

Repeat RNA expansion disorders of the nervous system: post-transcriptional mechanisms and therapeutic strategies

Joshua L. Schwartz , Krysten Leigh Jones & Gene W. Yeo

To cite this article: Joshua L. Schwartz , Krysten Leigh Jones & Gene W. Yeo (2020): Repeat RNA expansion disorders of the nervous system: post-transcriptional mechanisms and therapeutic strategies, *Critical Reviews in Biochemistry and Molecular Biology*, DOI: [10.1080/10409238.2020.1841726](https://doi.org/10.1080/10409238.2020.1841726)

To link to this article: <https://doi.org/10.1080/10409238.2020.1841726>



© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 10 Nov 2020.



Submit your article to this journal [↗](#)



Article views: 147





View related articles [↗](#)



View Crossmark data [↗](#)

Repeat RNA expansion disorders of the nervous system: post-transcriptional mechanisms and therapeutic strategies

Joshua L. Schwartz , Krysten Leigh Jones and Gene W. Yeo 

Department of Cellular and Molecular Medicine, University of California San Diego, La Jolla, CA, USA

ABSTRACT

Dozens of incurable neurological disorders result from expansion of short repeat sequences in both coding and non-coding regions of the transcriptome. Short repeat expansions underlie microsatellite repeat expansion (MRE) disorders including myotonic dystrophy (DM1, CUG_{50-3,500} in *DMPK*; DM2, CCTG_{75-11,000} in *ZNF9*), fragile X tremor ataxia syndrome (FXTAS, CGG₅₀₋₂₀₀ in *FMR1*), spinal bulbar muscular atrophy (SBMA, CAG₄₀₋₅₅ in *AR*), Huntington's disease (HD, CAG₃₆₋₁₂₁ in *HTT*), C9ORF72-amyotrophic lateral sclerosis (ALS)/frontotemporal dementia (FTD and C9-ALS/FTD, GGGGCC in *C9ORF72*), and many others, like ataxias. Recent research has highlighted several mechanisms that may contribute to pathology in this heterogeneous class of neurological MRE disorders – bidirectional transcription, intranuclear RNA foci, and repeat associated non-AUG (RAN) translation – which are the subject of this review. Additionally, many MRE disorders share similar underlying molecular pathologies that have been recently targeted in experimental and pre-clinical contexts. We discuss the therapeutic potential of versatile therapeutic strategies that may selectively target disrupted RNA-based processes and may be readily adaptable for the treatment of multiple MRE disorders. Collectively, the strategies under consideration for treatment of multiple MRE disorders include reducing levels of toxic RNA, preventing RNA foci formation, and eliminating the downstream cellular toxicity associated with peptide repeats produced by RAN translation. While treatments are still lacking for the majority of MRE disorders, several promising therapeutic strategies have emerged and will be evaluated within this review.

ARTICLE HISTORY

Received 25 June 2020
Revised 1 October 2020
Accepted 21 October 2020

KEYWORDS



MRE; RNA; protein; repeat expansion; RNA foci; RAN translation; rCas; ASO

Introduction to MREs in neurological disorders

Microsatellites, also known as simple sequence repeats or short tandem repeats, are interspersed throughout the human genome where they comprise nearly 3% of total sequence (Hannan 2018; Nguyen et al. 2019). These short tandem repeats are highly polymorphic in repeat length among the human population and pathological expansion of several unique repeat sequences in both coding and non-coding regions of the genome has been associated with dozens of neurological disorders (Table 1) (Hannan 2018; Paulson 2018). Nearly all of these clinically distinct microsatellite repeat expansion (MRE) disorders share hallmark disruptions in RNA metabolism that result from synthesis of repeat rich RNA, although emerging evidence is beginning to uncover how repeat sequences may modulate gene transcription, as well (Rohilla and Gagnon 2017). Since pathological repeat expansions of short, often trinucleotide repeat sequences within many different genes

converge upon key cellular processes dysregulated in several MRE disorders, understanding how MRE RNA exacerbates neuropathology should facilitate development of versatile therapeutics to treat multiple MRE disorders (Rohilla and Gagnon 2017). As many MRE disorders share similar post-transcriptional pathologies, it is conceivable that molecular therapies may be readily adaptable for targeting neuropathology in a variety of MRE disorders.

Pathological MRE within multiple genes expressed in neurons can yield diverse pathological consequences that are clinically distinct for each disorder and may affect unique neuronal populations. Nearly all MRE disorders have been linked to transcription of the MRE, as opposed to pathological perturbation of genomic or chromatin landscapes, but hypermethylation and transcriptional silencing of the *FMR1* locus in Fragile X syndrome (FXS) presents an interesting exception associated with neurodevelopmental conditions, such

CONTACT Gene W. Yeo  geneyeo@ucsd.edu  Department of Cellular and Molecular Medicine, University of California San Diego, La Jolla, CA, USA

© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

Table 1. Summary of MRE disorders of the nervous system, highlighting repeat sequence, repeat size threshold, host gene, neuropathology, and clinical presentation.

Disorder	Repeat	Repeat length (healthy/pathological)	Host gene	Neuropathology	Clinical presentation
1. DM1	CTG	5–38/50–1500	DMPK 3' UTR	Brain atrophy, white and gray matter abnormalities	Neuropsychiatric disturbances, cognitive deficits, sleepiness, fatigue, mood disorder, emotion and memory problems
2. DM2	CCTG	<30/75–11,000	ZNF9 Intron 1	Brain atrophy, white and gray matter abnormalities	Cognitive impairment, intellectual disability, sleepiness and fatigue
3. FXTAS	CGG	20–45/55–200	FMR1 5' UTR	Brain atrophy, white matter lesions, cerebellar volume loss, peripheral neuropathy	Ataxia, cognitive decline, parkinsonism, autonomic dysfunction, short term memory loss
4. FXS	CGG	20–45/>200	FMR1 5' UTR	Hippocampal anomalies, enlarged ventricles, thinning of corpus callosum, aberrant pruning of dendritic spines	Epilepsy, intellectual impairment, autism, macroorchidism, long face and ears
5. HD	CAG	10–26/36–121	HTT Exon 1	Degeneration of caudate, putamen, cortex, and substantia nigra	Involuntary movements, cognitive anomalies, psychiatric disturbances, depression
6. HDL2	CTG	6–28/>41	JPH-3 3'UTR	Prominent neuronal atrophy of striatum and cortex	Movement, emotional, and cognitive anomalies
7. C9-ALS/FTD	GGGGCC	2–10/30(?)–	Intron 1	Frontotemporal lobar dysfunction, motor neuron dysfunction	Dementia, cognitive impairment, changes in personality, behavior, mood, language ability
8. SCA3	CAG	<44/52–82	ATXN3 Exon 10	Neuronal atrophy of brain stem, cerebellum, and basal ganglia	Cerebellar ataxia, parkinsonism, peripheral neuropathy
9. SCA8	CTG	15–50/71–1300	KLH1 3'UTR	Cerebellar atrophy	Progressive ataxia
10. SCA10	ATTCT	10–29/800–4500	SCA10, Intron	White matter atrophy and degeneration of gray matter in cerebellum, brain stem, and thalamus	Ataxia, dysarthria, dysphagia, seizures, anxiety
11. SCA12	CAG	<51/>51	PPP2R2B, 5'UTR	Cerebral and/or cerebellar atrophy	Ataxia, seizures, dementia
12. EPM1	CCCCGCCCGCG	2–3/30–75	CSTB, promoter	Ataxia, incoordination, intention tremor, dysarthria, dementia	Ataxia, epilepsy,
13. DRPLA	CAG	7–25/49–88	ATN1, ORF/Exon 5	Atrophy of dentaorubral and pallidolusian system, cerebral white matter and brain stem damage,	Ataxia, epilepsy, cognitive impairment
14. SBMA	CAG	11–24/40–62	AR, ORF	Degeneration of lower motor neurons and muscle atrophy	Gynecomastia, testicular atrophy, androgen insensitivity
15. SCA1	CAG	6–39/40–83	ATXN1, ORF	Cerebellar and brain stem atrophy, particularly Purkinje neurons	Ataxia, dysarthria, hypotonia, dysphagia
16. SCA2	CAG	15–29/34–59	ATXN2, ORF	Purkinje cell loss, decreased neuronal arborization, atrophy of brain stem and substantia nigra	Ataxia, nystagmus, saccadic eye movements, parkinsonism
17. SCA6	CAG	4–16/21–30	CACNA1A	Degeneration of Purkinje and/or granule cells	Ataxia, dysarthria, nystagmus
18. SCA7	CAG	4–34/35–300	ATXN7, ORF	Gliosis and loss of neurons and myelination in cerebellum, inferior olivary, dentate, and pontine nuclei	Ataxia, dysarthria, dysphagia, retinal atrophy, blindness
19. SCA17	CAG	25–44/45–66	TBP, ORF	Atrophy of striatum and cerebellum	Ataxia, dementia, involuntary movements, and dystonia
20. FRAXE MR	CCG	6–25/> 200	AFF2, 5'?		Mild intellectual disability, speech delay, hyperactivity
21. FRA12A MR	CGG	6–23/?	DIP2B, 5' UTR		Seizures, intellectual disability
22. SCA31	TGGAA	0 ?/>110	BEAN/TK2, Intron	Moderate cerebellar atrophy with Purkinje degeneration and dendritic abnormality	Ataxia, dysarthria, hypotonia, nystagmus
23. SCA36	GGCCTG	5–14/650–2500	NOP56, Intron 1	Cerebellar atrophy, loss of Purkinje cells	Ataxia, tongue fasciculations, nystagmus, hyperreflexia
24. FECD	CTG	10–37/>50	TCF4, Intron	Degeneration of corneal endothelium, deposition of extracellular matrix in cornea	Impaired vision from guttae
25. FRDA	GAA	8–33/>90	FXN, Intron	Peripheral neuropathy, atrophy of cervical spinal cord and cerebellum	Ataxia, hearing loss, dysarthria, muscle weakness
26. FRA7A	CGG	5–22/>85	ZNF713, Intron		Autism spectrum disorder

as autism and intellectual disability (Shin et al. 2009; Todd and Paulson 2010; Hagerman et al. 2017; Rohilla and Gagnon 2017; Misra et al. 2018; Swinnen et al. 2020). Indeed, expanded CGG repeat tracts in *FMR1* RNA may hybridize to its genomic locus forming

RNA:DNA duplexes that promote epigenetic silencing through polycomb group complexes around week 11 of human gestation (Colak et al. 2014; Kumari and Usdin 2014). Transcribing repeat rich sequences may give rise to numerous histopathological mechanisms,

and two additional major drivers of disease pathogenesis across several MRE disorders include repeat-associated non-AUG (RAN) translation and formation of intranuclear RNA foci that primarily sequester key RNA binding proteins (RBPs) from endogenous gene regulatory functions (Lin et al. 2010; Koole et al. 2014; Schmidt and Pearson 2016; Rohilla and Gagnon 2017; Swinnen et al. 2020). Expansion of endogenous polyglutamine-tracts within protein coding sequences also contributes to neuropathologies that share similarities to those seen following toxic RAN translation, but polyglutamine expansions are inherently more limited by underlying sequence constraints than the sequence diversity that enables RAN translation. While RNA repeats may be invariably toxic to multiple cell types, several studies have highlighted the selective vulnerability of neurons to RNA repeats, which likely underlies cognitive, behavioral, and motor symptoms in neurological MRE disorders (Wenzel et al. 2010; Ariza et al. 2015; Bavassano et al. 2017; Jimenez-Sanchez et al. 2017; Selvaraj et al. 2018).

Indeed, while somatic mosaicism and genetic anticipation account for differences in the precise number of repeating sequence units present in any given patient cell, the selective neuronal vulnerability to MREs is hypothesized to emerge from neurons' highly complex morphologies with unique activity-dependent and developmental requirements for spatiotemporally restricted changes in gene expression (McMurray 2010; Roselli and Caroni 2015; Fu et al. 2018; Misra et al. 2018; Nussbacher et al. 2019). Disruptions to homeostatic controls of neuronal gene expression in response to age, stress, pathological repeat length, or environmental changes may underlie the aberrant executive and cognitive dysfunction present in patients with MRE disorders. Consistent with this hypothesis, numerous *in vitro* and *in vivo* experiments have shown that repeat rich transcript accumulation positively correlates with time and underlying repeat unit length (Todd and Paulson 2010; Nelson et al. 2013; Gendron and Petrucelli 2018). These two factors strikingly influence age of disease onset and severity across several different MRE disorders, although not all, underscoring the need to develop therapies for those genetically identifiable patient populations of such disorders (Haeusler et al. 2016; Paulson 2018).

Although researchers have made significant advances in understanding the molecular underpinnings of neuropathology in MRE disorders, translation of these insights into therapies for patients suffering from MRE disorders is lagging (Nussbacher et al. 2019). Pathological MRE within many neuronal genes yields

diverse pathological consequences that are clinically distinct for each individual disorder and may affect different neuronal populations. RAN translation or RNA foci formation are hallmarks of many MRE disorders, yet, upon examination, often with more sensitive tools or reagents, many MRE disorders display signs of both RAN translation and RNA foci formation (Cleary and Ranum 2014). Given that similar molecular and cellular pathologies have been observed to underlie several MRE disorders, developing therapies to eliminate repeat RNA, block RNA foci formation, or prevent RAN translation may have widespread applicability for the treatment of multiple MRE disorders (Rohilla and Gagnon 2017). Select therapeutic strategies that have been considered here include eliminating toxic RNA species, masking toxicity of repeat RNA, and blocking RAN translation-linked toxicity. These strategies have been tested with a variety of agents, such as antisense oligonucleotides, transcription-blocking Cas9, RNA-targeting Cas fusion proteins, engineered RNA binding proteins, and small molecules, which will be discussed in subsequent sections of this review.

Mechanisms underlying MRE disorders of the nervous system

With the advent of next-generation genetic sequencing and the development of animal and cellular models of neurological disorders, it is now clear that impairments to neuronal RNA metabolism underlie numerous unique neuropathologies (Maziuk et al. 2017; Nussbacher et al. 2019). Indeed, widespread dysregulation of RNA metabolism has been observed in several neurodegenerative and neurodevelopmental disorders, highlighting the fundamental importance of homeostatic control of neuronal gene expression for cognition. The focus of this section comprises known and emerging roles of dysregulated RNA metabolism driving pathology in MRE disorders of the nervous system. Multiple, non-exclusive pathological mechanisms contribute to MRE disorders and a single MRE disorder may result from several unique disruptions to RNA biology. A summary of the repeat lengths, genetic basis, neuropathology, and clinical presentation associated with each of the MREs discussed within this review is provided in Table 1. In reviewing prominent mechanisms underpinning neurological MREs, we first highlight the RNA-based processes that are the most well-characterized (i.e. bidirectional transcription, RNA foci formation, and RAN translation), and then discuss emerging post-transcriptional mechanisms of neuropathology, such as RNA phase transitions and disruption of nucleocytoplasmic

shuttling. Many of these mechanisms are intricately linked with one another and can co-occur within a single cell or potentially even different cells of a patient. Rarely would a single post-transcriptional mechanism underlie the entire phenotypic presentation of a neurological MRE disorder, and often, multiple pathological mechanisms contribute to a given disease phenotype.

Known post-transcriptional mechanisms underlying MRE disorders

Synthesis of repeat rich RNA is associated with a diverse array of dominantly inherited neuropathologies (Rohilla and Gagnon 2017; Misra et al. 2018; Swinnen et al. 2020). The most well-known RNA-based mechanisms underlying neuropathology in a variety of genetically diverse MREs include those related to bidirectional

transcription (BT), to sequestration of proteins contained within repeat rich RNA foci, and to toxicity of RAN translation products (Figure 1). These three pathological mechanisms we highlight occur across diverse compositions of MRE sequences, while other neuropathological mechanisms, such as in-frame CAG-mediated polyglutamine expansions within protein coding regions, have more specific sequence context constraints that drive pathology and have been thoroughly reviewed elsewhere (Lieberman et al. 2019). Intriguingly, emerging research is uncovering evidence that bidirectionally transcribed CAG repeats can serve as substrates for intranuclear protein sequestration and RAN translation in a variety of MRE disorders also characterized by polyglutamine protein expansion (Cleary et al. 2018; Nguyen et al. 2019). Indeed, each MRE disorder is characterized by a unique combination of

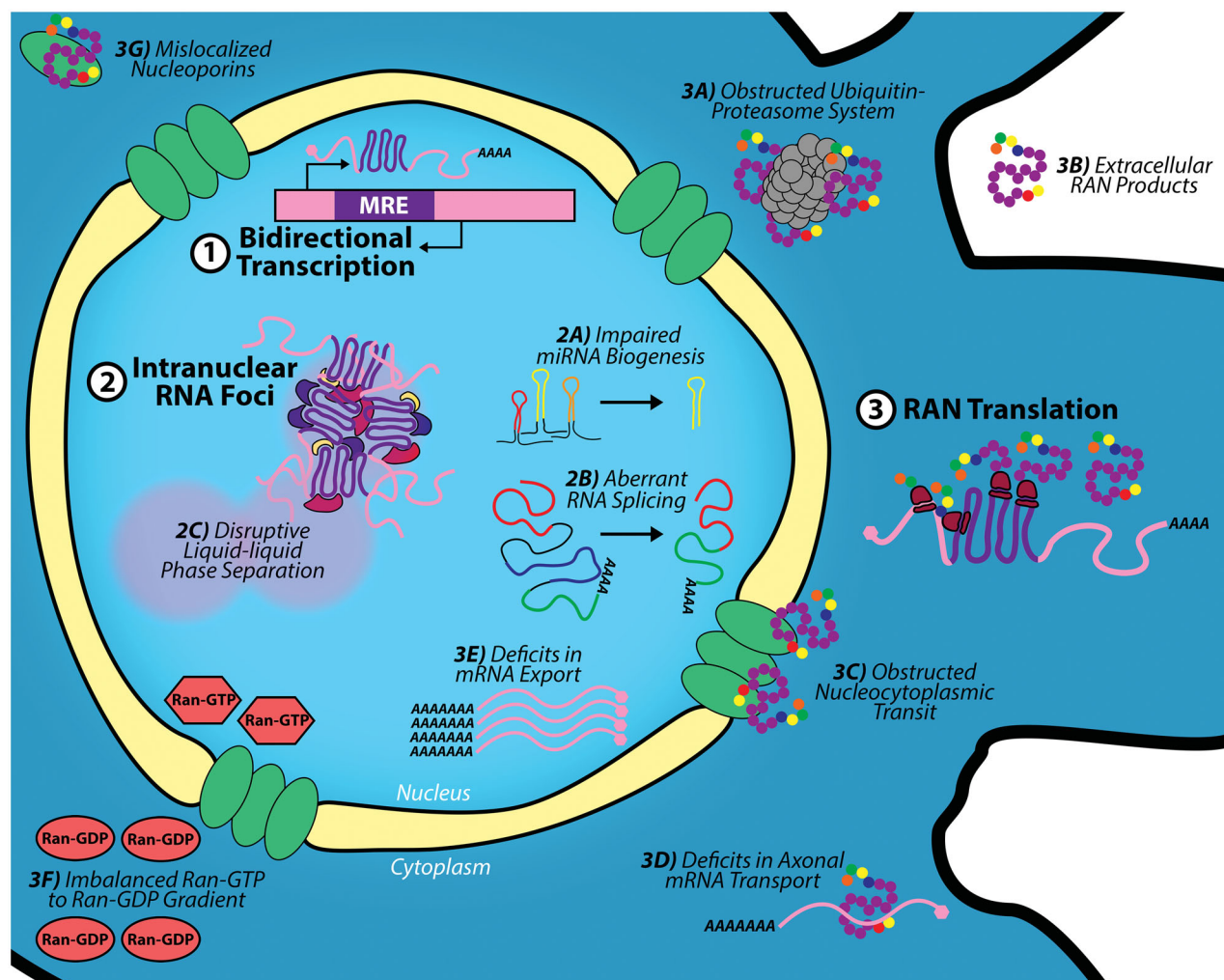


Figure 1. Schematic of a representative neuron, displaying key neuropathologies associated with MRE disorders. Starting in the nucleus, (1) bidirectional transcription initiates a cascade of RNA-based pathologies, including (2) intranuclear RNA foci, which can impair (2A) miRNA biogenesis, (2B) mRNA splicing, (2C) and phase separation, as well as (3) RAN translation, the products of which can impair (3A) ubiquitin-proteasome system, (3B) extracellular environments, (3C) nucleocytoplasmic transit, (3D) axonal mRNA transport, (3E) mRNA export, (3F) Ran gradients, and (3G) nucleoporin localization.

repeat RNA sequence, pathological repeat unit length threshold, host gene, transcript levels, post-transcriptional processing events (e.g. splicing, editing), and cell types that express the repeat RNA, which is summarized in Table 1. While these distinctions selectively influence disease pathogenesis, the well-characterized mechanisms reviewed in this section – BT, sequestration of key proteins within RNA foci, and disruption of cellular functions by RAN translation products – highlight the molecular underpinnings of disease pathogenesis for many clinically heterogeneous MRE disorders of the nervous system.

Historically, spinal bulbar muscular atrophy (SBMA) and FXS were among the first identified MRE disorders their molecular characterizations drove much of our understanding of pathological MRE (La Spada et al. 1991; Arnold and Merry 2019). Both of the host genes underlying SBMA and FXS, *AR* (androgen receptor) and *FMR1*, respectively, are located on the X chromosome, providing early clues as to mechanisms of inheritance. As researchers began to uncover the molecular basis of these disorders, it emerged that the repeats underlying these two disorders are located within two different intragenic contexts. Translation of CAG trinucleotide expansions within the coding sequence of the *AR* gene yields toxic polyglutamine tracts within androgen receptors, a characteristic of SBMA. This mechanism contrasts with the numerous disease mechanisms reported to arise from MRE of CCG within the 5' UTR of *FMR1*, a gene encoding a critical regulator of neuronal development, FMRP. Indeed, depending upon the precise length of genomic CAG:CCG expansions in the *FMR1* promoter region, two clinically distinct neuropathologies can result characteristic of either FXS or fragile X tremor/ataxia syndrome (FXTAS) (Hagerman 2013; Hagerman et al. 2017). Since the discovery of the *AR* and *FMRP* genes and their pathogenic repeats, expansions in dozens of disease causing intragenic repeats have been identified (Hannan 2018).

Collectively, the repeat sequences that give rise to neurological MREs are diverse in nucleotide composition and intragenic location. One striking bias, however, is the preponderance of trinucleotide (Rodriguez et al 2020) repeat expansions compared to other short RNA repeats, such as di-, tetra-, penta-, and hexanucleotide sequences (McMurray 2010; Shimada et al. 2016; Paulson 2018). Another general feature of MRE sequences includes the prevalence of C and G, compared to A or U, and the ability of MRE RNA to form higher order A-form like duplexes, alternative secondary structures, or even G-quadruplexes (Fratta et al. 2012; Cammas and Millevoi 2017; Hale et al. 2019). While GC-

rich sequences are suspected to contribute to pathological higher order secondary structures, the precise significance of trinucleotide sequences, particularly those outside of canonical coding sequences, remains largely unknown (Lin et al. 2010; Cammas and Millevoi 2017; Rohilla and Gagnon 2017). Despite the heterogeneity in repeat sequence composition, location, and length, similar pathological mechanisms can emerge from expansions of endogenously short repeats within neuronal genes (Rohilla and Gagnon 2017; Zhang and Ashizawa 2017; Hannan 2018; Misra et al. 2018; Paulson 2018; Swinnen et al. 2020).

Another recognizable feature of neurological MREs is the generally increased lengths of pathogenic repeat tracts within UTRs (e.g. >1000 in DM1, SCA10, or SCA36) compared to those within coding sequences (e.g. 35–80 in many polyQ disorders), which may be a consequence of evolutionary pressures to maintain functional open reading frames of host genes key to neuronal homeostasis (Shimada et al. 2016; Paulson 2018). Moreover, the length of repeat expansions often inversely correlates with disease severity and onset, especially for polyQ disorders, but varies based on several factors (Orr and Zoghbi 2007; Shimada et al. 2016). Indeed, interruptions to repeat sequences, such as by single nucleotide insertions or substitutions within MREs, can significantly modify severity and onset of several MRE disorders (Yrigollen et al. 2012; McFarland et al. 2013; Cumming et al. 2018). Ultimately, the clinical presentation for each neurological MRE disorder can be influenced by multiple neuropathological factors unique to each individual MRE, including the host gene, repeat sequence, repeat length threshold, and level of intracellular transcript expression, among other well-characterized, non-exclusive mechanisms that are prevalent across other neurological MREs characterized by diverse repeat sequence compositions (Rohilla and Gagnon 2017; Paulson 2018).

Bidirectional transcription

A previously overlooked phenomenon that is now recognized to occur in at least 20 MRE disorders is bidirectional transcription (BT) of the repeat locus, whereby both the canonical repeat and its antisense (AS) repeat are co-expressed (Budworth and McMurray 2013). The earliest reports of AS transcription in MRE disorders were in DM1 (Cho et al. 2005) and SCA8 (Moseley et al. 2006), but BT has since been observed in FXTAS, HD, HDL2, SCA7, and C9 (Cho et al. 2005; Moseley et al. 2006; Cleary et al. 2018; Nguyen et al. 2019). The precise contributions of AS transcripts to pathology is still

under active investigation, as many AS transcripts are often produced at lower levels than their sense counterparts (Moseley et al. 2006; Budworth and McMurray 2013). Indeed, the creation of AS transcripts may occur across the majority of the human genome (Barman et al. 2019). AS transcription of MRE loci may also regulate sense gene expression by a variety of mechanisms at the epigenetic, transcriptional, and post-transcriptional levels. Consequently, understanding the interplay between repeat lengths and transcriptional or translational output would be an exciting area of further exploration that may uncover therapeutic targets of broad relevance to MRE disorders. Within select cells, some AS transcripts are co-expressed with their sense counterparts, such as those from G4C2 expansions within C9orf72, the leading genetic cause of amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (C9-ALS/FTD). This co-expression pattern may highlight a shared role between sense and AS transcripts in disease pathogenesis. In other MRE disorders with bidirectional transcription, for example, the AS transcript spanning the CGG repeat of FMR1, *FMR1-AS*, may undergo differential alternatively splicing in FXTAS premutation carriers compared to neurotypical controls (Ladd et al. 2007). This differential splicing suggests a potential post-transcriptional contribution to FXTAS severity that remains to be thoroughly explored. Indeed, *FMR1-AS* levels are elevated in premutation carriers, compared to controls, and since *FMR1-AS* is polyadenylated, exported to cytoplasm, and contains a putative ORF encoding a polyproline stretch (Ladd et al. 2007), more research is needed to completely understand the roles of naturally occurring AS transcripts and how alterations in AS RNA processing may contribute to disease phenotypes.

RNA foci

Transcripts of repeat rich RNA can form dynamic aggregates that disrupt cellular function, often on a multisystem level, by sequestering RBPs and other essential proteins from native locations (Zhang and Ashizawa 2017). These RNA foci have been observed in a variety of MRE disorders, including FXTAS, C9orf72-ALS/FTD, DM1, DM2, FECD, HD, HDL2, SBMA, SCA-3, -8, -10, -31, and -36 (Rohilla and Gagnon 2017; Gendron and Petrucelli 2018; Ishikawa and Nagai 2019; Matthaei et al. 2019; Nussbacher et al. 2019; Paulson 2018; Swinnen et al. 2020). The specific size, number, subcellular localization, and composition of pathogenic ribonucleoprotein foci may vary across and within MRE disorders, likely contributing to varying degrees of

impairment to neuronal RNA processing, splicing, transport, or translation, gene regulatory processes critical for healthy neuronal and cognitive function (Markmiller et al. 2018). For example, in cerebellar tissue from SCA8 patients, single large CUG intranuclear foci were often found in interneurons and Bergmann glia, while multiple small nuclear CUG foci were found in Purkinje cells, suggesting differences in assembly and distribution of these types of CUG inclusions across cell types of the brain (Daughters et al. 2009). CUG sense foci associated with *DMPK* from DM1 patients, however, have been found within both the nucleus and cytoplasm, as were antisense transcripts *DMPKAS*, although it's unclear cytoplasmic AS foci impart toxicity (Zhang and Ashizawa 2017). As multiple MRE disorders are characterized by RNA foci, a challenge remains in identifying foci that contribute to disease compared to foci that may provide adaptive benefits, such as in response to stress or aging.

Following transcription, several repeat-expanded RNA transcripts are retained within the nucleus, where they colocalize with interacting RBPs and RNAs to form microscopic inclusions. Neurons of individuals with healthy repeat lengths typically produce repeat RNA that is exported to the cytoplasm for proper cellular function if present within an exon or appropriately spliced if the repeat is intronic. These key RNA metabolic processes are often disrupted in repeat expansion disorders, whereby genomic repeats may augment transcript accumulation or saturate the endogenous RNA splicing, miRNA biogenesis, or nuclear export machinery responsible for translocating repeat RNA to the cytoplasm (Figure 1). By exogenously expressing repeat sequences of various lengths within native or artificial gene contexts and then performing RNA fluorescence *in situ* hybridization (FISH), multiple labs identified sequence-specific repeat length thresholds underlying RNA foci formation (Urbanek and Krzyzosiak 2016; Paulson 2018). These findings largely supported observations based on tissue biopsies from patients with repeat expansions and permitted mechanistic investigations of RNA foci formation, dynamics, diversity, and toxicity. For example, following transfection of plasmids expressing variable length CGG repeats linked to FXTAS, foci formation was absent in the subthreshold repeat length contexts, was weak in the smallest expansion length condition, and was highly pronounced in the pre-mutation range (Sellier et al. 2010). Similarly, when researchers expressed transcripts with ALS- and FTD-linked G4C2 of either 8, 38, or 72 repeats, foci and neurotoxicity were only observed at pathologically expanded repeat lengths (Lee et al. 2013).

Additionally, by exploiting the temporal control of exogenous repeat RNA expression, researchers could compare percent colocalization of multiple RBPs over time, and identified Sam68 as an early interactor within CGG foci, followed subsequently by MBNL1 and hnRNP-G, for example, Sellier et al. (2010). Sam68 is a primarily nuclear RBP that modulates RNA 3' processing and alternative splicing, but hnRNP-G and MBNL1 also regulate processing and splicing of distinct RNA populations. Moreover, comparison of Sam68 colocalization with additional exogenously expressed MRE-linked repeat sequences revealed a striking specificity of Sam68 for CGG repeats (Sellier et al. 2010), suggesting sequence-specific biophysical properties underlying foci formation and dynamics across MRE disorders.

Furthermore, numerous groups have attempted to identify compositions of RNA foci interactomes by a variety of approaches (Paul et al. 2011; Ishiguro et al. 2017). Early attempts to uncover the composition of intranuclear foci in FXTAS relied upon fluorescent activated sorting to isolate endogenous ubiquitin and crystallin positive inclusions. Isolating inclusions from FXTAS patient brain tissue for proteomic analysis, although crude and difficult to rigorously control, revealed enrichment of RBPs that may affect critical gene regulatory processes in neurons (Iwahashi et al. 2006). These initial studies have subsequently been largely supported by more sensitive, higher throughput, and well controlled experiments (Iwahashi et al. 2006; Ma et al. 2019). For example, by flowing mouse brain lysate over a column of variable length CGG repeats, researchers found proteins enriched in the pathogenic length contexts include regulators of miRNA biogenesis, which bind, but do not cleave, the CGG repeat RNA hairpins (Sellier et al. 2013). Complementary strategies such as RNA pulldowns followed by mass spectrometry have been undertaken to identify components of G4C2 foci in C9ORF72-ALS/FTD or CUG foci in DM1, HDL2, SCA8, and others (Ishiguro et al. 2017; Zhang and Ashizawa 2017). Intriguingly, *in vitro* studies have identified MBNL1 as an RBP enriched within CUG RNA foci across various MRE disorders, suggesting critical roles for MBNL1 in maintaining neuronal health and homeostasis (Mankodi et al. 2001; Fardaei et al. 2002; Rudnicki et al. 2007; Daughters et al. 2009). While splicing factors like MBNL1 have been independently validated by different groups to be a major constituent of DM1 foci, for example, there have also been notable differences in candidate foci interactomes that have emerged across approaches or between *in vivo* and *in vitro* contexts, highlighting the importance of validating candidate proteins sequestered within ribonuclear foci, through

methods like FISH combined with immunofluorescence on patient-derived models.

Beyond disrupting miRNA biogenesis, MRE-linked RNA foci sequester diverse repertoires of RBPs, including those that establish landscapes of alternative splicing and protein synthesis or are responsible for distributing RNA to distant neuronal locations, such as synapses or axon terminals (Ishiguro et al. 2017). Sequestration of splicing factors within pathogenic repeat RNA foci give rise to what has been called splicingopathies, including DM1, FXTAS, SCA8, C9ORF72-ALS/FTD, and Fuchs endothelial corneal dystrophy (FECD). Some of the splicing factors found within (1) C9 foci include hnRNP-A1/F/H/U, SRSF2, PUR α , SF2, and ADARB2; (2) DM1 foci include MBNL1-3, CUGBP1, and hnRNP-H/F; (3) FXTAS foci include MBNL1, SRSF-1/4/5/6/7/10, hnRNP-A1/A2/B1/A3/C/D/E1/G/M, PUR α , and Sam68; (4) SCA8 foci include MBNL1; and (5) FECD foci include MBNL (Zhang and Ashizawa 2017). Recruitment of diverse repertoires of RBPs into repeat RNA foci raises the possibility that transcriptional cargoes normally associated with these RBPs may also be mislocalized within intranuclear inclusions and contribute to underlying disease pathology. Indeed, mislocalization of MBNL1 within CUG RNA foci disrupts critical splicing events in regulators of neuronal health and plasticity, such as insulin receptor IR2 and chloride channel CLC2 (Pettersson et al. 2015). Researchers hypothesize that disruptions in MBNL1-mediated splicing profiles drive DM1 pathology, but additional cellular toxicities downstream of CUG repeat RNA production have been reported (Botta et al. 2007; Onishi et al. 2008; Perbellini et al. 2011; Batra et al. 2014). Identification of highly penetrant disrupted splicing events associated with RBP sequestration by pathogenic ribonuclear foci in specific MRE disorders will be necessary to develop effective therapeutic strategies that modulate alternative splicing.

Differences in foci formation, cell type-specific sequestration of RBPs, subcellular localization, and the composition of inclusions likely contribute to variations in disease symptoms and severity. To this end, the pathogenicity of RNA foci observed in C9orf72 and FXTAS have received considerable attention. While CGG repeat RNA in FXTAS form intranuclear RNA inclusions that appear to drive disease pathology, intranuclear G4C2 repeat RNA foci have been proposed to impart an adaptive response. Indeed, cytoplasmically localized G4C2 repeat RNA foci can be localized to neurites, where it triggers branching defects and disrupts transit of RNA transport granules to distant subcellular depots, such as dendritic spines (Burguete et al. 2015). For

example, in rat primary neurons transfected with expanded G4C2 repeat plasmids, neuritic G4C2 foci colocalized with FMRP, an observation associated with increased synaptic protein levels of PSD95, a FMRP target gene (Burguete et al. 2015). Since dendritic spines represent some of the earliest sites of atrophy across various neurodegenerative disorders, sequestration of RBPs that establish neuronal polarity or subcellular RNA distribution within ribonuclear or cytoplasmic foci may have profound consequences for neuronal health.

RAN translation

Another post-transcriptional mechanism suspected to underlie many MRE disorders includes repeat-associated non-AUG (RAN) translation (Cleary et al. 2018; Banez-Coronel and Ranum 2019; Nguyen et al. 2019; Nussbacher et al. 2019). First observed in degenerating Purkinje neurons of cerebellar biopsies from SCA8 patients, RAN translation has been reported in at least nine MRE disorders, including C9-ALS/FTD, DM1, FXTAS, HD, HDL2, and SCA3, among others (Zu et al. 2011; Paulson 2018; Banez-Coronel and Ranum 2019; Swinnen et al. 2020). By initiating translation at non-canonical start codons (e.g. CUG) or by ribosomal slipping, RAN translation can occur in all three ribosomal reading frames and within coding sequences or “untranslated” regions. RAN translation is also commonly observed on antisense transcripts. Based on *in vitro* and *in vivo* data representing several MRE disorders, expression of RAN translation products generally increases as the number of repeats increases. RAN translation products from trinucleotide repeat expansions are predominantly homopolymeric (e.g. polyalanine [polyA], polyG, polyQ, etc.) mono-peptide repeats (MPRs), but hexameric GGG-GCC repeat units found in C9-ALS/FTD encode a variety of dipeptide repeats (DPRs). MRE disorders such as SCA31 and FECD are characterized by even more complex peptide repeat sequences. Each RAN translation product will have unique biophysical and biochemical properties that will influence molecular interactions and sub-cellular distribution. Current data suggest that neurodegenerative MRE disorders with phenotypes most strongly linked to RAN translation include C9-ALS/FTD, HD, FXTAS, and SCA8, but researchers are actively exploring the pathogenicity of RAN translation in other MRE diseases where peptide repeats have been identified and likely underlie select neurological phenotypes as well, such as FECD, DM1/2, and SCA31, for example Banez-Coronel and Ranum (2019) and Nguyen et al. (2019).

RAN translation products are predominantly localized to the cytoplasm or perinuclear spaces where they can form aggregates and sequester biomolecules, in much the same way as repeat rich RNA can recruit RNA and proteins through homotypic and heterotypic interactions, respectively. RAN proteins are found in multiple regions of patient brains, but not always in regions experiencing visible signs of neuronal atrophy, suggesting the possibility that RAN product aggregation may have adaptive, as well as pathological consequences. The hypothesis that RAN production may be an adaptive response was recently supported by the observation that native, endogenous RAN translation of healthy CGG repeat lengths in the FMR1 5'UTR of FMRpolyG inhibits downstream production of FMRP from the canonical AUG codon (Rodriguez et al. 2020). While highly intriguing, it remains to be determined how widespread native RAN translation occurs in healthy repeat lengths and how RAN translation may influence AUG-initiating translation of other host genes.

The underlying mechanisms as well as the pathological consequences of RAN translation have gained considerable attention since their discovery nearly a decade ago, and the mechanistic insights and number of diseases linked to RAN are likely to increase as improvements are made to current detection methods, such as antibodies to mono- and di-peptide repeats. Canonical translation is an elegantly orchestrated process at all steps of protein synthesis from initiation, elongation, and termination, but physiological exceptions to canonical AUG initiation sites exist, including internal ribosomal entry sites and near-cognate start codons, like CUG. Indeed, an upstream CUG may underlie m7G cap- and eIF4a-dependent mechanisms of C9orf72 RAN translation (Green et al. 2017). While near-cognate start codons and ribosomal slippage along repeat tracts may account for multiple homopolymeric translation products in other MRE disorders, a m7G cap- and 40S ribosomal-dependent scanning mechanism appears to underlie RAN translation of CGG repeat transcripts, as well.

Differences in translation initiation across MRE disorders would indicate mechanistic differences in RAN production to be considered when designing therapeutic inhibitors, for example. A screen in yeast for modifiers of toxicity associated with RAN translation of C9ORF-ALS/FTD uncovered a role for RPS25 specifically in RAN translation of expanded disease-linked repeats, such as G4C2 and CAG, but not in canonical AUG-initiated translation. Moreover, when expression of RPS25 was reduced either in a drosophila model of C9orf72-expansion or in human model of patient-derived motor

neurons, survival was extended *in vivo* and *in vitro*, respectively (Yamada et al. 2019). Notably, the increased survival phenotypes occur independent of reductions in intranuclear C9 RNA foci, suggesting a therapeutic benefit solely from reducing intranuclear RAN translation products, such as poly(GR) and poly(PR) (Yamada et al. 2019). This observation also underscores the potential adaptive roles intranuclear sequestration of G4C2 RNA may have in preventing cytoplasmic translation and toxic accumulation of DPRs, a hypothesis that should be explored for other MRE disorders.

As researchers study the pathological consequences of RAN translation, some RAN products (e.g. R-rich ones) have emerged as more toxic than others (e.g. polyGA). In elegant experiments using synthetic RAN translation products or repeat RNA constructs that either (1) lack AUG start codons, (2) contain frequent premature stop codons, (3) contain disease-modifying and clinically-linked repeat-interrupting single nucleotide insertions, (4) contain canonical AUG start codons instead of endogenous near-cognate start codons, or (5) encode RAN products but lack hairpin-forming RNA secondary structures, researchers have made tremendous insights into the neuropathological consequences of RAN translation, both dependent and independent of intranuclear RNA foci formation. For example, *in vitro* expression of DMPK1-linked CUG repeat expansions in the absence of an upstream AUG start codon results in production of polyL, polyC, and polyA. Moreover, cellular toxicity of repeat RNA has been augmented in a number of studies by mutating upstream near-cognate start codons to canonical AUG start codons (Lopez-Gonzalez et al. 2016; Choi et al. 2019), presumably by increasing initiation and ultimately synthesis of RAN products. Although multiple RAN polymers may be synthesized *in vitro*, RAN translation products have so far been identified in patient tissue from C9orf72-ALS/FTD, DM2, HD, FECD, FXTAS, SCA8, and SCA31. The patient-derived RAN products may result from diverse reading frames or AS transcripts and may occur within focalized neuroanatomical deposits or specific cell types.

Some of the most well characterized RAN translation products include DPRs from C9orf72-ALS/FTD. DPRs in C9orf72 are found within clinically significant cytoplasmic inclusions throughout patient brains (Mann et al. 2013) that may reflect sites of stress or elevated neuronal activity (Green et al. 2017). Although accurate quantification of relative DPR abundance remains challenging with existing antibody-based detection methods, DPRs are generally thought to be expressed at lower levels in the spinal cord than the brain, but considerable variation exists in focal deposition of RAN

translation products across patients and within individual brains. In terms of structural and biophysical properties, poly-GA forms relatively uncharged dense inclusions, while poly-PR and -GR, for example, form highly charged and polarized flexible coils that likely disrupt processes critical for cellular health and homeostasis (Freibaum and Taylor 2017). Understanding the individual and synergistic effects of DPR production is an active area of research and may provide clues into the pathogenic basis of many MRE disorders.

Emerging post-transcriptional mechanisms underlying MRE disorders

While significant clinical and preclinical research efforts have advanced our understanding of how widespread bidirectional transcription, RNA foci formation, and RAN translation are across MRE disorders, emerging research has identified additional cellular pathologies downstream of repeat RNA production that may occur in multiple MRE disorders. For example, groups are beginning to uncover biophysical and biochemical properties of pathogenic repeat RNA within intracellular environments, such as liquid-liquid phase separation, and the cellular consequences of such disruptions on processes like stress granule formation and regulation of gene repression, ribosomal biogenesis within nucleoli, and transit of nucleocytoplasmic cargoes. Often, these pathologies only manifest above a critical number of repeat units that is generally comparable to the pathogenic length in MRE patients. As with many of the aforementioned MRE-associated pathologies, several post-transcriptional mechanisms may underlie a given MRE and these mechanisms may be closely interrelated or disrupt similar fundamental cellular processes.

Liquid-liquid phase separation

Aberrant liquid-liquid phase separation (LLPS) has emerged as another pathological mechanism that may arise following synthesis of expanded RNA repeats or peptide repeats. While biomolecular processes within many cellular organelles are compartmentalized from the cytoplasm or nucleoplasm by lipid bilayer membranes, some organelles (e.g. nucleoli, stress granules, etc.) rely upon liquid-liquid phase separation to maintain subcellular organization and integrity. For example, nucleoli are dynamic and highly prominent membraneless structures found within eukaryotic nuclei that provide remarkably organized RNA- and protein-rich environments for ribosome biogenesis, a complex metabolic process enabled by the multivalent

interactions of many proteins and nucleic acids. The formation and organization of membraneless organelles have received considerable recent attention and have provided a foundation to explore LLPS in other contexts, such as disease pathology, where much less is understood about how LLPS is established or what the consequences of disrupted LLPS may be to neuronal function.

Many proteins prone to undergo LLPS, especially those found within stress granules or RNA processing bodies, contain intrinsically disordered regions (IDRs). These include many proteins with previously established roles in neurodegenerative disorders, such as FUS and TDP-43. These low complexity regions typically lack hydrophobic residues and are enriched for polar and charged amino acids, but also those amino acids that enhance molecular interactions, such as through pi-stacking (Gabryelczyk et al. 2019). Common residues found within IDRs include G, S, Q, P, E, K, and R. Repeating dipeptide units like RG, FG, SY, and YG are often found, as well (Brangwynne et al. 2015; Feng et al. 2019). Post-translational modification of specific amino acid side chains, such as by phosphorylation, may also drive LLPS, but one consistent driver of liquid–gel condensation is increased valency or molecular interactions (Aumiller et al. 2016; Brangwynne et al. 2015). Based on a variety of experiments using purified components (e.g. RNA or protein), the molecular interactions that promote LLPS can be augmented by a variety of factors, including increasing local concentrations of substrates, adding repeating units, and changing electrostatic interactions by altering salt concentrations (Brangwynne et al. 2015; Zhang et al. 2015a; Aumiller et al. 2016; Jain and Vale 2017; Choi et al. 2020). These *in vitro* tests revealed a striking level of reversibility of LLPS, but also one that diminished with repeated cycles of mixing and demixing. The impaired disassembly of RBP droplets and their resemblance to neuropathological deposits prompted many to question if these more solid-like structures may have adverse consequences and whether cells have mechanisms in place to promote reversibility of LLPS that prevent pathological aggregate formation.

Nucleocytoplasmic transport

Nuclear membranes are a defining feature of eukaryotes, providing biophysical separation for transcription within the nucleus and translation within the cytoplasm. For this reason, disruptions to nucleocytoplasmic transport (NCT) or nuclear membrane integrity may

have dire consequences to cellular viability. To this end, an increasing number of studies from multiple groups have shown that several neurodegenerative disorders, including those not characterized by MRE, display disruptions to NCT. NCT in vertebrates is a tightly regulated process, coordinated by nuclear pore complexes (NPCs), large macromolecular assemblies composed of hundreds of copies of dozens of low complexity rich nucleoporin (NUP) proteins (Boehringer and Bowser 2018), and a high nuclear to cytoplasmic RanGTP ratio that maintains nuclear export. NUPs are some of the longest lived proteins within mammalian neurons *in vivo*, hinting at possible link between age-related declines in NCT integrity and neuronal survival, and impaired nucleocytoplasmic Ran gradients have been observed in patient cells (Ward et al. 2014; Zhang et al. 2015b). These observations further support the involvement of NCT in neurodegenerative diseases.

Additional evidence linking MRE toxicity to NCT defects was provided by separate genetic screens in drosophila and yeast for modifiers of C9-ALS/FTD-linked toxicity (Freibaum et al. 2015; Jovičić et al. 2015). Both screens revealed striking enrichments for components of NCT. While the precise contribution disruptions to NCT, RanGTP gradients, or nuclear envelopes have on MRE disease pathogenesis remains to be determined, NCT dysfunction may result from toxicities linked to sequestration by repeat RNA or peptide repeats, such as from RAN translation. Thus, while not suspected to be an underlying driver of disease pathogenesis, compromised NCT is predicted to influence advanced disease phenotypes in a variety of MRE disorders.

Disruption of neuronal NCT by repeat RNA has been most well characterized in C9-ALS/FTD, but evidence suggests a basis for RNA-mediated disruption of NCT in FXTAS, as well. A high throughput screen of more than 4000 human ORFs expressed in yeast found several candidate proteins that preferentially bound C9-ALS/FTD-linked G4C2. Comparing fold enrichment of proteins bound to G4C2 compared to a scrambled G4C2 RNA sequence that is also predicted to form G-quadruplexes permitted researchers to make conservative predictions of G4C2 interacting proteins (Hu et al. 2009; Donnelly et al. 2013). RanGAP1, an activator of GTPases, represents one such G4C2-interacting protein with known roles in regulating NCT and which was subsequently found to modify C9-ALS/FTD phenotype when knocked down in drosophila (Zhang et al. 2015b). By pharmacologically inhibiting G4C2-mediated sequestration of RanGAP1 with TMPyP4, a porphyrin compound that disrupts G-quadruplexes, researchers rescued an *in vitro*

nuclear transport phenotype as well as a G4C2-dependent rough eye phenotype in *Drosophila*.

In addition to C9-ALS/FTD, FXTAS represents another MRE disorder reported to have RNA-mediated disruption of nuclear integrity or NCT. While the biomolecular composition of inclusions within FXTAS remain an area of active interest, recent proteomic characterizations of these foci suggest Fmr1 premutation transcripts may colocalize with many proteins that regulate signaling across the nuclear membrane (Tassone et al. 2004). For example, two of the most highly enriched proteins within FXTAS inclusions are SUMO2 and p62, which regulate intracellular signaling between nuclear and cytoplasmic compartments (Chen et al. 2006; Hewitt et al. 2016). Indeed, SUMOylation is well known to regulate NCT (Flotho and Melchior 2013), but also embryogenesis. In one elegant *in vivo* mouse study, overexpression of SUMO2 was shown to rescue survival phenotypes in SUMO1 loss of function mutants (Wang et al. 2014). Moreover, RanGAP1 is a major post-translational target of SUMO1 with well characterized roles in NCT, further highlighting the extent to which disruptions to NCT should be explored in other MRE disorders, and underscoring the potential widespread disruption of RanGAP1, or potentially RanGTP gradients, across MRE disorders.

In terms of peptide repeats that disrupt NCT in MRE disorders, several clinical examples have emerged, including mHTT in HD, FMRpolyG in FXTAS, and PR and GA dipeptides from C9-ALS/FTD. Like RNA-based disruptions of NCT, peptide-based disruptions of NCT frequently result from sequestration of key regulators of NCT or nuclear envelope integrity. Based on human data and evidence from mouse models, intranuclear polyQ-expanded mHTT foci appear to sequester NUP62, nuclear export factor GLE1, THOC2, which is a member of the TREX (transcription/export) complex, and RanGAP1, which likely underlies impaired export of polyA⁺ RNA in patient cells or disrupted nuclear envelope morphology (Gasset-Rosa et al. 2017; Grima et al. 2017). FMRpolyG in FXTAS has also been reported to disrupt morphology of nuclear lamin, as have dipeptide repeats from C9-ALS/FTD. Indeed, heterotypic interactions between intrinsically disordered PR DPRs and the LCD domains of NUPs, primarily FG repeats, may drive impaired morphology of nuclear envelopes in C9-ALS/FTD patient cells (Shi et al. 2017). PA C9-ALS/FTD DPRs may also compromise nuclear integrity by a combination of intranuclear sequestration of RanGAP1, and an otherwise transmembrane component of the nuclear pore, POM121 (Zhang et al. 2016). Given the range of peptide repeats produced by a variety of MRE disorders

and the widespread observation of NCT dysfunction in neurodegenerative disease, mechanisms underlying impaired NCT warrant further investigation, especially to identify therapeutic targets.

Outlook on mechanisms underlying MRE disorders

While the majority of MRE disorders discovered involve neurodegeneration, an increasing number are suspected to underlie neurodevelopmental disorders, such as autism and intellectual disabilities, yet our understanding of why neurons remain selectively vulnerable, especially older ones, remains incomplete. Significant progress has been made in terms of identifying specific pathways that underlie disease pathogenesis for particular MRE disorders, and multiple mechanisms are increasingly observed within individual MRE disorders upon testing, although their precise pathogenic contributions often remain to be determined. Some observations have challenged preexisting models of neurodegenerative disorders solely explained by gain or loss of function mechanisms, as we are finding that many repeat expansions not only influence transcription or translation of underlying host genes, but also exert effects at the level of RNA. For example, C9-ALS/FTD is characterized by sense and antisense intranuclear RNA transcripts. These RNAs may form intranuclear foci or undergo RAN translation, each with variable contributions to neuronal toxicity. Neuronal atrophy may be influenced by a variety of processes, such as age, stress, or cell type specific expression patterns of repeat RNA, and likely influences whether C9 carriers are diagnosed with ALS, FTD, or both. The discovery of widespread TDP-43 mislocalization in multiple neurodegenerative disorders has highlighted the potential link between neuronal atrophy and mislocalization of select RBPs that critically orchestrate neuronal gene expression. Additionally, understanding how epigenetic or epitranscriptomic mechanisms like methylation of DNA or RNA, respectively, influence repeat RNA production, turnover, foci formation, or RAN translation would be exciting areas for future research exploration and may explain why some elderly MRE carriers lack apparent disease symptomatology or exhibit delayed onset.

Therapeutic strategies to target underlying MRE pathology

While many advances have been made in understanding key neuropathological mechanisms underlying various MRE disorders, there is still a surprising dearth of clinically approved therapies targeting these

mechanisms. Indeed, while most approved treatments for MRE disorders provide only limited symptomatic amelioration, there are currently several therapeutic strategies under preclinical investigation that target the underlying pathology of MRE disorders (Dickey and La Spada 2018; Egorova and Bezprozvanny 2019; Ishikawa and Nagai 2019; Panza et al. 2020). By targeting the underlying pathology, researchers aim to prevent the spectrum of disease phenotypes associated with a specific neuropathological MRE (Rohilla and Gagnon 2017). Moreover, many proposed therapeutics target post-transcriptional pathologies that share similarities to those seen in several other MRE disorders. These features may enable identification of robust therapeutic modalities that can be readily adapted for the treatment of alternative MRE disorders characterized by similar cellular impairments. For example, as some neurological MRE disorders share the same MRE sequence, such as CAG in multiple spinocerebellar ataxias (SCAs) and Huntington's disease, it is conceivable that a single therapy targeting the repeat site can be readily adapted for multiple CAG-repeat expansion disorders (Kotowska-Zimmer et al. 2020). For the purposes of this review, we will focus on therapeutic strategies that suppress repeat RNA levels, inhibit RNA foci formation, and block RAN-mediated toxicity, as these are major drivers of disease pathogenesis across several previously discussed neurological MRE disorders with diverse underlying repeat sequence compositions. Encouragingly, the preclinical success of such therapeutic strategies has been evaluated by researchers experimenting with a variety of *in vitro* or *in vivo* preclinical models of MRE diseases using highly versatile therapeutic agents, such as antisense oligonucleotides (ASOs), engineered RBPs, and small molecules (SMs).

One straightforward strategy that both prevents formation of RNA foci and blocks translation of RAN peptides is to selectively suppress toxic repeat RNA, as opposed to the neurotypical allele of a healthy polymorphic repeat length. Methods to suppress MRE RNA include inhibiting transcription, eliminating the MRE transcript, or promoting degradation of MRE RNA, but these approaches could have deleterious consequences for brain function if translation of the host gene were critical for neuronal development, homeostasis, or plasticity, for example (e.g., *FMR1* in FXS and FXTAS or *HTT* in HD). Other strategies that specifically target the repeat sequences such as by RNA editing are also appealing, but would require significant improvements to specificity and delivery before translation to the clinic. Additionally, since CAG repeat length-dependent aberrant splicing of *HTT* exon 1 may generate

pathological, polyadenylated, repeat-containing *HTT* exon 1, MRE therapies targeting sequences present in non-repeat containing exons may be insufficient, so long as repeat-containing exons are translated independent of targeted sequence elements (Gipson et al. 2013; Sathasivam et al. 2013). For these reasons, alternative strategies that specifically prevent downstream toxicities linked to repeat RNA production, translation, or accumulation are highly desired, as well.

One such strategy that avoids repeat RNA elimination includes "RNA masking," whereby the repeat RNA sequences are competitively bound by ASOs or engineered RBPs that prevent formation of secondary structures that underlie sequestration of critical proteins within RNA foci. As opposed to eliminating the RNA, successfully masking repeat RNA sequences, such as with ASOs, provides another appealing strategy to prevent RNA hairpin formation, RNA foci formation, or RAN translation, while still potentially enabling translation of the underlying host gene. A final therapeutic strategy under consideration involves elimination of RAN translation products either by clearance with antibodies or ASO-mediated knockdown of genes necessary for RAN translation. Together, these strategies to eliminate pathogenicity in MRE disorders rely upon a variety of therapeutics, including ASOs, antibodies, small molecules, and engineered or repurposed nucleic acid binding proteins, such as DNA-targeting Cas9 or RNA-targeting Cas proteins (engineered rCas9, Cas13Rx).

Background on potential, versatile, bioactive tools for MRE disorders

Some therapeutic modalities under consideration for potential treatment of MRE disorders are recent products of scientific discovery and protein engineering (e.g. Cas-based strategies), but others, such as small molecules, antibodies, and ASOs, have been tested in various disease contexts for decades. Perhaps the most well-studied therapeutic agent discussed in subsequent sections involves designer oligonucleotides that recognize specific antisense target sequences to impart a variety of RNA-based functions. Based on the molecular architecture of sequence complementarity, for example, diverse RNA metabolic reactions may occur. Some ASOs direct target RNA elimination by RNase H upon sequence complementarity across approximately 20 nucleotides, while others promote RISC-mediated translational suppression following near perfect complementarity to a sequence of 7 or 8 nucleotides within the 5' sequence of target mRNA, mimicking the mechanism of action of microRNA. For some MRE transcripts, ASOs

may be used to modify alternative splice site selection to produce mature mRNA either lacking expanded repeat tracts or retaining introns within coding sequences, introducing premature termination codons that mark the repeat mRNA as a substrate for nonsense mediated decay (NMD). Alternatively, some ASOs could potentially be used to inhibit miRNA-mediated gene silencing by RISC of key dosage sensitive proteins that may otherwise be sequestered within intranuclear RNA foci, further expanding the therapeutic relevance of ASOs for the treatment of MRE disorders.

In addition to more historically well-characterized therapeutic agents like ASOs, small molecules, or antibodies, engineered RBPs have recently emerged as appealing candidates to rescue RNA-based phenotypes associated with MRE expansion. Although engineered variants of human expressed proteins, such as zinc finger family members and PUFs, are predicted to engender more limited immune responses than Cas-based therapies, achieving robust target specificity and binding efficiencies needed for manipulation of MRE gene expression can be time consuming and presents several design challenges. In terms of adaptability and scalability, however, Cas proteins represent one of the most attractive candidates to engineer for modulating MRE gene expression.

Initially discovered as prokaryotic genomic defense surveyors, CRISPR-Cas (clustered regularly interspaced short palindromic repeats and CRISPR-associated proteins) systems provide an adaptive immune system for bacteria and archaea to defend against pathogens, such as phages. CRISPR-Cas defends host cells by eliminating foreign genomic material, which is achieved upon recognition of sequences complementary to those present in previously encountered pathogens. Specifically, Cas-based adaptive immune responses involve storing molecular memories of pathogens as 25–35 bp sequences within a CRISPR array, which, when transcribed and post-transcriptional processed, complexes with Cas proteins to eliminate foreign nucleic acid. At least two CRISPR-Cas systems have evolved, each differing in CRISPR-RNA processing and CRISPR-Cas locus organization, but the defining feature is the required number of Cas effectors. Smaller and presumably more easily packaged into viral delivery systems, CRISPR class 2 systems only require one effector protein, such as Cas9 or Cas13. Compared to Cas9 which naturally evolved to eliminate foreign DNA genomes, the smaller Cas13 targets foreign RNA genomes. Recently, our lab and others have engineered the DNA-targeting Cas9 to recognize RNA and by fusing either GFP or nucleases to RNA-targeting Cas9 (rCas9),

researchers can track or eliminate specific transcripts within cells, respectively (Nelles et al. 2016; Batra et al. 2017). The versatility of the Cas system to target diverse and specific nucleic acid species combined with its potential to be fused to effector molecules, such as nucleases, but potentially helicases or RNA editors as well, make RNA-targeting Cas proteins an exciting area of exploration for future MRE therapies.

Reductions of toxic RNA levels

Suppression of repeat RNA transcript levels is a straightforward strategy for treating MRE disorders characterized by toxic RNA gain of function (Figure 2), but specifically targeting the expanded allele can be challenging (Swinnen et al. 2020). Leading approaches to suppress repeat RNA include (1) inhibiting transcription of the MRE allele, (2) eliminating MRE RNA, and (3) modulation of repeat RNA levels through alternative splicing; however, these approaches would be ineffective for disorders sensitive to loss of the repeat-expanded allele, which likely includes many ataxias. To treat MRE disorders characterized by gene dosage sensitivity, additional RNA-elimination therapies may rely upon generation of functionally compensatory alternative splice variants lacking repeat expansion sequences, for example. The preclinical success of versatile and promising therapeutics will be evaluated further within this section, as these factors can reduce toxic RNA levels through transcriptional inhibition, elimination of MRE RNA, or modulation of repeat RNA levels.

Inhibition of MRE transcription is an attractive mechanism to prevent many post-transcriptional MRE toxicities and can be associated with reduced RNA polymerase function, repression of transcription factors, activation of repressors, and remodeling of histones, among other mechanisms. Therapeutic reductions in MRE transcription can be achieved with small molecules that globally inhibit transcription, such as HDAC inhibitors; however, such approaches may present significant side effects. Despite these challenges, one remarkable preclinical example includes use of the general transcription inhibitor Actinomycin D at low, nontoxic levels for the selective treatment of DM1 in mice (Siboni et al. 2015). Actinomycin D is an FDA-approved chemotherapeutic that preferentially binds GC-rich DNA sequences, such as the CTG expansion characteristic of DM1. When mouse models of DM1 were treated with low, nontoxic doses of Actinomycin D, researchers observed reduced CUG repeat RNA levels and a partial restoration of mRNA splicing profiles (Siboni et al. 2015). This raises the intriguing possibility that sub-chemotherapeutic

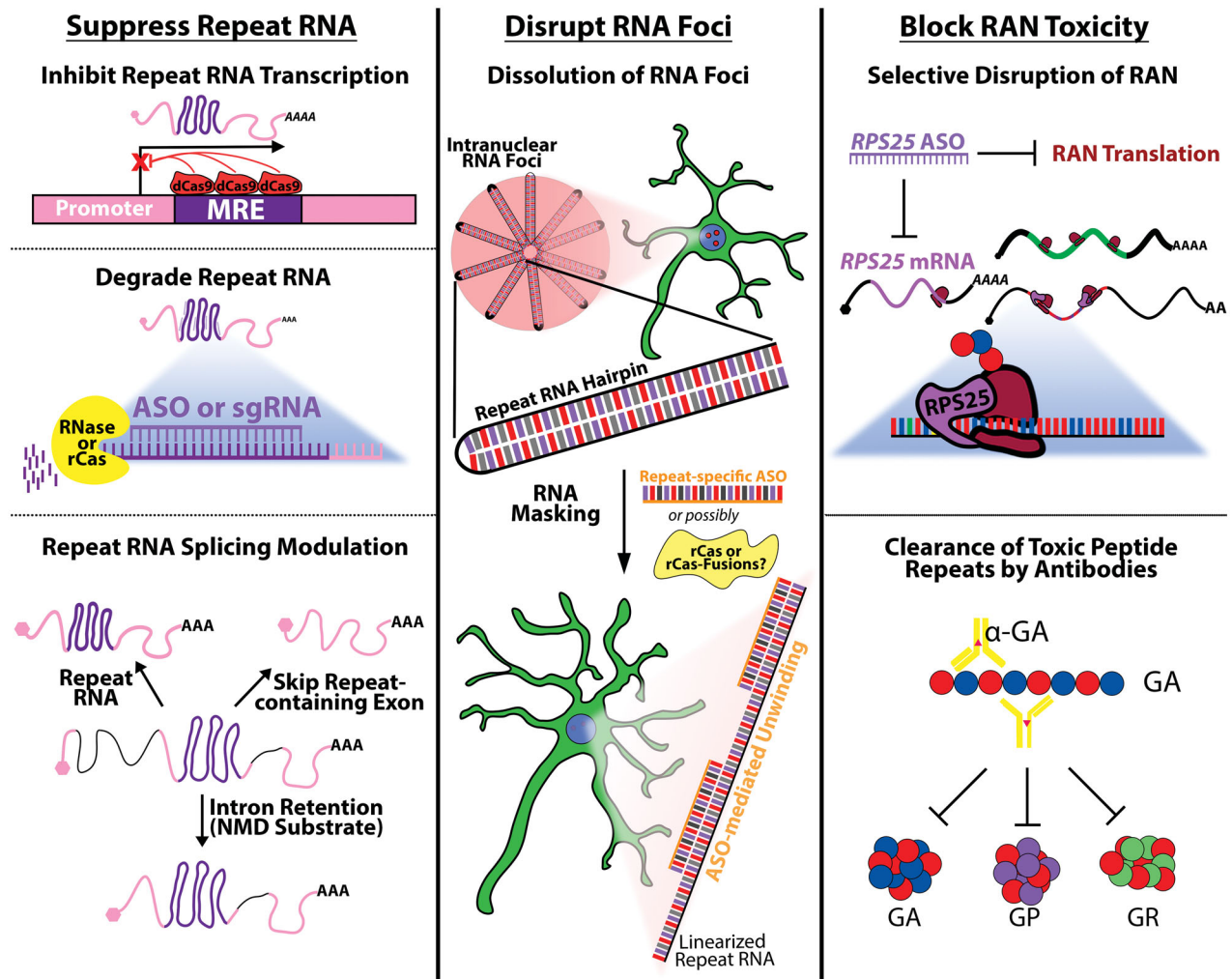


Figure 2. Schematic depicting major therapeutic strategies and associated approaches for treating molecular pathologies associated with neurological MRE disorders. To suppress repeat RNA, repeat RNA synthesis could be inhibited, such as with catalytically inactive dCas9. Alternatively, repeat RNA could be targeted for degradation by RNaseH through ASOs or by fusions of nucleases to rCas through sgRNA. Additionally, modulation of repeat-containing pre-RNA splicing may prevent accumulation of mature repeat-containing RNA by promoting either skipping of the repeat-containing exon or retention of an intron to mark the transcript as a substrate for nonsense mediated decay, for example. To disrupt RNA foci formation independent of eliminating repeat transcripts, ASOs can be used to inhibit pathogenic secondary structures that recruit RBPs and give rise to intranuclear foci. To block RAN toxicity, two strategies are shown that either inhibits a critical regulator of RAN translation, but not canonical translation, or promotes clearance of RAN peptides with antibodies.

doses of ActD may exhibit preclinical success for the treatment of other CG-rich repeat expansion disorders, such as FXTAS, but given their roles as global inhibitors of transcription, selective inhibitors of MRE transcription or repeat RNA levels are also desired.

To this end, designer drugs and small molecule chemical screens have been employed for selective reductions of repeat RNA levels. In contrast to small molecules that displace the RBPs sequestered by CUG repeat RNA, Cugamycin recognizes the three-dimensional structure of CUG repeat expansions present in DM1 and promotes transcript degradation through cleavage, similar to Bleomycin-based cleavage of RNA and DNA (Angelbello et al. 2019). The demonstration

that designer drugs can selectively degrade toxic repeat RNA highlights this strategy's potential for treatment of other repeat expansion disorders. In addition to drugs designed to directly target RNA, chemical screens have identified potent modulators of repeat RNA toxicity. Unexpectedly, for example, such tests revealed that multiple inhibitors of microtubule function selectively reduce toxic CUG repeat RNA levels and partially rescue splicing profiles in a DM1 HeLa model (Reddy et al. 2019). Subsequent tests of the FDA-approved microtubule inhibitor colchicine in DM1 mice and patient cells revealed a selective modulation of CUG repeat RNA levels, likely a result of impaired cytoskeletal and nucleoskeletal complexes, raising the possibility that

colchicine may be useful in treatment of other repeat expansion disorders (Reddy et al. 2019).

Successful and specific *in vitro* suppression of repeat RNA levels at the level of transcription has also been achieved in select cellular and animal models of DM by expression of catalytically dead DNA-targeting Cas9 (dCas9) and sgRNA (Pinto et al. 2017). Although long-term efficacy, multi-systemic delivery, and therapeutic safety have yet to be demonstrated in animal models, injection of AAV6-dSaCas9 encoding CAG-targeting sgRNA reduced myotonia within targeted muscle fibers in a mouse model of DM1 (Pinto et al. 2017). Additionally, dCas9 may rescue molecular phenotypes associated with C9-ALS/FTD (Pinto et al. 2017). Presumably, dCas9:sgRNA complexes tile along MRE genomic loci and this avidity sterically hinders elongation of RNA polymerases independent of double strand breaks, a significant therapeutic concern for active Cas nucleases. These dCas9 tools have since been further adapted to target repeat RNA directly for elimination as well as genomic loci associated with various MRE disorders, highlighting the versatility of dCas9 for direct RNA elimination and transcriptional inhibition.

One way dCas9 can modulate transcription of target loci is through fusion to epigenetic modifiers to create molecular tools often referred to as either CRISPRi, for Cas-based inhibitors of transcription, or CRISPRa systems, for Cas-based activators of transcription (Brezgin et al. 2019). For example, dCas9-KRAB promotes CRISPRi-based silencing of genes by catalyzing formation of repressive chromatin marks, and this tool has effectively rescued *in vitro* phenotypes present within models of DM1, DM2, and C9-ALS/FTD either by reducing toxic MRE transcription or by identifying genetic modifiers of neurodegeneration through high throughput screens (Kramer et al. 2018; Ikeda et al. 2020). In contrast to transcriptional inhibition, dCas9-Tet1 promotes CRISPRa-based gene activation by catalyzing oxidation of repressive methylated cytosines within the underlying DNA, such as those present within the CGG trinucleotide stretch characteristic of FXS (Liu et al. 2018). Earlier tests of global pharmacological inhibitors of epigenetic silencing demonstrated feasibility of restoring *FMR1* transcript levels, but dCas9-Tet1 provides a target-specific strategy (Kumari and Usdin 2016). It should be noted that while researchers demonstrated remarkable rescue of multiple *in vitro* and *in vivo* FXS phenotypes dependent upon targeting dCas9-Tet1 to the *FMR1* promoter, the consequences of prolonged CGG-repeat expanded *FMR1* expression in these models remain unknown, providing a unique tool

to investigate CGG MRE pathobiology (Kumari and Usdin 2016; Liu et al. 2018).

Preclinical regulation of transcription has also been achieved by engineered proteins such as zinc finger proteins, ZFPs (Garriga-Canut et al. 2012). In mouse models of HD, for example, virally expressed ZFPs targeting genomic CAG repeats reduced expression of mutant *HTT* RNA and protein with minimal reductions in levels of non-pathological, endogenous CAG-containing transcripts (Garriga-Canut et al. 2012), suggesting feasibility of selectively targeting expanded repeat sequences. Similar to the CRISPRi system, KRAB repressor domains fused to ZFPs or other genome targeting agents, such as TALENs, may promote epigenetic silencing of target loci.

In addition to transcriptional inhibition, MRE RNA can be suppressed post-transcriptionally by direct elimination of RNA. Current methods to eliminate repeat RNA include oligonucleotides (ONs), siRNAs, and engineered variants of Cas9 and RBPs. ONs and siRNAs can promote degradation of transcripts harboring complementary target sites through RNase H1 and RISC machinery, respectively, both of which are broadly expressed in neurological tissue (Liang et al. 2017; Rinaldi and Wood 2018). ONs targeting pathological MRE transcripts have shown preclinical promise with *in vitro* models of SCA2, DM1, and FECD, while siRNA-based targeting of MRE transcripts has shown promise in models of HD, C9-ALS/FTD, and multiple SCAs.

The earliest and most common method of ON-targeted degradation relies upon recruitment of RNase H by ASOs, but additional ON-based methods rely upon recruitment of RISC by either siRNA to degrade bound targets or by miRNA to suppress protein synthesis. Currently, comparisons between efficacy of siRNA and ASOs are limited, but should be forthcoming as ON chemical modifications become optimized and more ON therapies advance through clinical trials (Chi et al. 2017). Likewise, comparisons of therapeutic benefits between ONs that target only the specific repeat-expanded allele versus both wild-type and expanded alleles remain inconclusive (Leavitt and Tabrizi 2020). Indeed, since designing effective allele-specific ONs relies upon identifying single nucleotide polymorphisms (SNPs) and can present significant challenges, it is encouraging that allele-specific and nonspecific targeting strategies have demonstrated significant therapeutic potential in MRE disorders such as HD, DM1, and ataxias (Kordasiewicz et al. 2012; Thornton 2014; Cepeda and Tong 2018; Egorova and Bezprozvanny 2019; van Cruchten et al. 2019; McLoughlin et al. 2020).

For example, *in vivo* administration of ASOs targeting both alleles of DMPK mRNA for degradation by RNase H reduced levels of ribonuclear foci, improved global splicing profiles and for specific MBNL1 target exons, and partially rescued myotonia in a mouse model of DM1 (Wheeler et al. 2012). Additionally, siRNAs and ASOs have been developed that non-allele specifically target *ATXN2*, expansions of which are linked to SCA2. Intriguingly, reducing *ATXN2* levels demonstrated therapeutic potential for the treatment of ALS linked to TDP-43 (Becker and Gitler 2018). This observation may further highlight the versatility and potential applications for successfully developed ASOs that degrade repeat RNAs beyond disorders defined by MRE.

In addition to novel ON-based methods, a recent tool that has been developed to target specific RNA transcripts includes engineered variants of Cas9. These engineered variants of Cas9 that target RNA have also demonstrated preclinical potential to eliminate repeat RNA. Our lab developed variants of SpCas9 to either label or eliminate RNA by engineering a catalytically inactivated RNA-targeting Cas9 fused with a fluorophore (i.e. GFP) or nuclease (i.e. PIN), respectively (Nelles et al. 2016; Batra et al. 2017). By co-expressing an MRE disease re-pet sequence and dCas9-PIN along with a targeting or non-targeting sgRNA, RNA FISH revealed dramatic reductions in the fraction of cells containing RNA foci in conditions expressing targeting sgRNA compared to conditions expressing non-targeting sgRNA. Specifically, guiding rCas9-PIN to repeat RNA reduced levels of RNA foci associated with *in vitro* models of DM, C9-ORF72, and polyQ disorders, like HD (Batra et al. 2017). These findings suggest widespread versatility in terms of targeting diverse repeat RNA sequences with modular effector domains that could be fused to rCas.

In a recent demonstration from our group, we observed that sustained expression of virally-encoded rCas9-PIN targeting expanded CAG repeat RNA reversed several phenotypes in neonatal and adult mouse models of DM1 (Batra et al. 2020). Our group noted that intramuscular or systemic injections of AAV-encoding rCas9-PIN with CAG-targeting sgRNA resulted in expression that endured for nearly 3 months with sustained specificity for expanded CAG repeats. While future experiments will seek to more thoroughly characterize *in vivo* phenotypes, identify off-target effects, and mitigate the inherent immunogenicity from expression of non-self Cas9 fusions, this work highlights the feasibility of prolonged AAV-mediated expression of rCas9-PIN for targeting a variety of MRE disorders (Batra et al. 2020).

More recently, smaller Cas proteins that target RNA have been discovered and subjected to preclinical tests of eliminating RNA-based neuropathology. Engineered variants of several Cas proteins are currently under exploration in preclinical contexts, but one of the most well characterized includes Cas13 (Paul and Montoya 2020; Smargon et al. 2020; Xu et al. 2020). Compared to rCas9, Cas13 naturally evolved to recognize RNA substrates and has the added benefit of a smaller size, which facilitates packaging into viruses, like AAV (Xu et al. 2020). Engineered variants of Cas13 with improved subcellular distribution and specificity for RNA elimination or tracking in mammalian cells have also been developed (Abudayyeh et al. 2017; Cox et al. 2017). Additionally, catalytically inactive dCas13 can block pathological splice site inclusion within transcripts like *MAPT*, the gene encoding the microtubule-associated protein Tau, which is associated with multiple neurological disorders (Konermann et al. 2018). While not tested in MRE diseases yet, this provides a potentially interesting avenue of research especially given the versatility of catalytically active Cas13dRX-NLS in targeting different types of RNA as well as double stranded products (Konermann et al. 2018). Current RNA-targeting Cas proteins are appealing agents to rescue RNA pathologies *in vitro* that require limited, if any, protein engineering to achieve target specificity, but rCas proteins can present significant immunological challenges for human therapeutics. Indeed, as up to 67% of the population may have preexisting antibodies to SpCas9 (Simhadri et al. 2018; Charlesworth et al. 2019), alternative approaches that avoid immune activation are desired.

Engineering human RBPs may present less of a threat to immune activation (Charlesworth et al. 2019), but remains technically challenging as a new protein or domain must be created per target site as opposed to simply designing a new complementary sgRNA for CRISPR methods (Gilbert et al. 2014; Xu et al. 2020). Like Cas proteins though, human RBPs they can be fused to effector domains to directly degrade their targets or modulate splicing (Klug 2010; Quenault et al. 2011; Sulej et al. 2012; Chen and Varani 2013; Yang et al. 2017). In addition to ZFPs, Pumilio/fem-3-binding factor (PUF) proteins have been engineered to suppress transcript levels. Indeed, a PUF code for RNA recognition has been proposed to target specific RNA sequences. PUF proteins play diverse RNA metabolic and physiological roles across eukaryotic development and their binding domains and targets have been well characterized, revealing a bias for domains that bound A, U, or G, but not C. By directed evolution for PUF variants that specifically bound cytosine, the RNA target recognition

code by PUFs was expanded to now include all four ribonucleotides, raising exciting possibilities for treatment of RNA-based disorders (Filipovska and Rackham 2011; Filipovska et al. 2011).

Another way to achieve RNA suppression is through modulation of RNA splicing, which has primarily been accomplished with ASOs, but recent engineering has extended to dCasRx, as well (Konermann et al. 2018). RNA splicing could, for example, promote mRNA decay through intron retention and nonsense-mediated decay. Alternatively, normal RNA function can be restored using ASOs to modulate alternative splicing, a method shown to be effective for the MRE disorder spinal cerebellar ataxia 3 (SCA3) (Evers et al. 2013; Toonen et al. 2017; McIntosh et al. 2019). In SCA3, the polyQ MRE occurs at the second to last exon of Ataxin-3, exon 10. Different groups have targeted chemically modified ASOs to pre-mRNA to produce mature mRNA transcripts lacking the pathogenic repeat expansion sequence. Instead of relying upon degradation of the repeat containing transcript, these ASOs mask splicing sites which lead to the exclusion of the exon 10 while maintaining the correct reading frame (Toonen et al. 2017; McIntosh et al. 2019). This leads to the reduction of the small, aggregate prone, toxic, polyQ-containing proteins, while retaining normal ubiquitination function of Ataxin-3 (Toonen et al. 2017).

Disruption of RNA foci

Although the exact contributions of RNA foci to onset and severity of MRE disease phenotypes remain under investigation, RNA foci are pathological hallmarks of several neurodegenerative disorders. In many MRE disorders characterized by ribonuclear foci, higher order, repeat RNA structures tend to promote focal sequestration of critical proteins. Sequestration of proteins – like the splicing factor MBNL1 within DM1-linked foci or the microRNA microprocessor within FXTAS-linked foci – is suspected to be a key pathological mechanism in several MRE disorders, including SCA2, FECD, DM1, FXTAS, and C9-ALS/FTD. Thus, approaches that dissolve toxic foci may provide attractive therapeutics for multiple MRE disorders. One promising approach that may disrupt formation of RNA foci across a variety of MRE disorders includes disruption of underlying RNA structures, such as by RNA masking with ASOs or alternatively, by RNA hairpin unwinding with rCas fusions to helicases. Other leading strategies associated with reduced RNA foci formation, which may potentially be independent of reduced repeat RNA levels, include small molecules that modulate neuronal signaling, although the

mechanisms remain less well understood. While several promising therapies have emerged that prevent the formation of RNA foci, presumably, many such strategies may enable potentially deleterious nuclear export of repeat expanded RNAs. Depending on the therapeutic approach, by dissolving intranuclear RNA foci, cytoplasmic repeat RNAs may have the potential to aggregate or provide substrates for RAN translation, so caution is needed in understanding long-term physiological consequences of such therapeutic approaches.

Multiple groups have shown preclinical reductions of RNA–protein foci associated with MRE disorders. More than a decade ago, researchers rescued cellular phenotypes associated with a mouse model of DM1 by selectively preventing MRE RNA foci formation. Specifically, CUG repeat RNA characteristic of DM1 was competitively bound by CAG antisense morpholino oligonucleotides that do not trigger RNaseH-transcript elimination, but instead, disrupt CUG repeat RNA hairpins. By preventing CUG repeat hairpin formation, pathological substrates for otherwise sequestered RBPs, such as MBNL1, were eliminated, along with the characteristic RNA foci. Indeed, following CAG25 morpholino, CUG repeat was found to be diffuse, as opposed to punctate, within the nucleus (Wheeler et al. 2009). Additionally, masking ASOs reduced toxic foci in other disorders, such as FECD (Hu et al. 2018) and C9-ALS (Donnelly et al. 2013), and preliminary data *in vitro* solely targeting CAG repeats have shown promise by reducing foci in HD (Urbanek et al. 2017). Similarly, reductions in toxic foci, independent of changes in repeat RNA level, have been found for Fuchs' endothelial corneal dystrophy (FECD) using masking ASOs on patient cells and corneal explants from FECD patients (Hu et al. 2018). Moreover, owing to the similarities between repeat sequences across MRE disorders, it may be possible to develop a single, highly versatile oligonucleotide that rescues pathological foci formation in multiple MRE disorders, like SCA3, HD, and DRPLA (Fischer et al. 2016).

Other strategies that may reduce RNA foci include those that target neuronal signaling pathways, such as by altering neuronal activity or stress responses. Indeed, as a growing body of evidence suggests prolonged neuronal activity or neural stress can disrupt homeostatic LLPS, therapeutics that support healthy biophysical phase separation may be able to disrupt pathological foci formation. For example, activation of the Nrf2-antioxidant stress pathway by small molecules improved multiple MRE disease phenotypes by reducing formation of polyglutamine aggregates in models of SCA3, SCA17, SBMA, and HD (Oliveira et al. 2015; Bott et al. 2016; Brandes and Gray 2020). SSRIs and

antipsychotic drugs represent other small molecules that may suppress pathological foci formation by regulating neuronal signaling pathways.

Elimination of RAN toxicity

The increasing number of MRE disorders characterized by toxic peptide repeats presents a therapeutic challenge that has been addressed by removal of RAN translation products themselves or knockdown of RBPs that are required for RAN translation to occur. While many therapies related to RAN have focused on DPRs in C9-ALS/FTD, such therapies may be readily adaptable for treatments of other MREs with RAN toxicity, such as SCA8, FXTAS, DM1, and HD. To this end, antibodies have been discovered that selectively target GA dipeptides for clearance, and while such a targeted strategy may seem impractical for MRE disorders characterized by multiple DPRs, recent results encouragingly suggest toxicities from other DPRs were reduced as well. Other broad therapeutic strategies aim to eliminate RAN toxicity by first identifying proteins specific to RAN translation and then disrupting that gene's function in a disease context.

Two proteins shown to promote RAN translation include RPS25 and SRSF1. The small ribosomal subunit protein RPS25 was previously discussed and was identified as part of a screen in yeast for modifiers of C9 DPR toxicity (Yamada et al. 2019). RPS25 selectively regulates RAN translation of G4C2 repeats in yeast and patient cells and RPS25 knockdown can suppress polyGP toxicity in patient derived iPSCs or extend lifespan in drosophila models of C9-ALS/FTD. Although it remains unknown whether reducing RPS25 expression may benefit other MRE disorders, the screen and subsequent counterscreen strategy to identify selective regulators of RAN translation, as opposed to global translation, could be applied to other MRE disorders. For those such MRE disorders, such a strategy would potentially uncover modifiers that could also be readily targeted for gene knockdown by ASO. In addition to RPS25, SRSF1 has emerged as a recent candidate that promotes RAN translation of C9orf72 by exporting G4C2 RNA from the nucleus to the cytoplasm (Hautbergue et al. 2017). Since only 0.5–2% of the transcriptome is reportedly exported by SRSF1, a protein suspected to have highly redundant functions with proteins SRSF2–6, loss of SRSF1 is suspected to have minimal deleterious side effects. Indeed, SRSF1 knockdown noticeably extended lifetime and improved locomotor function in flies, consistent with the reduced motor

neuron death and astrocytic toxicity phenotypes observed in patient-derived cells.

Outlook on therapeutic strategies that target post-transcriptional MRE pathology

Formation of RNA foci and translation of toxic RAN peptides drive pathology in a variety of MRE disorders. Therapies that inhibit, eliminate, or correct these processes may rescue underlying disease pathology and provide therapeutic benefit. Recent therapeutic strategies that have emerged to treat post-transcriptional pathologies in MRE disorders include testing oligonucleotides (e.g. ASOs) to eliminate toxic RNA or interrupt RBP sequestration, engineering or repurposing nucleic acid binding proteins to eliminate RNA, optimizing small molecules to disrupt repeat RNA structure and pathogenicity, and developing antibodies to eliminate toxic RAN translation products. Each of these approaches has shown preclinical promise and translation of these tests into preclinical models, particularly *in vivo* ones, is desired.

Disclosure statement

GWY is co-founder, member of the Board of Directors, on the SAB, equity holder, and paid consultant for Locanabio and Eclipse BiolInnovations. GWY is a visiting professor at the National University of Singapore. *GWY's interests have been reviewed and approved by the University of California, San Diego in accordance with its conflict of interest policies.* The authors declare no other competing financial interests.

Funding

This work was partially supported by the NIH (NS103172, AG069098, EY029166) to G.W.Y.

ORCID

Joshua L. Schwartz  <http://orcid.org/0000-0002-4263-2829>
Gene W. Yeo  <http://orcid.org/0000-0002-0799-6037>

References

- Abudayyeh OO, Gootenberg JS, Essletzbichler P, Han S, Joung J, Belanto JJ, Verdine V, Cox DBT, Kellner MJ, Regev A, et al. 2017. RNA targeting with CRISPR-Cas13. *Nature*. 550(7675):280–284.
- Angelbello AJ, Rzuczek SG, McKee KK, Chen JL, Olafson H, Cameron MD, Moss WN, Wang ET, Disney MD. 2019. Precise small-molecule cleavage of an r(CUG) repeat expansion in a myotonic dystrophy mouse model. *Proc Natl Acad Sci USA*. 116(16):7799–7804.

- Ariza J, Steward C, Rueckert F, Widdison M, Coffman R, Afjei A, Noctor SC, Hagerman R, Hagerman P, Martínez-Cerdeño V. 2015. Dysregulated iron metabolism in the choroid plexus in fragile X-associated tremor/ataxia syndrome. *Brain Res.* 1598:88–96.
- Arnold FJ, Merry DE. 2019. Molecular mechanisms and therapeutics for SBMA/Kennedy's disease. *Neurotherapeutics.* 16(4):928–947.
- Aumiller WM, Cakmak FP, Davis BW, Keating CD. 2016. RNA-based coacervates as a model for membraneless organelles: formation, properties, and interfacial liposome assembly. *Langmuir.* 32(39):10042–10053.
- Banez-Coronel M, Ranum LPW. 2019. Repeat-associated non-AUG (RAN) translation: insights from pathology. *Lab Invest.* 99(7):929–942.
- Barman P, Reddy D, Bhaumik SR. 2019. Mechanisms of antisense transcription initiation with implications in gene expression, genomic integrity and disease pathogenesis. *Noncoding RNA.* 5:11.
- Batra R, Charizanis K, Manchanda M, Mohan A, Li M, Finn DJ, Goodwin M, Zhang C, Sobczak K, Thornton CA, et al. 2014. Loss of MBNL leads to disruption of developmentally regulated alternative polyadenylation in RNA-mediated disease. *Mol Cell.* 56(2):311–322.
- Batra R, Nelles DA, Pirie E, Blue SM, Marina RJ, Wang H, Chaim IA, Thomas JD, Zhang N, Nguyen V, et al. 2017. Elimination of toxic microsatellite repeat expansion RNA by RNA-targeting Cas9. *Cell.* 170(5):899–912 e810.
- Batra R, Nelles DA, Roth DM, Krach F, Nutter CA, Tadokoro T, Thomas JD, Sznajder LJ, Blue SM, Gutierrez HL, et al. 2020. The sustained expression of Cas9 targeting toxic RNAs reverses disease phenotypes in mouse models of myotonic dystrophy type 1. *Nat Biomed Eng.* <https://www.nature.com/articles/s41551-020-00607-7#citeas>
- Bavassano C, Eigentler A, Stanika R, Obermair GJ, Boesch S, Dechant G, Nat R. 2017. Bicistronic CACNA1A gene expression in neurons derived from spinocerebellar ataxia type 6 patient-induced pluripotent stem cells. *Stem Cells Dev.* 26(22):1612–1625.
- Becker LA, Gitler AD. 2018. Ataxin-2 is Droppin' some knowledge. *Neuron.* 98(4):673–675.
- Boehringer A, Bowser R. 2018. RNA nucleocytoplasmic transport defects in neurodegenerative diseases. *Adv Neurobiol.* 20:85–101.
- Bott LC, Badders NM, Chen KL, Harmison GG, Bautista E, Shih CC, Katsuno M, Sobue G, Taylor JP, Dantuma NP, et al. 2016. A small-molecule Nrf1 and Nrf2 activator mitigates polyglutamine toxicity in spinal and bulbar muscular atrophy. *Hum Mol Genet.* 25(10):1979–1989.
- Botta A, Vallo L, Rinaldi F, Bonifazi E, Amati F, Biancolella M, Gambardella S, Mancinelli E, Angelini C, Meola G, et al. 2007. Gene expression analysis in myotonic dystrophy: indications for a common molecular pathogenic pathway in DM1 and DM2. *Gene Expr.* 13(6):339–351.
- Brandes MS, Gray NE. 2020. NRF2 as a therapeutic target in neurodegenerative diseases. *ASN Neuro.* 12:1759091419899782
- Brangwynne CP, Tompa P, Pappu RV. 2015. Polymer physics of intracellular phase transitions. *Nature Phys.* 11(11):899–904.
- Brezgin S, Kostyusheva A, Kostyushev D, Chulanov V. 2019. Dead Cas systems: types, principles, and applications. *Int J Mol Sci.* 20:6041.
- Budworth H, McMurray CT. 2013. Bidirectional transcription of trinucleotide repeats: roles for excision repair. *DNA Repair (Amst).* 12(8):672–684.
- Burguete AS, Almeida S, Gao FB, Kalb R, Akins MR, Bonini NM. 2015. GGGGCC microsatellite RNA is neuritically localized, induces branching defects, and perturbs transport granule function. *Elife.* 4:e08881.
- Cammass A, Millevoi S. 2017. RNA G-quadruplexes: emerging mechanisms in disease. *Nucleic Acids Res.* 45(4):1584–1595.
- Cepeda C, Tong XP. 2018. Huntington's disease: from basic science to therapeutics. *CNS Neurosci Ther.* 24(4):247–249.
- Charlesworth CT, Deshpande PS, Dever DP, Camarena J, Lemgart VT, Cromer MK, Vakulskas CA, Collingwood MA, Zhang L, Bode NM, et al. 2019. Identification of preexisting adaptive immunity to Cas9 proteins in humans. *Nat Med.* 25(2):249–254.
- Chen A, Wang PY, Yang YC, Huang YH, Yeh JJ, Chou YH, Cheng JT, Hong YR, Li SS. 2006. SUMO regulates the cytoplasmic nuclear transport of its target protein Daxx. *J Cell Biochem.* 98(4):895–911.
- Chen Y, Varani G. 2013. Engineering RNA-binding proteins for biology. *Febs J.* 280(16):3734–3754.
- Chi X, Gatti P, Papoian T. 2017. Safety of antisense oligonucleotide and siRNA-based therapeutics. *Drug Discov Today.* 22(5):823–833.
- Cho DH, Thienes CP, Mahoney SE, Analau E, Filippova GN, Tapscott SJ. 2005. Antisense transcription and heterochromatin at the DM1 CTG repeats are constrained by CTCF. *Mol Cell.* 20(3):483–489.
- Choi JM, Holehouse AS, Pappu RV. 2020. Physical principles underlying the complex biology of intracellular phase transitions. *Annu Rev Biophys.* 49:107–133.
- Choi SY, Lopez-Gonzalez R, Krishnan G, Phillips HL, Li AN, Seeley WW, Yao WD, Almeida S, Gao FB. 2019. C9ORF72-ALS/FTD-associated poly(GR) binds Atp5a1 and compromises mitochondrial function in vivo. *Nat Neurosci.* 22(6):851–862.
- Cleary JD, Pattamatta A, Ranum LPW. 2018. Repeat-associated non-ATG (RAN) translation. *J Biol Chem.* 293(42):16127–16141.
- Cleary JD, Ranum LP. 2014. Repeat associated non-ATG (RAN) translation: new starts in microsatellite expansion disorders. *Curr Opin Genet Dev.* 26:6–15.
- Colak D, Zaninovic N, Cohen MS, Rosenwaks Z, Yang WY, Gerhardt J, Disney MD, Jaffrey SR. 2014. Promoter-bound trinucleotide repeat mRNA drives epigenetic silencing in fragile X syndrome. *Science.* 343(6174):1002–1005.
- Cox DBT, Gootenberg JS, Abudayyeh OO, Franklin B, Kellner MJ, Joung J, Zhang F. 2017. RNA editing with CRISPR-Cas13. *Science.* 358(6366):1019–1027.
- Cumming SA, Hamilton MJ, Robb Y, Gregory H, McWilliam C, Cooper A, Adam B, McGhie J, Hamilton G, Herzyk P, et al.; Scottish Myotonic Dystrophy Consortium. 2018. De novo repeat interruptions are associated with reduced somatic instability and mild or absent clinical features in myotonic dystrophy type 1. *Eur J Hum Genet.* 26(11):1635–1647.

- Daughters RS, Tuttle DL, Gao W, Ikeda Y, Moseley ML, Ebner TJ, Swanson MS, Ranum LP. 2009. RNA gain-of-function in spinocerebellar ataxia type 8. *PLoS Genet.* 5(8):e1000600.
- Dickey AS, La Spada AR. 2018. Therapy development in Huntington disease: from current strategies to emerging opportunities. *Am J Med Genet A.* 176(4):842–861.
- Donnelly CJ, Zhang PW, Pham JT, Heusler AR, Mistry NA, Vidensky S, Daley EL, Poth EM, Hoover B, Fines DM, et al. 2013. RNA toxicity from the ALS/FTD C9ORF72 expansion is mitigated by antisense intervention. *Neuron.* 80(2): 415–428.
- Egorova PA, Bezprozvanny IB. 2019. Molecular mechanisms and therapeutics for spinocerebellar ataxia type 2. *Neurotherapeutics.* 16(4):1050–1073.
- Evers MM, Tran HD, Zalachoras I, Pepers BA, Meijer OC, den Dunnen JT, van Ommen GJ, Aartsma-Rus A, van Roon-Mom WM. 2013. Ataxin-3 protein modification as a treatment strategy for spinocerebellar ataxia type 3: removal of the CAG containing exon. *Neurobiol Dis.* 58:49–56.
- Fardaei M, Rogers MT, Thorpe HM, Larkin K, Hamshire MG, Harper PS, Brook JD. 2002. Three proteins, MBNL, MBLL and MBXL, co-localize in vivo with nuclear foci of expanded-repeat transcripts in DM1 and DM2 cells. *Hum Mol Genet.* 11(7):805–814.
- Feng Z, Chen X, Wu X, Zhang M. 2019. Formation of biological condensates via phase separation: Characteristics, analytical methods, and physiological implications. *J Biol Chem.* 294(40):14823–14835.
- Filipovska A, Rackham O. 2011. Designer RNA-binding proteins: new tools for manipulating the transcriptome. *RNA Biol.* 8(6):978–983.
- Filipovska A, Razif MF, Nygard KK, Rackham O. 2011. A universal code for RNA recognition by PUF proteins. *Nat Chem Biol.* 7(7):425–427.
- Fiszer A, Ellison-Klimontowicz ME, Krzyzosiak WJ. 2016. Silencing of genes responsible for polyQ diseases using chemically modified single-stranded siRNAs. *Acta Biochim Pol.* 63(4):759–764.
- Flotho A, Melchior F. 2013. Sumoylation: a regulatory protein modification in health and disease. *Annu Rev Biochem.* 82: 357–385.
- Fratra P, Mizielińska S, Nicoll AJ, Zloh M, Fisher EM, Parkinson G, Isaacs AM. 2012. C9orf72 hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes. *Sci Rep.* 2:1016.
- Freibaum BD, Lu Y, Lopez-Gonzalez R, Kim NC, Almeida S, Lee KH, Badders N, Valentine M, Miller BL, Wong PC, et al. 2015. GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. *Nature.* 525(7567):129–133.
- Freibaum BD, Taylor JP. 2017. The role of dipeptide repeats in C9ORF72-related ALS-FTD. *Front Mol Neurosci.* 10:35.
- Fu HJ, Hardy J, Duff KE. 2018. Selective vulnerability in neurodegenerative diseases. *Nat Neurosci.* 21(10):1350–1358.
- Gabryelczyk B, Cai H, Shi X, Sun Y, Swinkels PJM, Salentini S, Pervushin K, Miserez A. 2019. Hydrogen bond guidance and aromatic stacking drive liquid-liquid phase separation of intrinsically disordered histidine-rich peptides. *Nat Commun.* 10(1):5465.
- Garriga-Canut M, Agustin-Pavon C, Herrmann F, Sanchez A, Dierssen M, Fillat C, Isalan M. 2012. Synthetic zinc finger repressors reduce mutant huntingtin expression in the brain of R6/2 mice. *Proc Natl Acad Sci USA.* 109(45): E3136–3145.
- Gasset-Rosa F, Chillon-Marinás C, Goginashvili A, Atwal RS, Artates JW, Tabet R, Wheeler VC, Bang AG, Cleveland DW, Lagier-Tourenne C. 2017. Polyglutamine-expanded huntingtin exacerbates age-related disruption of nuclear integrity and nucleocytoplasmic transport. *Neuron.* 94(1): 48–57 e44.
- Gendron TF, Petrucelli L. 2018. Disease mechanisms of C9ORF72 repeat expansions. *Cold Spring Harb Perspect Med.* 8(4):a024224.
- Gilbert LA, Horlbeck MA, Adamson B, Villalta JE, Chen Y, Whitehead EH, Guimaraes C, Panning B, Ploegh HL, Bassik MC, et al. 2014. Genome-scale CRISPR-mediated control of gene repression and activation. *Cell.* 159(3):647–661.
- Gipson TA, Neueder A, Wexler NS, Bates GP, Housman D. 2013. Aberrantly spliced HTT, a new player in Huntington's disease pathogenesis. *RNA Biol.* 10(11): 1647–1652.
- Green KM, Glineburg MR, Kearse MG, Flores BN, Linsalata AE, Fedak SJ, Goldstrohm AC, Barmada SJ, Todd PK. 2017. RAN translation at C9orf72-associated repeat expansions is selectively enhanced by the integrated stress response. *Nat Commun.* 8(1):2005.
- Grima JC, Daigle JG, Arbez N, Cunningham KC, Zhang K, Ochaba J, Geater C, Morozko E, Stocksdale J, Glatzer JC, et al. 2017. Mutant Huntingtin disrupts the nuclear pore complex. *Neuron.* 94(1):93–107 e106.
- Haeusler AR, Donnelly CJ, Rothstein JD. 2016. The expanding biology of the C9orf72 nucleotide repeat expansion in neurodegenerative disease. *Nat Rev Neurosci.* 17(6): 383–395.
- Hagerman P. 2013. Fragile X-associated tremor/ataxia syndrome (FXTAS): pathology and mechanisms. *Acta Neuropathol.* 126(1):1–19.
- Hagerman RJ, Berry-Kravis E, Hazlett HC, Bailey DB, Jr., Moine H, Kooy RF, Tassone F, Gantois I, Sonenberg N, Mandel JL, et al. 2017. Fragile X syndrome. *Nat Rev Dis Primers.* 3: 17065.
- Hale MA, Johnson NE, Berglund JA. 2019. Repeat-associated RNA structure and aberrant splicing. *Biochim Biophys Acta Gene Regul Mech.* 1862(11-12):194405.
- Hannan AJ. 2018. Tandem repeats mediating genetic plasticity in health and disease. *Nat Rev Genet.* 19(5):286–298.
- Hautbergue GM, Castelli LM, Ferraiuolo L, Sanchez-Martinez A, Cooper-Knock J, Higginbottom A, Lin YH, Bauer CS, Dodd JE, Myszczyńska MA, et al. 2017. SRSF1-dependent nuclear export inhibition of C9ORF72 repeat transcripts prevents neurodegeneration and associated motor deficits. *Nat Commun.* 8:16063.
- Hewitt G, Carroll B, Sarallah R, Correia-Melo C, Ogradnik M, Nelson G, Otten EG, Manni D, Antrobus R, Morgan BA, et al. 2016. SQSTM1/p62 mediates crosstalk between autophagy and the UPS in DNA repair. *Autophagy.* 12(10): 1917–1930.
- Hu J, Rong Z, Gong X, Zhou Z, Sharma VK, Xing C, Watts JK, Corey DR, Mootha VV. 2018. Oligonucleotides targeting TCF4 triplet repeat expansion inhibit RNA foci and missplicing in Fuchs' dystrophy. *Hum Mol Genet.* 27(6): 1015–1026.
- Hu S, Xie Z, Onishi A, Yu X, Jiang L, Lin J, Rho HS, Woodard C, Wang H, Jeong JS, et al. 2009. Profiling the human

- protein-DNA interactome reveals ERK2 as a transcriptional repressor of interferon signaling. *Cell*. 139(3):610–622.
- Ikeda M, Taniguchi-Ikeda M, Kato T, Shinkai Y, Tanaka S, Hagiwara H, Sasaki N, Masaki T, Matsumura K, Sonoo M, et al. 2020. Unexpected mutations by CRISPR-Cas9 CTG repeat excision in myotonic dystrophy and use of CRISPR interference as an alternative approach. *Mol Ther Methods Clin Dev*. 18:131–144.
- Ishiguro T, Sato N, Ueyama M, Fujikake N, Sellier C, Kanegami A, Tokuda E, Zamiri B, Gall-Duncan T, Mirceta M, et al. 2017. Regulatory role of RNA chaperone TDP-43 for RNA misfolding and repeat-associated translation in SCA31. *Neuron*. 94(1):108–124 e107.
- Ishikawa K, Nagai Y. 2019. Molecular mechanisms and future therapeutics for spinocerebellar ataxia type 31 (SCA31). *Neurotherapeutics*. 16(4):1106–1114.
- Iwahashi CK, Yasui DH, An HJ, Greco CM, Tassone F, Nannan K, Babineau B, Lebrilla CB, Hagerman RJ, Hagerman PJ. 2006. Protein composition of the intranuclear inclusions of FXTAS. *Brain*. 129(Pt 1):256–271.
- Jain A, Vale RD. 2017. RNA phase transitions in repeat expansion disorders. *Nature*. 546(7657):243–247.
- Jimenez-Sanchez M, Licitra F, Underwood BR, Rubinsztein DC. 2017. Huntington's disease: mechanisms of pathogenesis and therapeutic strategies. *Cold Spring Harb Perspect Med*. 7(7):a024240.
- Jovičić A, Mertens J, Boeynaems S, Bogaert E, Chai N, Yamada SB, Paul JW, Sun S, Herdy JR, Bieri G, et al. 2015. Modifiers of C9orf72 dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS. *Nat Neurosci*. 18(9):1226–1229.
- Klug A. 2010. The discovery of zinc fingers and their applications in gene regulation and genome manipulation. *Annu Rev Biochem*. 79:213–231.
- Konermann S, Lotfy P, Brideau NJ, Oki J, Shokhirev MN, Hsu PD. 2018. Transcriptome engineering with RNA-targeting type VI-D CRISPR effectors. *Cell*. 173(3):665–676 e614.
- Koole W, van Schendel R, Karambelas AE, van Heteren JT, Okihara KL, Tijsterman M. 2014. A polymerase theta-dependent repair pathway suppresses extensive genomic instability at endogenous G4 DNA sites. *Nat Commun*. 5: 3216.
- Kordasiewicz HB, Stanek LM, Wancewicz EV, Mazur C, McAlonis MM, Pytel KA, Artates JW, Weiss A, Cheng SH, Shihabuddin LS, et al. 2012. Sustained therapeutic reversal of Huntington's disease by transient repression of huntingtin synthesis. *Neuron*. 74(6):1031–1044.
- Kotowska-Zimmer A, Ostrovska Y, Olejniczak M. 2020. Universal RNAi triggers for the specific inhibition of mutant Huntingtin, atrophin-1, ataxin-3, and ataxin-7 expression. *Mol Ther Nucleic Acids*. 19:562–571.
- Kramer NJ, Haney MS, Morgens DW, Jovicic A, Couthouis J, Li A, Ousey J, Ma R, Bieri G, Tsui CK, et al. 2018. CRISPR-Cas9 screens in human cells and primary neurons identify modifiers of C9ORF72 dipeptide-repeat-protein toxicity. *Nat Genet*. 50(4):603–612.
- Kumari D, Usdin K. 2014. Polycomb group complexes are recruited to reactivated FMR1 alleles in Fragile X syndrome in response to FMR1 transcription. *Hum Mol Genet*. 23(24): 6575–6583.
- Kumari D, Usdin K. 2016. Sustained expression of FMR1 mRNA from reactivated fragile X syndrome alleles after treatment with small molecules that prevent trimethylation of H3K27. *Hum Mol Genet*. 25(17):3689–3698.
- La Spada AR, Wilson EM, Lubahn DB, Harding AE, Fischbeck KH. 1991. Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature*. 352(6330): 77–79.
- Ladd PD, Smith LE, Rabaia NA, Moore JM, Georges SA, Hansen RS, Hagerman RJ, Tassone F, Tapscott SJ, Filippova GN. 2007. An antisense transcript spanning the CGG repeat region of FMR1 is upregulated in premutation carriers but silenced in full mutation individuals. *Hum Mol Genet*. 16(24):3174–3187.
- Leavitt BR, Tabrizi SJ. 2020. Antisense oligonucleotides for neurodegeneration. *Science*. 367(6485):1428–1429.
- Lee YB, Chen HJ, Peres JN, Gomez-Deza J, Attig J, Stalekar M, Troakes C, Nishimura AL, Scotter EL, Vance C, et al. 2013. Hexanucleotide repeats in ALS/FTD form length-dependent RNA foci, sequester RNA binding proteins, and are neurotoxic. *Cell Rep*. 5(5):1178–1186.
- Liang XH, Sun H, Nichols JG, Crooke ST. 2017. RNase H1-dependent antisense oligonucleotides are robustly active in directing RNA cleavage in both the cytoplasm and the nucleus. *Mol Ther*. 25(9):2075–2092.
- Lieberman AP, Shakkottai VG, Albin RL. 2019. Polyglutamine repeats in neurodegenerative diseases. *Annu Rev Pathol*. 14:1–27.
- Lin Y, Dent SY, Wilson JH, Wells RD, Napierala M. 2010. R loops stimulate genetic instability of CTG/CAG repeats. *Proc Natl Acad Sci USA*. 107(2):692–697.
- Liu XS, Wu H, Krzisch M, Wu X, Graef J, Muffat J, Hnisz D, Li CH, Yuan B, Xu C, et al. 2018. Rescue of fragile X syndrome neurons by DNA methylation editing of the FMR1 gene. *Cell*. 172(5):979–992 e976.
- Lopez-Gonzalez R, Lu Y, Gendron TF, Karydas A, Tran H, Yang D, Petrucelli L, Miller BL, Almeida S, Gao FB. 2016. Poly(GR) in C9ORF72-related ALS/FTD compromises mitochondrial function and increases oxidative stress and DNA damage in iPSC-derived motor neurons. *Neuron*. 92(2):383–391.
- Ma L, Herren AW, Espinal G, Randol J, McLaughlin B, Martinez-Cerdeno V, Pessah IN, Hagerman RJ, Hagerman PJ. 2019. Composition of the intranuclear inclusions of fragile X-associated Tremor/Ataxia syndrome. *Acta Neuropathol Commun*. 7(1):143.
- Mankodi A, Urbinati CR, Yuan QP, Moxley RT, Sansone V, Krym M, Henderson D, Schalling M, Swanson MS, Thornton CA. 2001. Muscleblind localizes to nuclear foci of aberrant RNA in myotonic dystrophy types 1 and 2. *Hum Mol Genet*. 10(19):2165–2170.
- Mann DM, Rollinson S, Robinson A, Bennion Callister J, Thompson JC, Snowden JS, Gendron T, Petrucelli L, Masuda-Suzukake M, Hasegawa M, et al. 2013. Dipeptide repeat proteins are present in the p62 positive inclusions in patients with frontotemporal lobar degeneration and motor neurone disease associated with expansions in C9ORF72. *Acta Neuropathol Commun*. 1:68.
- Markmiller S, Soltanieh S, Server KL, Mak R, Jin W, Fang MY, Luo EC, Krach F, Yang D, Sen A, et al. 2018. Context-dependent and disease-specific diversity in protein interactions within stress granules. *Cell*. 172(3):590–604 e513.
- Matthaei M, Hribek A, Clahsen T, Bachmann B, Cursiefen C, Jun AS. 2019. Fuchs endothelial corneal dystrophy: clinical,

- genetic, pathophysiologic, and therapeutic aspects. *Annu Rev Vis Sci.* 5:151–175.
- Maziuk B, Ballance HI, Wolozin B. 2017. Dysregulation of RNA binding protein aggregation in neurodegenerative disorders. *Front Mol Neurosci.* 10:89.
- McFarland KN, Liu J, Landrian I, Gao R, Sarkar PS, Raskin S, Moscovich M, Gatto EM, Teive HA, Ochoa A, et al. 2013. Paradoxical effects of repeat interruptions on spinocerebellar ataxia type 10 expansions and repeat instability. *Eur J Hum Genet.* 21(11):1272–1276.
- McIntosh CS, Aung-Htut MT, Fletcher S, Wilton SD. 2019. Removal of the polyglutamine repeat of Ataxin-3 by redirecting pre-mRNA processing. *Int J Mol Sci.* 20:5434.
- McLoughlin HS, Moore LR, Paulson HL. 2020. Pathogenesis of SCA3 and implications for other polyglutamine diseases. *Neurobiol Dis.* 134:104635.
- McMurray CT. 2010. Mechanisms of trinucleotide repeat instability during human development. *Nat Rev Genet.* 11(11):786–799.
- Misra C, Lin F, Kalsotra A. 2018. Deregulation of RNA metabolism in microsatellite expansion diseases. *Adv Neurobiol.* 20:213–238.
- Moseley ML, Zu T, Ikeda Y, Gao W, Mosemiller AK, Daughters RS, Chen G, Weatherspoon MR, Clark HB, Ebner TJ, et al. 2006. Bidirectional expression of CUG and CAG expansion transcripts and intranuclear polyglutamine inclusions in spinocerebellar ataxia type 8. *Nat Genet.* 38(7):758–769.
- Nelles DA, Fang MY, O'Connell MR, Xu JL, Markmiller SJ, Doudna JA, Yeo GW. 2016. Programmable RNA tracking in live cells with CRISPR/Cas9. *Cell.* 165(2):488–496.
- Nelson DL, Orr HT, Warren ST. 2013. The unstable repeats—three evolving faces of neurological disease. *Neuron.* 77(5):825–843.
- Nguyen L, Cleary JD, Ranum LPW. 2019. Repeat-associated non-ATG translation: molecular mechanisms and contribution to neurological disease. *Annu Rev Neurosci.* 42:227–247.
- Nussbacher JK, Tabet R, Yeo GW, Lagier-Tourenne C. 2019. Disruption of RNA metabolism in neurological diseases and emerging therapeutic interventions. *Neuron.* 102(2):294–320.
- Oliveira AM, Cardoso SM, Ribeiro M, Seixas RS, Silva AM, Rego AC. 2015. Protective effects of 3-alkyl luteolin derivatives are mediated by Nrf2 transcriptional activity and decreased oxidative stress in Huntington's disease mouse striatal cells. *Neurochem Int.* 91:1–12.
- Onishi H, Kino Y, Morita T, Futai E, Sasagawa N, Ishiura S. 2008. MBNL1 associates with YB-1 in cytoplasmic stress granules. *J Neurosci Res.* 86(9):1994–2002.
- Orr HT, Zoghbi HY. 2007. Trinucleotide repeat disorders. *Annu Rev Neurosci.* 30:575–621.
- Panza F, Lozupone M, Seripa D, Daniele A, Watling M, Giannelli G, Imbimbo BP. 2020. Development of disease-modifying drugs for frontotemporal dementia spectrum disorders. *Nat Rev Neurol.* 16(4):213–228.
- Paul B, Montoya G. 2020. CRISPR-Cas12a: functional overview and applications. *Biomed J.* 43(1):8–17.
- Paul S, Dansithong W, Jog SP, Holt I, Mittal S, Brook JD, Morris GE, Comai L, Reddy S. 2011. Expanded CUG repeats dysregulate RNA splicing by altering the stoichiometry of the muscleblind 1 complex. *J. Biol. Chem.* 286(44):38427–38438.
- Paulson H. 2018. Repeat expansion diseases. *Handb Clin Neurol.* 147:105–123.
- Perbellini R, Greco S, Sarra-Ferraris G, Cardani R, Capogrossi MC, Meola G, Martelli F. 2011. Dysregulation and cellular mislocalization of specific miRNAs in myotonic dystrophy type 1. *Neuromuscul Disord.* 21(2):81–88.
- Pettersson OJ, Aagaard L, Jensen TG, Damgaard CK. 2015. Molecular mechanisms in DM1 – a focus on foci. *Nucleic Acids Res.* 43(4):2433–2441.
- Pinto BS, Saxena T, Oliveira R, Mendez-Gomez HR, Cleary JD, Denes LT, McConnell O, Arboleda J, Xia G, Swanson MS, et al. 2017. Impeding transcription of expanded microsatellite repeats by deactivated Cas9. *Mol Cell.* 68(3):479–490 e475.
- Quenault T, Lithgow T, Traven A. 2011. PUF proteins: repression, activation and mRNA localization. *Trends Cell Biol.* 21(2):104–112.
- Reddy K, Jenquin JR, McConnell OL, Cleary JD, Richardson JI, Pinto BS, Haerle MC, Delgado E, Planco L, Nakamori M, et al. 2019. A CTG repeat-selective chemical screen identifies microtubule inhibitors as selective modulators of toxic CUG RNA levels. *Proc Natl Acad Sci USA.* 116(42):20991–21000.
- Rinaldi C, Wood MJA. 2018. Antisense oligonucleotides: the next frontier for treatment of neurological disorders. *Nat Rev Neurol.* 14(1):9–21.
- Rodriguez CM, Wright SE, Kearse MG, Haenfler JM, Flores BN, Liu Y, Ifrim MF, Glineburg MR, Krans A, Jafar-Nejad P, et al. 2020. A native function for RAN translation and CGG repeats in regulating fragile X protein synthesis. *Nat Neurosci.* 23(3):386–397.
- Rohilla KJ, Gagnon KT. 2017. RNA biology of disease-associated microsatellite repeat expansions. *Acta Neuropathol Commun.* 5(1):63.
- Roselli F, Caroni P. 2015. From intrinsic firing properties to selective neuronal vulnerability in neurodegenerative diseases. *Neuron.* 85(5):901–910.
- Rudnicki DD, Holmes SE, Lin MW, Thornton CA, Ross CA, Margolis RL. 2007. Huntington's disease-like 2 is associated with CUG repeat-containing RNA foci. *Ann Neurol.* 61(3):272–282.
- Sathasivam K, Neueder A, Gipson TA, Landles C, Benjamin AC, Bondulich MK, Smith DL, Faull RL, Roos RA, Howland D, et al. 2013. Aberrant splicing of HTT generates the pathogenic exon 1 protein in Huntington disease. *Proc Natl Acad Sci USA.* 110(6):2366–2370.
- Schmidt MHM, Pearson CE. 2016. Disease-associated repeat instability and mismatch repair. *DNA Repair (Amst).* 38:117–126.
- Sellier C, Freyermuth F, Tabet R, Tran T, He F, Ruffenach F, Alunni V, Moine H, Thibault C, Page A, et al. 2013. Sequestration of DROSHA and DGCR8 by expanded CGG RNA repeats alters microRNA processing in fragile X-associated tremor/ataxia syndrome. *Cell Rep.* 3(3):869–880.
- Sellier C, Rau F, Liu Y, Tassone F, Hukema RK, Gattoni R, Schneider A, Richard S, Willemsen R, Elliott DJ, et al. 2010. Sam68 sequestration and partial loss of function are associated with splicing alterations in FXTAS patients. *Embo J.* 29(7):1248–1261.
- Selvaraj BT, Livesey MR, Zhao C, Gregory JM, James OT, Cleary EM, Chouhan AK, Gane AB, Perkins EM, Dando O, et al. 2018. C9ORF72 repeat expansion causes vulnerability

- of motor neurons to Ca²⁺-permeable AMPA receptor-mediated excitotoxicity. *Nat Commun.* 9(1):347.
- Shi KY, Mori E, Nizami ZF, Lin Y, Kato M, Xiang S, Wu LC, Ding M, Yu Y, Gall JG, et al. 2017. Toxic PRn poly-dipeptides encoded by the C9orf72 repeat expansion block nuclear import and export. *Proc Natl Acad Sci USA.* 114(7): E1111–E1117.
- Shimada MK, Sanbonmatsu R, Yamaguchi-Kabata Y, Yamasaki C, Suzuki Y, Chakraborty R, Gojobori T, Imanishi T. 2016. Selection pressure on human STR loci and its relevance in repeat expansion disease. *Mol Genet Genomics.* 291(5): 1851–1869.
- Shin J, Charizanis K, Swanson MS. 2009. Pathogenic RNAs in microsatellite expansion disease. *Neurosci Lett.* 466(2): 99–102.
- Siboni RB, Nakamori M, Wagner SD, Struck AJ, Coonrod LA, Harriott SA, Cass DM, Tanner MK, Berglund JA. 2015. Actinomycin D specifically reduces expanded CUG repeat RNA in myotonic dystrophy models. *Cell Rep.* 13(11): 2386–2394.
- Simhadri VL, McGill J, McMahon S, Wang J, Jiang H, Sauna ZE. 2018. Prevalence of pre-existing antibodies to CRISPR-associated nuclease Cas9 in the USA population. *Mol Ther Methods Clin Dev.* 10:105–112.
- Smargon AA, Shi YJ, Yeo GW. 2020. RNA-targeting CRISPR systems from metagenomic discovery to transcriptomic engineering. *Nat Cell Biol.* 22(2):143–150.
- Sulej AA, Tuszynska I, Skowronek KJ, Nowotny M, Bujnicki JM. 2012. Sequence-specific cleavage of the RNA strand in DNA-RNA hybrids by the fusion of ribonuclease H with a zinc finger. *Nucleic Acids Res.* 40(22):11563–11570.
- Swinnen B, Robberecht W, Van Den Bosch L. 2020. RNA toxicity in non-coding repeat expansion disorders. *Embo J.* 39(1):e101112.
- Tassone F, Iwahashi C, Hagerman PJ. 2004. FMR1 RNA within the intranuclear inclusions of fragile X-associated tremor/ataxia syndrome (FXTAS). *RNA Biol.* 1(2):103–105.
- Thornton CA. 2014. Myotonic dystrophy. *Neurol Clin.* 32(3): 705–719, viii.
- Todd PK, Paulson HL. 2010. RNA-mediated neurodegeneration in repeat expansion disorders. *Ann Neurol.* 67(3): 291–300.
- Toonen LJA, Rigo F, van Attikum H, van Roon-Mom WMC. 2017. Antisense oligonucleotide-mediated removal of the polyglutamine repeat in spinocerebellar Ataxia type 3 mice. *Mol Ther Nucleic Acids.* 8:232–242.
- Urbanek MO, Fiszer A, Krzyzosiak WJ. 2017. Reduction of Huntington's disease RNA foci by CAG repeat-targeting reagents. *Front Cell Neurosci.* 11:82.
- Urbanek MO, Krzyzosiak WJ. 2016. RNA FISH for detecting expanded repeats in human diseases. *Methods.* 98: 115–123.
- van Cruchten RTP, Wieringa B, Wansink DG. 2019. Expanded CUG repeats in DMPK transcripts adopt diverse hairpin conformations without influencing the structure of the flanking sequences. *RNA.* 25(4):481–495.
- Wang LL, Wansleben C, Zhao SL, Miao P, Paschen W, Yang W. 2014. SUMO2 is essential while SUMO3 is dispensable for mouse embryonic development. *EMBO Rep.* 15(8): 878–885.
- Ward ME, Taubes A, Chen R, Miller BL, Sephton CF, Gelfand JM, Minami S, Boscardin J, Martens LH, Seeley WW, et al. 2014. Early retinal neurodegeneration and impaired Ran-mediated nuclear import of TDP-43 in progranulin-deficient FTLD. *J Exp Med.* 211(10):1937–1945.
- Wenzel HJ, Hunsaker MR, Greco CM, Willemsen R, Berman RF. 2010. Ubiquitin-positive intranuclear inclusions in neuronal and glial cells in a mouse model of the fragile X premutation. *Brain Res.* 1318:155–166.
- Wheeler TM, Leger AJ, Pandey SK, MacLeod AR, Nakamori M, Cheng SH, Wentworth BM, Bennett CF, Thornton CA. 2012. Targeting nuclear RNA for in vivo correction of myotonic dystrophy. *Nature.* 488(7409):111–115.
- Wheeler TM, Sobczak K, Lueck JD, Osborne RJ, Lin X, Dirksen RT, Thornton CA. 2009. Reversal of RNA dominance by displacement of protein sequestered on triplet repeat RNA. *Science.* 325(5938):336–339.
- Xu X, Hulshoff MS, Tan X, Zeisberg M, Zeisberg EM. 2020. CRISPR/Cas derivatives as novel gene modulating tools: possibilities and in vivo applications. *Int J Mol Sci.* 21: 3038.
- Yamada SB, Gendron TF, Niccoli T, Genuth NR, Grosely R, Shi Y, Glaria I, Kramer NJ, Nakayama L, Fang S, et al. 2019. RPS25 is required for efficient RAN translation of C9orf72 and other neurodegenerative disease-associated nucleotide repeats. *Nat Neurosci.* 22(9):1383–1388.
- Yang P, Wang Y, Hoang D, Tinkham M, Patel A, Sun MA, Wolf G, Baker M, Chien HC, Lai KN, et al. 2017. A placental growth factor is silenced in mouse embryos by the zinc finger protein ZFP568. *Science.* 356(6339):757–759.
- Yrigollen CM, Durbin-Johnson B, Gane L, Nelson DL, Hagerman R, Hagerman PJ, Tassone F. 2012. AGG interruptions within the maternal FMR1 gene reduce the risk of offspring with fragile X syndrome. *Genet Med.* 14(8): 729–736.
- Zhang H, Elbaum-Garfinkle S, Langdon EM, Taylor N, Occhipinti P, Bridges AA, Brangwynne CP, Gladfelter AS. 2015a. RNA controls PolyQ protein phase transitions. *Mol Cell.* 60(2):220–230.
- Zhang K, Donnelly CJ, Haeusler AR, Grima JC, Machamer JB, Steinwald P, Daley EL, Miller SJ, Cunningham KM, Vidensky S, et al. 2015b. The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. *Nature.* 525(7567):56–61.
- Zhang N, Ashizawa T. 2017. RNA toxicity and foci formation in microsatellite expansion diseases. *Curr Opin Genet Dev.* 44:17–29.
- Zhang YJ, Gendron TF, Grima JC, Sasaguri H, Jansen-West K, Xu YF, Katzman RB, Gass J, Murray ME, Shinohara M, et al. 2016. C9ORF72 poly(GA) aggregates sequester and impair HR23 and nucleocytoplasmic transport proteins. *Nat Neurosci.* 19(5):668–677.
- Zu T, Gibbens B, Doty NS, Gomes-Pereira M, Huguet A, Stone MD, Margolis J, Peterson M, Markowski TW, Ingram MA, et al. 2011. Non-ATG-initiated translation directed by microsatellite expansions. *Proc Natl Acad Sci USA.* 108(1): 260–265.