenerative Diseases. In *Neuron* (Vol. 96, Issue 2, pp. 285–297). Cell Press.

<https://doi.org/10.1016/j.neuron.2017.07.029>

Kino, Y., Washizu, C., Oma, Y., Onishi, H., Nezu, Y., Sasagawa, N., Nukina, N., & Ishiura, S. (2009).

MBNL and CELF proteins regulate alternative splicing of the skeletal muscle chloride channel CLCN1. *Nucleic Acids Research*, *37*(19), 6477–6490. <https://doi.org/10.1093/nar/gkp681>

Klein, B. J., Bose, D., Baker, K. J., Yusoff, Z. M., Zhang, X., & Murakami, K. S. (2011). RNA

polymerase and transcription elongation factor Spt4/5 complex structure. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(2), 546–550.

<https://doi.org/10.1073/pnas.1013828108>

Komar, A. A., & Merrick, W. C. (2020). A retrospective on EIF2A—and not the alpha subunit of EIF2.

In *International Journal of Molecular Sciences* (Vol. 21, Issue 6). MDPI AG.

<https://doi.org/10.3390/ijms21062054>

Konieczny, P., Stepniak-Konieczna, E., & Sobczak, K. (2014). MBNL proteins and their target RNAs,

interaction and splicing regulation. *Nucleic Acids Research*, *42*(17), 10873–10887.

<https://doi.org/10.1093/nar/gku767>

Kozel, C., Thompson, B., Hustak, S., Moore, C., Nakashima, A., Singh, C. R., Reid, M., Cox, C.,

Papadopoulos, E., Luna, R. E., Anderson, A., Tagami, H., Hiraishi, H., Slone, E. A., Yoshino, K.

I., Asano, M., Gillaspie, S., Nietfeld, J., Perchet, J. P., ... Asano, K. (2016). Overexpression of

eIF5 or its protein mimic 5MP perturbs eIF2 function and induces ATF4 translation through

delayed re-initiation. *Nucleic Acids Research*, *44*(18), 8704–8713.

<https://doi.org/10.1093/nar/gkw559>

Kozłowski, P., de Mezer, M., & Krzyzosiak, W. J. (2010). Trinucleotide repeats in human genome and

exome. *Nucleic Acids Research*, *38*(12), 4027–4039. <https://doi.org/10.1093/nar/gkq127>

Kramer, N. J., Carlomagno, Y., Zhang, Y. J., Almeida, S., Cook, C. N., Gendron, T. F., Prudencio, M.,

Van Blitterswijk, M., Belzil, V., Couthouis, J., Paul, J. W., Goodman, L. D., Daugherty, L., Chew,

J., Garrett, A., Pregent, L., Jansen-West, K., Tabassian, L. J., Rademakers, R., ... Gitler, A. D.

(2016). Spt4 selectively regulates the expression of C9orf72 sense and antisense mutant

transcripts. *Science*, *353*(6300), 708–712. <https://doi.org/10.1126/science.aaf7791>

Krol, J., Fiszer, A., Mykowska, A., Sobczak, K., de Mezer, M., & Krzyzosiak, W. J. (2007).

Ribonuclease Dicer Cleaves Triplet Repeat Hairpins into Shorter Repeats that Silence Specific

Targets. *Molecular Cell*, *25*(4), 575–586. <https://doi.org/10.1016/j.molcel.2007.01.031>

- Kuyumcu-Martinez, N. M., Wang, G. S., & Cooper, T. A. (2007). Increased Steady-State Levels of CUGBP1 in Myotonic Dystrophy 1 Are Due to PKC-Mediated Hyperphosphorylation. *Molecular Cell*, 28(1), 68–78. <https://doi.org/10.1016/j.molcel.2007.07.027>
- Lagier-Tourenne, C., Baughn, M., Rigo, F., Sun, S., Liu, P., Li, H. R., Jiang, J., Watt, A. T., Chun, S., Katz, M., Qiu, J., Sun, Y., Ling, S. C., Zhu, Q., Polymenidou, M., Drenner, K., Artates, J. W., McAlonis-Downes, M., Markmiller, S., ... Ravits, J. (2013). Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration. *Proceedings of the National Academy of Sciences of the United States of America*, 110(47). <https://doi.org/10.1073/pnas.1318835110>
- Landry, D. M., Hertz, M. I., & Thompson, S. R. (2009). RPS25 is essential for translation initiation by the Dicistroviridae and hepatitis C viral IRESs. *Genes and Development*, 23(23), 2753–2764. <https://doi.org/10.1101/gad.1832209>
- Lane, R. M., Smith, A., Baumann, T., Gleichmann, M., Norris, D., Bennett, C. F., & Kordasiewicz, H. (2018). Translating antisense technology into a treatment for Huntington's disease. In *Methods in Molecular Biology* (Vol. 1780, pp. 497–523). Humana Press Inc. [https://doi.org/10.1007/978-1-4939-7825-0\\_23](https://doi.org/10.1007/978-1-4939-7825-0_23)
- Langdon, E. M., Qiu, Y., Niaki, A. G., McLaughlin, G. A., Weidmann, C. A., Gerbich, T. M., Smith, J. A., Crutchley, J. M., Termini, C. M., Weeks, K. M., Myong, S., & Gladfelter, A. S. (2018). mRNA structure determines specificity of a polyQ-driven phase separation. *Science*, 360(6391), 922–927. <https://doi.org/10.1126/science.aar7432>
- Lawlor, K. T., O'Keefe, L. V., Samaraweera, S. E., van Eyk, C. L., McLeod, C. J., Maloney, C. A., Dang, T. H. Y., Suter, C. M., & Richards, R. I. (2011). Double-stranded RNA is pathogenic in Drosophila models of expanded repeat neurodegenerative diseases. *Human Molecular Genetics*, 20(19), 3757–3768. <https://doi.org/10.1093/hmg/ddr292>
- Lee, K. H., Zhang, P., Kim, H. J., Mitrea, D. M., Sarkar, M., Freibaum, B. D., Cika, J., Coughlin, M., Messing, J., Molliex, A., Maxwell, B. A., Kim, N. C., Temirov, J., Moore, J., Kolaitis, R. M., Shaw, T. I., Bai, B., Peng, J., Kriwacki, R. W., & Taylor, J. P. (2016). C9orf72 Dipeptide Repeats Impair the Assembly, Dynamics, and Function of Membrane-Less Organelles. *Cell*, 167(3), 774–788.e17. <https://doi.org/10.1016/j.cell.2016.10.002>
- Lee, Y. B., Chen, H. J., Peres, J. N., Gomez-Deza, J., Attig, J., Štálekár, M., Troakes, C., Nishimura,

- A. L., Scotter, E. L., Vance, C., Adachi, Y., Sardone, V., Miller, J. W., Smith, B. N., Gallo, J. M., Ule, J., Hirth, F., Rogelj, B., Houart, C., & Shaw, C. E. (2013). Hexanucleotide repeats in ALS/FTD form length-dependent RNA Foci, sequester RNA binding proteins, and are neurotoxic. *Cell Reports*, 5(5), 1178–1186. <https://doi.org/10.1016/j.celrep.2013.10.049>
- Li, L., Matsui, M., & Corey, D. R. (2016). Activating frataxin expression by repeat-targeted nucleic acids. *Nature Communications*, 7. <https://doi.org/10.1038/ncomms10606>
- Lim, K. R. Q., Maruyama, R., & Yokota, T. (2017). Eteplirsen in the treatment of Duchenne muscular dystrophy. In *Drug Design, Development and Therapy* (Vol. 11, pp. 533–545). Dove Medical Press Ltd. <https://doi.org/10.2147/DDDT.S97635>
- Linsalata, A. E., He, F., Malik, A. M., Glineburg, M. R., Green, K. M., Natla, S., Flores, B. N., Krans, A., Archbold, H. C., Fedak, S. J., Barmada, S. J., & Todd, P. K. (2019). DDX 3X and specific initiation factors modulate FMR 1 repeat-associated non-AUG-initiated translation. *EMBO Reports*, 1–18. <https://doi.org/10.15252/embr.201847498>
- Liquori, C. L., Ricker, K., Moseley, M. L., Jacobsen, J. F., Kress, W., Naylor, S. L., Day, J. W., & Ranum, L. P. W. (2001). Myotonic dystrophy type 2 caused by a CCTG expansion in intron I of ZNF9. *Science*, 293(5531), 864–867. <https://doi.org/10.1126/science.1062125>
- Liu, C. R., Chang, C. R., Chern, Y., Wang, T. H., Hsieh, W. C., Shen, W. C., Chang, C. Y., Chu, I. C., Deng, N., Cohen, S. N., & Cheng, T. H. (2012). Spt4 is selectively required for transcription of extended trinucleotide repeats. *Cell*, 148(4), 690–701. <https://doi.org/10.1016/j.cell.2011.12.032>
- Liu, H., Lu, Y. N., Paul, T., Periz, G., Banco, M. T., Ferré-D'Amaré, A. R., Rothstein, J. D., Hayes, L. R., Myong, S., & Wang, J. (2021). A Helicase Unwinds Hexanucleotide Repeat RNA G-Quadruplexes and Facilitates Repeat-Associated Non-AUG Translation. *Journal of the American Chemical Society*, 143(19), 7368–7379. <https://doi.org/10.1021/jacs.1c00131>
- Loomis, E. W., Sanz, L. A., Chédin, F., & Hagerman, P. J. (2014). Transcription-Associated R-Loop Formation across the Human FMR1 CGG-Repeat Region. *PLoS Genetics*, 10(4). <https://doi.org/10.1371/journal.pgen.1004294>
- Loughran, G., Sachs, M. S., Atkins, J. F., & Ivanov, I. P. (2012). Stringency of start codon selection modulates autoregulation of translation initiation factor eIF5. *Nucleic Acids Research*, 40(7), 2898–2906. <https://doi.org/10.1093/nar/gkr1192>
- Lunde, B. M., Moore, C., & Varani, G. (2007). RNA-binding proteins: Modular design for efficient

function. *Nature Reviews Molecular Cell Biology*, 8(6), 479–490.

<https://doi.org/10.1038/nrm2178>

Ma, L., Herren, A. W., Espinal, G., Randol, J., McLaughlin, B., Martinez-Cerdeño, V., Pessah, I. N., Hagerman, R. J., & Hagerman, P. J. (2019). Composition of the Intranuclear Inclusions of Fragile X-associated Tremor/Ataxia Syndrome. *Acta Neuropathologica Communications*, 7(1).  
<https://doi.org/10.1186/s40478-019-0796-1>

MacDonald, M. E., Ambrose, C. M., Duyao, M. P., Myers, R. H., Lin, C., Srinidhi, L., Barnes, G., Taylor, S. A., James, M., Groot, N., MacFarlane, H., Jenkins, B., Anderson, M. A., Wexler, N. S., Gusella, J. F., Bates, G. P., Baxendale, S., Hummerich, H., Kirby, S., ... Harper, P. S. (1993). A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell*, 72(6), 971–983. [https://doi.org/10.1016/0092-8674\(93\)90585-E](https://doi.org/10.1016/0092-8674(93)90585-E)

Maharana, S., Wang, J., Papadopoulos, D. K., Richter, D., Pozniakovsky, A., Poser, I., Bickle, M., Rizk, S., Guillén-Boixet, J., Franzmann, T. M., Jahnel, M., Marrone, L., Chang, Y. T., Sternecker, J., Tomancak, P., Hyman, A. A., & Alberti, S. (2018). RNA buffers the phase separation behavior of prion-like RNA binding proteins. *Science*, 360(6391), 918–921.  
<https://doi.org/10.1126/science.aar7366>

Malik, I., Kelley, C. P., Wang, E. T., & Todd, P. K. (2021). Molecular mechanisms underlying nucleotide repeat expansion disorders. *Nature Reviews Molecular Cell Biology* 2021, 1–19.  
<https://doi.org/10.1038/s41580-021-00382-6>

Malik, I., Tseng, Y., Wright, S. E., Zheng, K., Ramaiyer, P., Green, K. M., & Todd, P. K. (2021). SRSF protein kinase 1 modulates RAN translation and suppresses CGG repeat toxicity. *EMBO Molecular Medicine*, e14163. <https://doi.org/10.15252/emmm.202114163>

Mankodi, A. (2001). Muscleblind localizes to nuclear foci of aberrant RNA in myotonic dystrophy types 1 and 2. *Human Molecular Genetics*, 10(19), 2165–2170.  
<https://doi.org/10.1093/hmg/10.19.2165>

Mankodi, A., Lin, X., Blaxall, B. C., Swanson, M. S., & Thornton, C. A. (2005). Nuclear RNA foci in the heart in myotonic dystrophy. *Circulation Research*, 97(11), 1152–1155.  
<https://doi.org/10.1161/01.RES.0000193598.89753.e3>

Mankodi, A., Logigian, E., Callahan, L., McClain, C., White, R., Henderson, D., Krym, M., & Thornton, C. A. (2000). Myotonic dystrophy in transgenic mice expressing an expanded CUG repeat.

- Science*, 289(5485), 1769–1772. <https://doi.org/10.1126/science.289.5485.1769>
- Martí, E. (2016). RNA toxicity induced by expanded CAG repeats in Huntington's disease. *Brain Pathology*, 26(6), 779–786. <https://doi.org/10.1111/bpa.12427>
- Martinez-Rucobo, F. W., Sainsbury, S., Cheung, A. C. M., & Cramer, P. (2011). Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity. *EMBO Journal*, 30(7), 1302–1310. <https://doi.org/10.1038/emboj.2011.64>
- May, S., Hornburg, D., Schludi, M. H., Arzberger, T., Rentzsch, K., Schwenk, B. M., Grässer, F. A., Mori, K., Kremmer, E., Banzhaf-Strathmann, J., Mann, M., Meissner, F., & Edbauer, D. (2014). C9orf72 FTL/ALS-associated Gly-Ala dipeptide repeat proteins cause neuronal toxicity and Unc119 sequestration. *Acta Neuropathologica*, 128(4), 485–503. <https://doi.org/10.1007/s00401-014-1329-4>
- McMurray, C. T. (2010). Mechanisms of trinucleotide repeat instability during human development. *Nature Reviews Genetics*, 11(11), 786–799. <https://doi.org/10.1038/nrg2828>
- Meola, G., & Cardani, R. (2017). Myotonic dystrophy type 2 and modifier genes: an update on clinical and pathomolecular aspects. *Neurological Sciences*, 38(4), 535–546. <https://doi.org/10.1007/s10072-016-2805-5>
- Miller, J. W., Urbinati, C. R., Teng-umnuay, P., Stenberg, M. G., Byrne, B. J., Thornton, C. A., & Swanson, M. S. (2000). Recruitment of human muscleblind proteins to (CUG)<sub>n</sub> expansions associated with myotonic dystrophy. *The EMBO Journal*, 19(17), 4439–4448. <https://doi.org/10.1093/emboj/19.17.4439>
- Mintz, P. J., Patterson, S. D., Neuwald, A. F., Spahr, C. S., & Spector, D. L. (1999). Purification and biochemical characterization of interchromatin granule clusters. *The EMBO Journal*, 18(15), 4308–4320. <https://doi.org/10.1093/EMBOJ/18.15.4308>
- Mizielinska, S., Ridler, C. E., Balendra, R., Thoeng, A., Woodling, N. S., Grässer, F. A., Plagnol, V., Lashley, T., Partridge, L., & Isaacs, A. M. (2017). Bidirectional nucleolar dysfunction in C9orf72 frontotemporal lobar degeneration. *Acta Neuropathologica Communications*, 5(1), 29. <https://doi.org/10.1186/s40478-017-0432-x>
- Mollasalehi, N., Francois-Moutal, L., Scott, D. D., Tello, J. A., Williams, H., Mahoney, B., Carlson, J. M., Dong, Y., Li, X., Miranda, V. G., Gokhale, V., Wang, W., Barmada, S. J., & Khanna, M. (2020). An allosteric modulator of RNA binding targeting the N-terminal domain of TDP-43 yields

neuroprotective properties. *ACS Chemical Biology*, 15(11), 2854–2859.

<https://doi.org/10.1021/acscchembio.0c00494>

Montavon, T. C., Baldaccini, M., Lefèvre, M., Girardi, E., Chane-Woon-Ming, B., Messmer, M.,

Hammann, P., Chicher, J., & Pfeffer, S. (2021). Human DICER helicase domain recruits PKR and modulates its antiviral activity. In *PLoS Pathogens* (Vol. 17, Issue 5).

<https://doi.org/10.1371/journal.ppat.1009549>

Mori, K., Lammich, S., Mackenzie, I. R. A., Forné, I., Zilow, S., Kretzschmar, H., Edbauer, D.,

Janssens, J., Kleinberger, G., Cruts, M., Herms, J., Neumann, M., Van Broeckhoven, C.,

Arzberger, T., & Haass, C. (2013). 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 Neuropathologica*, 125(3), 413–423. <https://doi.org/10.1007/s00401-013-1088-7>

Mori, K., Weng, S.-M., Arzberger, T., May, S., Rentzsch, K., Kremmer, E., Schmid, B., Kretzschmar,

H. A., Cruts, M., Broeckhoven, C. Van, Haass, C., & Edbauer, D. (2013). The C9orf72

GGGGCC Repeat Is Translated into Aggregating Dipeptide-Repeat Proteins in FTL/ALS.

*Science*, 339(6125), 1335–1338. <https://doi.org/10.1126/SCIENCE.1232927>

Mori, K., Weng, S., Arzberger, T., May, S., Rentzsch, K., Broeckhoven, C. Van, Haass, C., &

Edbauer, D. (2013). *The C9orf72 GGGGCC Repeat Is Translated into Aggregating Dipeptide-Repeat Proteins in FTL/ALS*. 339(March), 1335–1339.

Müller-McNicoll, M., Botti, V., de Jesus Domingues, A. M., Brandl, H., Schwich, O. D., Steiner, M. C.,

Curk, T., Poser, I., Zarnack, K., & Neugebauer, K. M. (2016). SR proteins are NXF1 adaptors

that link alternative RNA processing to mRNA export. *Genes and Development*, 30(5), 553–566.

<https://doi.org/10.1101/gad.276477.115>

Mykowska, A., Sobczak, K., Wojciechowska, M., Kozłowski, P., & Krzyzosiak, W. J. (2011). CAG

repeats mimic CUG repeats in the misregulation of alternative splicing. *Nucleic Acids Research*, 39(20), 8938–8951. <https://doi.org/10.1093/nar/gkr608>

Nakamori, M., Sobczak, K., Puwanant, A., Eichinger, K., Pandya, S., Dekdebrun, J., Heatwole, R.,

Mcdermott, M. P., Chen, T., Cline, M., Tawil, R., Osborne, R. J., Wheeler, T. M., Iii, R. T. M., &

Thornton, C. A. (2013). Splicing biomarkers of disease severity in myotonic dystrophy. *Annals of Neurology*, 74(6), 862–872. <https://doi.org/10.1002/ana.23992>.Splicing

Nedelsky, N. B., & Taylor, J. P. (2019). Bridging biophysics and neurology: aberrant phase transitions

in neurodegenerative disease. *Nature Reviews Neurology*, 15(5), 272–286.

<https://doi.org/10.1038/s41582-019-0157-5>

- Neumann, M., Sampathu, D. M., Kwong, L. K., Truax, A. C., Micsenyi, M. C., Chou, T. T., Bruce, J., Schuck, T., Grossman, M., Clark, C. M., McCluskey, L. F., Miller, B. L., Masliah, E., Mackenzie, I. R., Feldman, H., Feiden, W., Kretzschmar, H. A., Trojanowski, J. Q., & Lee, V. M. Y. (2006). Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science*, 314(5796), 130–133. <https://doi.org/10.1126/science.1134108>
- Nishiyama, T., Yamamoto, H., Uchiumi, T., & Nakashima, N. (2007). Eukaryotic ribosomal protein RPS25 interacts with the conserved loop region in a dicistroviral intergenic internal ribosome entry site. *Nucleic Acids Research*, 35(5), 1514–1521. <https://doi.org/10.1093/nar/gkl1121>
- Nolin, S. L., Glicksman, A., Ersalesi, N., Dobkin, C., Brown, W. T., Cao, R., Blatt, E., Sah, S., Latham, G. J., & Hadd, A. G. (2015). Fragile X full mutation expansions are inhibited by one or more AGG interruptions in premutation carriers. *Genetics in Medicine*, 17(5), 358–364. <https://doi.org/10.1038/gim.2014.106>
- Pakravan, D., Michiels, E., Bratek-Skicki, A., De Decker, M., Van Lindt, J., Alsteens, D., Derclaye, S., Van Damme, P., Schymkowitz, J., Rousseau, F., Tompa, P., & Van Den Bosch, L. (2021). Liquid–Liquid Phase Separation Enhances TDP-43 LCD Aggregation but Delays Seeded Aggregation. *Biomolecules* 2021, Vol. 11, Page 548, 11(4), 548. <https://doi.org/10.3390/BIOM11040548>
- Patel, A., Lee, H. O., Jawerth, L., Maharana, S., Janel, M., Hein, M. Y., Stoyanov, S., Mahamid, J., Saha, S., Franzmann, T. M., Pozniakovski, A., Poser, I., Maghelli, N., Royer, L. A., Weigert, M., Myers, E. W., Grill, S., Drechsel, D., Hyman, A. A., & Alberti, S. (2015). A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell*, 162(5), 1066–1077. <https://doi.org/10.1016/J.CELL.2015.07.047/ATTACHMENT/97D8ABC3-F96E-46AA-AB27-AF35968CD1ED/MMC8.MP4>
- Paul, S., Dansithong, W., Jog, S. P., Holt, I., Mittal, S., Brook, J. D., Morris, G. E., Comai, L., & Reddy, S. (2011). Expanded CUG Repeats Dysregulate RNA Splicing by Altering the Stoichiometry of the Muscleblind 1 Complex. *Journal of Biological Chemistry*, 286(44), 38427–38438. <https://doi.org/10.1074/jbc.M111.255224>
- Paul, S., Dansithong, W., Kim, D., Rossi, J., Webster, N. J. G., Comai, L., & Reddy, S. (2006).

- Interaction of muscleblind, CUG-BP1 and hnRNP H proteins in DM1-associated aberrant IR splicing. *EMBO Journal*, 25(18), 4271–4283. <https://doi.org/10.1038/sj.emboj.7601296>
- Paulson, H. (2018). Repeat expansion diseases. *Handbook of Clinical Neurology*, 147, 105. <https://doi.org/10.1016/B978-0-444-63233-3.00009-9>
- Perdoni, F., Malatesta, M., Cardani, R., Giagnacovo, M., Mancinelli, E., Meola, G., & Pellicciari, C. (2009). RNA/MBNL1-containing foci in myoblast nuclei from patients affected by myotonic dystrophy type 2: An immunocytochemical study. *European Journal of Histochemistry*, 53(3), 151–158. <https://doi.org/10.4081/ejh.2009.e18>
- Perez, B. A., Shorrock, H. K., Banez-Coronel, M., Zu, T., Romano, L. EL, Laboissonniere, L. A., Reid, T., Ikeda, Y., Reddy, K., Gomez, C. M., Bird, T., Ashizawa, T., Schut, L. J., Brusco, A., Berglund, J. A., Hasholt, L. F., Nielsen, J. E., Subramony, S., & Ranum, L. P. (2021). CCG•CGG interruptions in high-penetrance SCA8 families increase RAN translation and protein toxicity. *EMBO Molecular Medicine*. <https://doi.org/10.15252/EMMM.202114095>
- Persichetti, F., Ambrose, C. M., Ge, P., McNeil, S. M., Srinidhi, J., Anderson, M. A., Jenkins, B., Barnes, G. T., Duyao, M. P., & Kanaley, L. (1995). Normal and expanded Huntington's disease gene alleles produce distinguishable proteins due to translation across the CAG repeat. *Molecular Medicine (Cambridge, Mass.)*, 1(4), 374–383. <https://doi.org/10.1007/bf03401575>
- Polymenidou, M. (2018). The RNA face of phase separation. *Science*, 360(6391), 859–860. <https://doi.org/10.1126/science.aat8028>
- Ramesh, N., Daley, E. L., Gleixner, A. M., Mann, J. R., Kour, S., Mawrie, D., Anderson, E. N., Kofler, J., Donnelly, C. J., Kiskinis, E., & Pandey, U. B. (2020). RNA dependent suppression of C9orf72 ALS/FTD associated neurodegeneration by Matrin-3. *Acta Neuropathologica Communications*, 8(1), 1–20. <https://doi.org/10.1186/s40478-020-01060-y>
- Ravel-Chapuis, A., Bélanger, G., Yadava, R. S., Mahadevan, M. S., DesGroseillers, L., Côté, J., & Jasmin, B. J. (2012). The RNA-binding protein Staufen1 is increased in DM1 skeletal muscle and promotes alternative pre-mRNA splicing. *Journal of Cell Biology*, 196(6), 699–712. <https://doi.org/10.1083/jcb.201108113>
- Reddy, K., Tam, M., Bowater, R. P., Barber, M., Tomlinson, M., Nichol Edamura, K., Wang, Y. H., & Pearson, C. E. (2011). Determinants of R-loop formation at convergent bidirectionally transcribed trinucleotide repeats. *Nucleic Acids Research*, 39(5), 1749–1762.

<https://doi.org/10.1093/nar/gkq935>

Rodriguez, C. M., Wright, S. E., Kears, M. G., Haeflner, J. M., Flores, B. N., Liu, Y., Ifrim, M. F., Glineburg, M. R., Krans, A., Jafar-Nejad, P., Sutton, M. A., Bassell, G. J., Parent, J. M., Rigo, F., Barmada, S. J., & Todd, P. K. (2020). A native function for RAN translation and CGG repeats in regulating fragile X protein synthesis. *Nature Neuroscience*, *23*(3), 386–397.

<https://doi.org/10.1038/s41593-020-0590-1>

Rodriguez, S., Sahin, A., Schrank, B. R., Al-Lawati, H., Costantino, I., Benz, E., Fard, D., Albers, A. D., Cao, L., Gomez, A. C., Evans, K., Ratti, E., Cudkowicz, M., Frosch, M. P., Talkowski, M., Sorger, P. K., Hyman, B. T., & Albers, M. W. (2021). Genome-encoded cytoplasmic double-stranded RNAs, found in C9ORF72 ALS-FTD brain, propagate neuronal loss. *Science Translational Medicine*, *13*(601), 1–14. <https://doi.org/10.1126/scitranslmed.aaz4699>

Rohilla, K. J., & Gagnon, K. T. (2017). RNA biology of disease-associated microsatellite repeat expansions. *Acta Neuropathologica Communications*, *5*(1), 63. <https://doi.org/10.1186/s40478-017-0468-y>

Rossi, S., Serrano, A., Gerbino, V., Giorgi, A., Di Francesco, L., Nencini, M., Bozzo, F., Schininà, M. E., Bagni, C., Cestra, G., Carri, M. T., Achsel, T., & Cozzolino, M. (2015). Nuclear accumulation of mRNAs underlies G4C2-repeat-induced translational repression in a cellular model of C9orf72 ALS. *Journal of Cell Science*, *128*(9), 1787–1799. <https://doi.org/10.1242/jcs.165332>

Sabater-Arcis, M., Bargiela, A., Furling, D., & Artero, R. (2020). miR-7 Restores Phenotypes in Myotonic Dystrophy Muscle Cells by Repressing Hyperactivated Autophagy. *Molecular Therapy - Nucleic Acids*, *19*, 278–292. <https://doi.org/10.1016/j.omtn.2019.11.012>

Saberi, S., Stauffer, J. E., Jiang, J., Garcia, S. D., Taylor, A. E., Schulte, D., Ohkubo, T., Schloffman, C. L., Maldonado, M., Baughn, M., Rodriguez, M. J., Pizzo, D., Cleveland, D., & Ravits, J. (2018). Sense-encoded poly-GR dipeptide repeat proteins correlate to neurodegeneration and uniquely co-localize with TDP-43 in dendrites of repeat-expanded C9orf72 amyotrophic lateral sclerosis. *Acta Neuropathologica*, *135*(3), 459–474. <https://doi.org/10.1007/s00401-017-1793-8>

Salcedo-Arellano, M. J., Dufour, B., McLennan, Y., Martinez-Cerdeno, V., & Hagerman, R. (2020). Fragile X Syndrome and associated disorders: clinical aspects and pathology. *Neurobiology of Disease*, *136*, 104740. <https://doi.org/10.1016/J.NBD.2020.104740>

Samaraweera, S. E., O'Keefe, L. V., Price, G. R., Venter, D. J., & Richards, R. I. (2013). Distinct roles

for toll and autophagy pathways in double-stranded RNA toxicity in a drosophila model of expanded repeat neurodegenerative diseases. *Human Molecular Genetics*, 22(14), 2811–2819.  
<https://doi.org/10.1093/hmg/ddt130>

Sanchez-Rangel, E., & Inzucchi, S. E. (2017). Metformin: clinical use in type 2 diabetes. In *Diabetologia* (Vol. 60, Issue 9, pp. 1586–1593). Springer Verlag. <https://doi.org/10.1007/s00125-017-4336-x>

Sareen, D., O'Rourke, J. G., Meera, P., Muhammad, A. K. M. G., Grant, S., Simpkinson, M., Bell, S., Carmona, S., Ornelas, L., Sahabian, A., Gendron, T., Petrucelli, L., Baughn, M., Ravits, J., Harms, M. B., Rigo, F., Frank Bennett, C., Otis, T. S., Svendsen, C. N., & Baloh, R. H. (2013). Targeting RNA foci in iPSC-derived motor neurons from ALS patients with a C9ORF72 repeat expansion. *Science Translational Medicine*, 5(208).  
<https://doi.org/10.1126/scitranslmed.3007529>

Scoles, D. R., & Pulst, S. M. (2018). Oligonucleotide therapeutics in neurodegenerative diseases. In *RNA Biology* (Vol. 15, Issue 6, pp. 707–714). Taylor and Francis Inc.  
<https://doi.org/10.1080/15476286.2018.1454812>

Sellier, C., Buijsen, R. A. M., He, F., Natla, S., Jung, L., Tropel, P., Gaucherot, A., Jacobs, H., Meziane, H., Vincent, A., Champy, M.-F., Sorg, T., Pavlovic, G., Wattenhofer-Donze, M., Birling, M.-C., Oulad-Abdelghani, M., Eberling, P., Ruffenach, F., Joint, M., ... Charlet-Berguerand, N. (2017). Translation of Expanded CGG Repeats into FMRpolyG Is Pathogenic and May Contribute to Fragile X Tremor Ataxia Syndrome. *Neuron*, 93(2), 331–347.  
<https://doi.org/10.1016/j.neuron.2016.12.016>

Sellier, C., Cerro-herreros, E., Blatter, M., Freyermuth, F., Gaucherot, A., Ruffenach, F., Sarkar, P., Puymirat, J., Udd, B., Day, J. W., Meola, G., Bassez, G., Fujimura, H., Takahashi, M. P., Schoser, B., & Furling, D. (2018). rbFOX1/MBNL1 competition for CCUG RNA repeats binding contributes to myotonic dystrophy type 1/type 2 differences. *Nature Communications*, 1–15.  
<https://doi.org/10.1038/s41467-018-04370-x>

Sellier, C., Freyermuth, F., Tabet, R., Tran, T., He, F., Ruffenach, F., Alunni, V., Moine, H., Thibault, C., Page, A., Tassone, F., Willemsen, R., Disney, M. D., Hagerman, P. J., Todd, P. K., & Charlet-Berguerand, N. (2013). Sequestration of DROSHA and DGCR8 by expanded CGG RNA Repeats Alters microRNA processing in fragile X-associated tremor/ataxia syndrome. *Cell*

*Reports*, 3(3), 869–880. <https://doi.org/10.1016/j.celrep.2013.02.004>

- Sellier, C., Rau, F., Liu, Y., Tassone, F., Hukema, R. K., Gattoni, R., Schneider, A., Richard, S., Willemsen, R., Elliott, D. J., Hagerman, P. J., & Charlet-Berguerand, N. (2010). Sam68 sequestration and partial loss of function are associated with splicing alterations in FXTAS patients. *EMBO Journal*, 29(7), 1248–1261. <https://doi.org/10.1038/emboj.2010.21>
- Shi, K. Y., Mori, E., Nizami, Z. F., Lin, Y., Kato, M., Xiang, S., Wu, L. C., Ding, M., Yu, Y., Gall, J. G., & McKnight, S. L. (2017). Toxic PRn poly-dipeptides encoded by the C9orf72 repeat expansion block nuclear import and export. *Proceedings of the National Academy of Sciences of the United States of America*, 114(7), E1111–E1117. <https://doi.org/10.1073/pnas.1620293114>
- Shi, Y., Yang, Y., Hoang, B., Bardeleben, C., Holmes, B., Gera, J., & Lichtenstein, A. (2016). Therapeutic potential of targeting IRES-dependent c-myc translation in multiple myeloma cells during ER stress. *Oncogene*, 35(8), 1015–1024. <https://doi.org/10.1038/onc.2015.156>
- Shin, Y., & Brangwynne, C. P. (2017). Liquid phase condensation in cell physiology and disease. In *Science* (Vol. 357, Issue 6357). American Association for the Advancement of Science. <https://doi.org/10.1126/science.aaf4382>
- Singh, C. R., Glineburg, M. R., Moore, C., Tani, N., Jaiswal, R., Zou, Y., Aube, E., Gillaspie, S., Thornton, M., Cecil, A., Hilgers, M., Takasu, A., Asano, I., Asano, M., Escalante, C. R., Nakamura, A., Todd, P. K., & Asano, K. (2021). Human oncoprotein 5MP suppresses general and repeat-associated non-AUG translation via eIF3 by a common mechanism. *Cell Reports*, 36(2), 109376. <https://doi.org/10.1016/j.celrep.2021.109376>
- Sobczak, K., Michlewski, G., De Mezer, M., Kierzek, E., Krol, J., Olejniczak, M., Kierzek, R., & Krzyzosiak, W. J. (2010). Structural diversity of triplet repeat RNAs. *Journal of Biological Chemistry*, 285(17), 12755–12764. <https://doi.org/10.1074/jbc.M109.078790>
- Sofola, O. A., Jin, P., Qin, Y., Duan, R., Liu, H., de Haro, M., Nelson, D. L., & Botas, J. (2007). RNA-Binding Proteins hnRNP A2/B1 and CUGBP1 Suppress Fragile X CGG Premutation Repeat-Induced Neurodegeneration in a Drosophila Model of FXTAS. *Neuron*, 55(4), 565–571. <https://doi.org/10.1016/j.neuron.2007.07.021>
- Sonobe, Y., Aburas, J., Krishnan, G., Fleming, A. C., Ghadge, G., Islam, P., Warren, E. C., Gu, Y., Kankel, M. W., Brown, A. E. X., Kiskinis, E., Gendron, T. F., Gao, F.-B., Roos, R. P., & Kratsios, P. (2021). A C. elegans model of C9orf72-associated ALS/FTD uncovers a conserved role for

eIF2D in RAN translation. *Nature Communications*, 12(1), 6025. <https://doi.org/10.1038/s41467-021-26303-x>

Sonobe, Y., Ghadge, G., Masaki, K., Sandoel, A., Fuchs, E., & Roos, R. P. (2018). Translation of dipeptide repeat proteins from the C9ORF72 expanded repeat is associated with cellular stress. *Neurobiology of Disease*, 116, 155–165. <https://doi.org/10.1016/j.nbd.2018.05.009>

Sun, X., Li, P. P., Zhu, S., Cohen, R., Marque, L. O., Ross, C. A., Pulst, S. M., Chan, H. Y. E., Margolis, R. L., & Rudnicki, D. D. (2015). Nuclear retention of full-length HTT RNA is mediated by splicing factors MBNL1 and U2AF65. *Scientific Reports*, 5. <https://doi.org/10.1038/srep12521>

Swami, M., Hendricks, A. E., Gillis, T., Massood, T., Mysore, J., Myers, R. H., & Wheeler, V. C. (2009). Somatic expansion of the Huntington's disease CAG repeat in the brain is associated with an earlier age of disease onset. *Human Molecular Genetics*, 18(16), 3039–3047.

<https://doi.org/10.1093/hmg/ddp242>

Swinnen, B., Bento-Abreu, A., Gendron, T. F., Boeynaems, S., Bogaert, E., Nuyts, R., Timmers, M., Scheveneels, W., Hersmus, N., Wang, J., Mizielińska, S., Isaacs, A. M., Petrucelli, L., Lemmens, R., Van Damme, P., Van Den Bosch, L., & Robberecht, W. (2018). A zebrafish model for C9orf72 ALS reveals RNA toxicity as a pathogenic mechanism. *Acta Neuropathologica*, 135(3), 427–443. <https://doi.org/10.1007/s00401-017-1796-5>

Sznajder, L. J., Michalak, M., Taylor, K., Cywoniuk, P., Kabza, M., Wojtkowiak-Szlachcic, A., Matłoka, M., Konieczny, P., & Sobczak, K. (2016). Mechanistic determinants of MBNL activity. *Nucleic Acids Research*, 44(21), 10326–10342. <https://doi.org/10.1093/nar/gkw915>

Sznajder, Ł. J., Thomas, J. D., Carrell, E. M., Reid, T., McFarland, K. N., Cleary, J. D., Oliveira, R., Nutter, C. A., Bhatt, K., Sobczak, K., Ashizawa, T., Thornton, C. A., Ranum, L. P. W., & Swanson, M. S. (2018). Intron retention induced by microsatellite expansions as a disease biomarker. *Proceedings of the National Academy of Sciences of the United States of America*, 115(16), 4234–4239. <https://doi.org/10.1073/pnas.1716617115>

Tabet, R., Schaeffer, L., Freyermuth, F., Jambeau, M., Workman, M., Lee, C. Z., Lin, C. C., Jiang, J., Jansen-West, K., Abou-Hamdan, H., Désaubry, L., Gendron, T., Petrucelli, L., Martin, F., & Lagier-Tourenne, C. (2018). CUG initiation and frameshifting enable production of dipeptide repeat proteins from ALS/FTD C9ORF72 transcripts. *Nature Communications*, 9(1), 1–14. <https://doi.org/10.1038/s41467-017-02643-5>

- Tabrizi, S. J., Leavitt, B. R., Landwehrmeyer, G. B., Wild, E. J., Saft, C., Barker, R. A., Blair, N. F., Craufurd, D., Priller, J., Rickards, H., Rosser, A., Kordasiewicz, H. B., Czech, C., Swayze, E. E., Norris, D. A., Baumann, T., Gerlach, I., Schobel, S. A., Paz, E., ... Lane, R. M. (2019). Targeting Huntingtin Expression in Patients with Huntington's Disease. *New England Journal of Medicine*, *380*(24), 2307–2316. <https://doi.org/10.1056/nejmoa1900907>
- Taneja, K. L., McCurrach, M., Schalling, M., Housman, D., & Singer, R. H. (1995). Foci of trinucleotide repeat transcripts in nuclei of myotonic dystrophy cells and tissues. *Journal of Cell Biology*, *128*(6), 995–1002. <https://doi.org/10.1083/jcb.128.6.995>
- Tang, L., Morris, J., Wan, J., Moore, C., Fujita, Y., Gillaspie, S., Aube, E., Nanda, J., Marques, M., Jangal, M., Anderson, A., Cox, C., Hiraishi, H., Dong, L., Saito, H., Singh, C. R., Witcher, M., Topisirovic, I., Qian, S. B., & Asano, K. (2017). Competition between translation initiation factor eIF5 and its mimic protein 5MP determines non-AUG initiation rate genome-wide. *Nucleic Acids Research*, *45*(20), 11941–11953. <https://doi.org/10.1093/nar/gkx808>
- Tassone, F., Beilina, A., Carosi, C., Albertosi, S., Bagni, C., Li, L., Glover, K., Bentley, D., & Hagerman, P. J. (2007). Elevated FMR1 mRNA in premutation carriers is due to increased transcription. *RNA*, *13*(4), 555–562. <https://doi.org/10.1261/rna.280807>
- Tassone, F., Hagerman, R. J., Loesch, D. Z., Lachiewicz, A., Taylor, A. K., & Hagerman, P. J. (2000). Fragile X males with unmethylated, full mutation trinucleotide repeat expansions have elevated levels of FMR1 messenger RNA. *American Journal of Medical Genetics*, *94*(3), 232–236. [https://doi.org/10.1002/1096-8628\(20000918\)94:3<232::AID-AJMG9>3.0.CO;2-H](https://doi.org/10.1002/1096-8628(20000918)94:3<232::AID-AJMG9>3.0.CO;2-H)
- Taylor, K., & Sobczak, K. (2020). Intrinsic regulatory role of RNA structural arrangement in alternative splicing control. *International Journal of Molecular Sciences*, *21*(14), 1–35. <https://doi.org/10.3390/ijms21145161>
- Thomas, M., White, R. L., & Davis, R. W. (1976). Hybridization of RNA to double stranded DNA: Formation of R loops. *Proceedings of the National Academy of Sciences of the United States of America*, *73*(7), 2294–2298. <https://doi.org/10.1073/pnas.73.7.2294>
- Tian, B., White, R. J., Xia, T., Welle, S., Turner, D. H., Mathews, M. B., & Thornton, C. A. (2000). Expanded CUG repeat RNAs form hairpins that activate the double-stranded RNA-dependent protein kinase PKR. *Rna*, *6*(1), 79–87. <https://doi.org/10.1017/S1355838200991544>
- Timchenko, N. A., Cai, Z. J., Welm, A. L., Reddy, S., Ashizawa, T., & Timchenko, L. T. (2001). RNA

- CUG Repeats Sequester CUGBP1 and Alter Protein Levels and Activity of CUGBP1. *Journal of Biological Chemistry*, 276(11), 7820–7826. <https://doi.org/10.1074/jbc.M005960200>
- Todd, P. K., Oh, S. Y., Krans, A., He, F., Sellier, C., Frazer, M., Renoux, A. J., Chen, K. chun, Scaglione, K. M., Basrur, V., Elenitoba-Johnson, K., Vonsattel, J. P., Louis, E. D., Sutton, M. A., Taylor, J. P., Mills, R. E., Charlet-Berguerand, N., & Paulson, H. L. (2013). CGG repeat-associated translation mediates neurodegeneration in fragile X tremor ataxia syndrome. *Neuron*, 78(3), 440–455. <https://doi.org/10.1016/j.neuron.2013.03.026>
- Trang, H., Stanley, S. Y., Thorner, P., Faghfoury, H., Schulze, A., Hawkins, C., Pearson, C. E., & Yoon, G. (2015). Massive CAG repeat expansion and somatic instability in maternally transmitted infantile spinocerebellar ataxia type 7. *JAMA Neurology*, 72(2), 219–223. <https://doi.org/10.1001/jamaneurol.2014.1902>
- Trottier, Y., Biancalana, V., & Mandel, J. L. (1994). Instability of CAG repeats in Huntington's disease: Relation to parental transmission and age of onset. *Journal of Medical Genetics*, 31(5), 377–382. <https://doi.org/10.1136/jmg.31.5.377>
- Tseng, Y.-J., Sandwith, S. N., Green, K. M., Chambers, A. E., Krans, A., Raimer, H. M., Sharlow, M. E., Reisinger, M. A., Richardson, A. E., Routh, E. D., Smaldino, M. A., Wang, Y.-H., Vaughn, J. P., Todd, P. K., & Smaldino, P. J. (2021). The RNA helicase DHX36/G4R1 modulates C9orf72 GGGGCC hexanucleotide repeat-associated translation. *Journal of Biological Chemistry*, 100914. <https://doi.org/10.1016/j.jbc.2021.100914>
- Tsoi, H., Lau, C. K., Lau, K. F., & Chan, H. Y. E. (2011). Perturbation of U2AF65/NXF1-mediated RNA nuclear export enhances RNA toxicity in polyQ diseases. *Human Molecular Genetics*, 20(19), 3787–3797. <https://doi.org/10.1093/hmg/ddr297>
- Tusi, S. K., Nguyen, L., Thangaraju, K., Li, J., Cleary, J. D., Zu, T., & Ranum, L. P. W. (2021). The alternative initiation factor eIF2A plays key role in RAN translation of myotonic dystrophy type 2 CCUG•CAGG repeats. *Human Molecular Genetics*, 30(11), 1020–1029. <https://doi.org/10.1093/hmg/ddab098>
- Uaesoontrachoon, K., Quinn, J. L., Tatem, K. S., Van Der Meulen, J. H., Yu, Q., Phadke, A., Miller, B. K., Gordish-Dressman, H., Ongini, E., Miglietta, D., & Nagaraju, K. (2014). Long-term treatment with naproxcinod significantly improves skeletal and cardiac disease phenotype in the mdx mouse model of dystrophy. *Human Molecular Genetics*, 23(12), 3239–3249.

<https://doi.org/10.1093/hmg/ddu033>

van Eyk, C. L., Samaraweera, S. E., Scott, A., Webber, D. L., Harvey, D. P., Mecinger, O., O'Keefe, L.

V., Cropley, J. E., Young, P., Ho, J., Suter, C., & Richards, R. I. (2019). Non-self mutation:

Double-stranded RNA elicits antiviral pathogenic response in a *Drosophila* model of expanded CAG repeat neurodegenerative diseases. *Human Molecular Genetics*, *28*(18), 3000–3012.

<https://doi.org/10.1093/hmg/ddz096>

Van Treeck, B., Protter, D. S. W., Matheny, T., Khong, A., Link, C. D., & Parker, R. (2018). RNA self-assembly contributes to stress granule formation and defining the stress granule transcriptome.

*Proceedings of the National Academy of Sciences of the United States of America*, *115*(11), 2734–2739. <https://doi.org/10.1073/pnas.1800038115>

Veloso, A., Kirkconnell, K. S., Magnuson, B., Biewen, B., Paulsen, M. T., Wilson, T. E., & Ljungman, M.

(2014). Rate of elongation by RNA polymerase II is associated with specific gene features and epigenetic modifications. *Genome Research*, *24*(6), 896–905.

<https://doi.org/10.1101/gr.171405.113>

Villate, O., Ibarluzea, N., Maortua, H., de la Hoz, A. B., Rodriguez-Revenga, L., Izquierdo-Álvarez, S.,

& Tejada, M. I. (2020). Effect of AGG Interruptions on FMR1 Maternal Transmissions. *Frontiers in Molecular Biosciences*, *7*(July), 1–6. <https://doi.org/10.3389/fmolb.2020.00135>

Wada, T., Takagi, T., Yamaguchi, Y., Ferdous, A., Imai, T., Hirose, S., Sugimoto, S., Yano, K.,

Hartzog, G. A., Winston, F., Buratowski, S., & Handa, H. (1998). DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. *Genes and Development*, *12*(3), 343–356.

<https://doi.org/10.1101/gad.12.3.343>

Wang, S., Latallo, M. J., Zhang, Z., Huang, B., Bobrovnikov, D. G., Dong, D., Livingston, N. M.,

Tjoeng, W., Hayes, L. R., Rothstein, J. D., Ostrow, L. W., Wu, B., & Sun, S. (2021). Nuclear export and translation of circular repeat-containing intronic RNA in C9ORF72-ALS/FTD. *Nature Communications*, *12*(1). <https://doi.org/10.1038/S41467-021-25082-9>

Wek, R. C., Jiang, H.-Y., & Anthony, T. G. (2006). Coping with stress: eIF2 kinases and translational control. *Biochemical Society Transactions*, *34*(1), 7. <https://doi.org/10.1042/bst20060007>

Wenninger, S., Cumming, S. A., Gutschmidt, K., Okkersen, K., Jimenez-Moreno, A. C., Daidj, F.,

Lochmüller, H., Hogarth, F., Knoop, H., Bassez, G., Monckton, D. G., Engelen, B. G. M. van, &

- Schoser, B. (2021). Associations Between Variant Repeat Interruptions and Clinical Outcomes in Myotonic Dystrophy Type 1. *Neurology: Genetics*, 7(2), e572.  
<https://doi.org/10.1212/NXG.0000000000000572>
- Wenzel, S., Schweimer, K., Rösch, P., & Wöhrl, B. M. (2008). The small hSpt4 subunit of the human transcription elongation factor DSIF is a Zn-finger protein with  $\alpha/\beta$  type topology. *Biochemical and Biophysical Research Communications*, 370(3), 414–418.  
<https://doi.org/10.1016/j.bbrc.2008.03.080>
- Westergard, T., McAvoy, K., Russell, K., Wen, X., Pang, Y., Morris, B., Pasinelli, P., Trotti, D., & Haeusler, A. (2019). Repeat-associated non-AUG translation in C9orf72-ALS/FTD is driven by neuronal excitation and stress. *EMBO Molecular Medicine*, 11(2).  
<https://doi.org/10.15252/EMMM.201809423>
- Wheeler, T. M., Sobczak, K., Lueck, J. D., Osborne, R. J., Lin, X., Dirksen, R. T., & Thornton, C. A. (2009). Reversal of RNA dominance by displacement of protein sequestered on triplet repeat RNA. *Science*, 325(5938), 336–339. <https://doi.org/10.1126/science.1173110>
- Wojciechowska, M., & Krzyzosiak, W. J. (2011). Cellular toxicity of expanded RNA repeats: Focus on RNA foci. *Human Molecular Genetics*, 20(19), 3811–3821. <https://doi.org/10.1093/hmg/ddr299>
- Wojtkowiak-Szlachcic, A., Taylor, K., Stepniak-Konieczna, E., Sznajder, L. J., Mykowska, A., Sroka, J., Thornton, C. A., & Sobczak, K. (2015). Short antisense-locked nucleic acids (all-LNAs) correct alternative splicing abnormalities in myotonic dystrophy. *Nucleic Acids Research*, 43(6), 3318–3331. <https://doi.org/10.1093/nar/gkv163>
- Wolozin, B., & Ivanov, P. (2019). Stress granules and neurodegeneration. *Nature Reviews Neuroscience*, 20(11), 649–666. <https://doi.org/10.1038/s41583-019-0222-5>
- Wong, J. J. L., Ritchie, W., Ebner, O. A., Selbach, M., Wong, J. W. H., Huang, Y., Gao, D., Pinello, N., Gonzalez, M., Baidya, K., Thoeng, A., Khoo, T. L., Bailey, C. G., Holst, J., & Rasko, J. E. J. (2013). Orchestrated intron retention regulates normal granulocyte differentiation. *Cell*, 154(3), 583–595. <https://doi.org/10.1016/j.cell.2013.06.052>
- Xu, Z., Poidevin, M., Li, X., Li, Y., Shu, L., Nelson, D. L., Li, H., Hales, C. M., Gearing, M., Wingo, T. S., & Jin, P. (2013). Expanded GGGGCC repeat RNA associated with amyotrophic lateral sclerosis and frontotemporal dementia causes neurodegeneration. *Proceedings of the National Academy of Sciences of the United States of America*, 110(19), 7778–7783.

<https://doi.org/10.1073/pnas.1219643110>

- Yamada, S. B., Gendron, T. F., Niccoli, T., Genuth, N. R., Grosely, R., Shi, Y., Glaria, I., Kramer, N. J., Nakayama, L., Fang, S., Dinger, T. J. I., Thoeng, A., Rocha, G., Barna, M., Puglisi, J. D., Partridge, L., Ichida, J. K., Isaacs, A. M., Petrucelli, L., & Gitler, A. D. (2019). RPS25 is required for efficient RAN translation of C9orf72 and other neurodegenerative disease-associated nucleotide repeats. *Nature Neuroscience*. <https://doi.org/10.1038/s41593-019-0455-7>
- Yang, H., & Hu, H. Y. (2016). Sequestration of cellular interacting partners by protein aggregates: implication in a loss-of-function pathology. *FEBS Journal*, *283*(20), 3705–3717. <https://doi.org/10.1111/febs.13722>
- Yrigollen, C. M., Martorell, L., Durbin-Johnson, B., Naudo, M., Genoves, J., Murgia, A., Polli, R., Zhou, L., Barbouth, D., Rupchock, A., Finucane, B., Latham, G. J., Hadd, A., Berry-Kravis, E., & Tassone, F. (2014). AGG interruptions and maternal age affect FMR1 CGG repeat allele stability during transmission. *Journal of Neurodevelopmental Disorders*, *6*(1). <https://doi.org/10.1186/1866-1955-6-24>
- Yu, Z., Teng, X., & Bonini, N. M. (2011). Triplet repeat-derived siRNAs enhance RNA-mediated toxicity in a drosophila model for myotonic dystrophy. *PLoS Genetics*, *7*(3), 1–11. <https://doi.org/10.1371/journal.pgen.1001340>
- Yuva-Aydemir, Y., Almeida, S., Krishnan, G., Gendron, T. F., & Gao, F. B. (2019). Transcription elongation factor AFF2/FMR2 regulates expression of expanded GGGGCC repeat-containing C9ORF72 allele in ALS/FTD. *Nature Communications*, *10*(1). <https://doi.org/10.1038/s41467-019-13477-8>
- Zhang, F., Bodycombe, N. E., Haskell, K. M., Sun, Y. L., Wang, E. T., Morris, C. A., Jones, L. H., Wood, L. D., & Pletcher, M. T. (2017). A flow cytometry-based screen identifies MBNL1 modulators that rescue splicing defects in myotonic dystrophy type I. *Human Molecular Genetics*, *26*(16), 3056–3068. <https://doi.org/10.1093/hmg/ddx190>
- Zhang, K., Donnelly, C. J., Haeusler, A. R., Grima, J. C., Machamer, J. B., Steinwald, P., Daley, E. L., Miller, S. J., Cunningham, K. M., Vidensky, S., Gupta, S., Thomas, M. A., Hong, I., Chiu, S. L., Haganir, R. L., Ostrow, L. W., Matunis, M. J., Wang, J., Sattler, R., ... Rothstein, J. D. (2015). The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. *Nature*, *525*(7567), 56–61. <https://doi.org/10.1038/nature14973>

- Zhang, N., & Ashizawa, T. (2017). RNA toxicity and foci formation in microsatellite expansion diseases. *Current Opinion in Genetics and Development*, *44*, 17–29.  
<https://doi.org/10.1016/j.physbeh.2017.03.040>
- Zhang, Y., Glineburg, M. R., Basrur, V., Conlon, K., Hall, D. A., & Todd, P. K. (2020). Near-cognate initiation generates FMRpolyG from CGG repeats in Fragile X associated Tremor Ataxia Syndrome. In *bioRxiv* (p. 2020.10.19.345736). bioRxiv.  
<https://doi.org/10.1101/2020.10.19.345736>
- Zhang, Y. J., Gendron, T. F., Ebbert, M. T. W., O'Raw, A. D., Yue, M., Jansen-West, K., Zhang, X., Prudencio, M., Chew, J., Cook, C. N., Daugherty, L. M., Tong, J., Song, Y., Pickles, S. R., Castanedes-Casey, M., Kurti, A., Rademakers, R., Oskarsson, B., Dickson, D. W., ... Petrucelli, L. (2018). Poly(GR) impairs protein translation and stress granule dynamics in C9orf72-associated frontotemporal dementia and amyotrophic lateral sclerosis. *Nature Medicine*, *24*(8), 1136–1142. <https://doi.org/10.1038/s41591-018-0071-1>
- Zhang, Y. J., Gendron, T. F., Grima, J. C., Sasaguri, H., Jansen-West, K., Xu, Y. F., Katzman, R. B., Gass, J., Murray, M. E., Shinohara, M., Lin, W. L., Garrett, A., Stankowski, J. N., Daugherty, L., Tong, J., Perkerson, E. A., Yue, M., Chew, J., Castanedes-Casey, M., ... Petrucelli, L. (2016). C9ORF72 poly(GA) aggregates sequester and impair HR23 and nucleocytoplasmic transport proteins. *Nature Neuroscience*, *19*(5), 668–677. <https://doi.org/10.1038/nn.4272>
- Zhang, Y., Yang, M., Duncan, S., Yang, X., Abdelhamid, M. A. S., Huang, L., Zhang, H., Benfey, P. N., Waller, Z. A. E., & Ding, Y. (2019). G-quadruplex structures trigger RNA phase separation. *Nucleic Acids Research*, *47*(22), 11746–11754. <https://doi.org/10.1093/nar/gkz978>
- Zu, T., Cleary, J. D., Liu, Y., Bañez-Coronel, M., Bubenik, J. L., Ayhan, F., Ashizawa, T., Xia, G., Clark, H. B., Yachnis, A. T., Swanson, M. S., & Ranum, L. P. W. (2017). RAN Translation Regulated by Muscleblind Proteins in Myotonic Dystrophy Type 2. *Neuron*, *95*(6), 1292-1305.e5. <https://doi.org/10.1016/j.neuron.2017.08.039>
- Zu, T., Gibbens, B., Doty, N. S., Gomes-Pereira, M., Huguet, A., Stone, M. D., Margolis, J., Peterson, M., Markowski, T. W., Ingram, M. A. C., Nan, Z., Forster, C., Low, W. C., Schoser, B., Somia, N. V., Clark, H. B., Schmechel, S., Bitterman, P. B., Gourdon, G., ... Ranum, L. P. W. (2011). Non-ATG-initiated translation directed by microsatellite expansions. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(1), 260–265.

<https://doi.org/10.1073/pnas.1013343108>

Zu, T., Guo, S., Bardhi, O., Ryskamp, D. A., Li, J., Tusi, S. K., Engelbrecht, A., Klippel, K., Chakrabarty, P., Nguyen, L., Golde, T. E., Sonenberg, N., & Ranum, L. P. W. (2020). Metformin inhibits RAN translation through PKR pathway and mitigates disease in C9orf72 ALS/FTD mice. *Proceedings of the National Academy of Sciences of the United States of America*, 117(31), 18591–18599. <https://doi.org/10.1073/pnas.2005748117>

Zu, T., Liu, Y., Bañez-Coronel, M., Reid, T., Pletnikova, O., Lewis, J., Miller, T. M., Harms, M. B., Falchook, A. E., Subramony, S. H., Ostrow, L. W., Rothstein, J. D., Troncoso, J. C., & Ranum, L. P. W. (2013). RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. *Proceedings of the National Academy of Sciences of the United States of America*, 110(51). <https://doi.org/10.1073/pnas.1315438110>