HHS Public Access

Author manuscript

Annu Rev Neurosci. Author manuscript; available in PMC 2019 August 08.

Published in final edited form as:

Annu Rev Neurosci. 2019 July 08; 42: 227–247. doi:10.1146/annurev-neuro-070918-050405.

Repeat Associated Non-ATG translation: molecular mechanisms and contribution to neurologic disease

Lien Nguyen[1,2,4], **John Douglas Cleary**[1,2,4], **Laura P.W. Ranum**[1,2,3,4,5,*]

[1]Center for NeuroGenetics, University of Florida, Gainesville, FL, USA, 32610

[2]Departments of Molecular Genetics & Microbiology, University of Florida, Gainesville, FL, USA, 32610

[3]Departments of Neurology, College of Medicine, University of Florida, Gainesville, FL, USA, 32610

[4]Genetics Institute, University of Florida, Gainesville, FL, USA, 32610

[5]McKnight Brain Institute, University of Florida, Gainesville, FL, USA, 32610

Abstract

Microsatellite mutations involving the expansion of tri-, tetra-, penta- or hexanucleotide repeats cause more than 40 different neurological disorders. Although, traditionally, the position of the repeat within or outside of an open reading frame has been used to focus research into disease mechanisms on protein loss-of-function, protein gain-of-function or RNA gain-of- function, the discoveries of bidirectional transcription and repeat associated non-ATG (RAN) have blurred these distinctions. Here we review what is known about RAN proteins in disease, the mechanisms by which they are produced and the novel therapeutic opportunities they provide.

Keywords

translation; microsatellite repeat diseases; RAN translation; toxicity; mechanism; therapy

1. EXPANDING MECHANISMS IN MICROSATELLITE EXPANSION DISORDERS

Microsatellite expansion mutations cause more than 40 neurological and neuromuscular diseases including myotonic dystrophy (DM), *C9orf72* amyotrophic lateral sclerosis/ frontotemporal dementia (C9-ALS/FTD), Huntington disease (HD), and multiple types of spinocerebellar ataxia (SCA). These tri-, tetra-, penta- or hexanucleotide repeats can expand when passed from generation to generation and also during an individual's lifespan through somatic instability (Cleary et al 2006, Lopez Castel et al 2010). These mutations can be located in protein coding regions, within introns or in the 5' or 3' untranslated regions

*Corresponding Author: Laura P.W. Ranum: Director, Center for NeuroGenetics, Kitzman Family Professor of Molecular Genetics and Microbiology, College of Medicine, University of Florida, 2033 Mowry Road, PO Box 103610, Gainesville, FL 32610-3610; ranum@ufl.edu; Tel. (352) 294-5209; Fax. (352) 273.8284.

(UTRs) of their respective genes (Gijselinck et al 2016, Harley et al 1992, Trottier et al 1994). In general, repeats located in non-coding gene regions can be 10-100 times longer than those found in open reading frames. In most expansion diseases longer repeat lengths correlate with decreased ages of onset and increased disease severity (Cleary et al 2018, Cleary & Ranum 2014, Gijselinck et al 2016, Harley et al 1992, Trottier et al 1994).

1.1. RNA gain-of-function, protein gain-of-function and protein loss-of-function

Traditionally, repeat expansion mutations have been thought to cause disease through toxic gain-of-function (GOF) properties of mutant proteins or RNAs or through protein loss-of-function (LOF) mechanisms. For dominantly inherited disorders, microsatellite expansion mutations located in non-coding regions have been typically considered to involve RNA GOF mechanisms. In this mechanism, GC-rich expansion transcripts form secondary structures that can sequester disease-specific RNA binding proteins and disrupt their normal functions. The best example of this RNA GOF mechanism is in myotonic dystrophy types 1 and 2. Hairpin forming CUG and CCUG expansion RNAs transcribed from expansion mutations located in the *DMPK* 3'UTR or *CNBP* intron 1, sequester muscleblind-like (MBNL) proteins in nuclear RNA foci resulting in the loss of MBNL function and downstream RNA processing abnormalities (Kanadia et al 2006, Lander et al 2001, Miller et al 2000, Mohan et al 2014, Nelson et al 2013, Scotti & Swanson 2016, Tóth et al 2000, Wheeler & Thornton 2007). In contrast, expansions located in protein coding regions can take on new functions and also disrupt normal functions proteins with examples including Huntington disease and spinocerebellar ataxia type 1 (Labbadia & Morimoto 2013, Orr & Zoghbi 2007). Recessively inherited diseases involving protein LOF mechanisms include the intronic GAA and 5' UTR CGG repeats which decrease expression of the *FXN* and *FMR1* genes involved in Friedreich ataxia (FA) and Fragile X syndrome (FMR), respectively (Bidichandani et al 1998, Campuzano et al 1996, Hagerman 2013, Hagerman & Hagerman 2013). The discoveries that many of these expansion mutations are bidirectionally expressed and that both coding and non-coding expansion mutations can express proteins in multiple reading frames by repeat associated non-ATG (RAN) translation has changed our fundamental understanding of the molecular mechanisms of repeat expansion disorders.

1.2. Bidirectional transcription

Bidirectional transcription at repeat expansion loci was first described for DM1 (Cho et al 2005) and SCA8 (Moseley et al 2006). In DM1, small RNAs expressed in the antisense direction were detected (Cho et al 2005) and in SCA8, antisense transcripts containing CAG expansion were shown to express a polyGln expansion protein that accumulates in human and mouse autopsy brains (Moseley et al 2006). Although most research and therapeutic strategies are focused on expansion RNAs or proteins expressed in the sense direction, antisense transcripts have also been reported in a growing number of diseases including FXTAS, HD, HDL2, SCA7, DM2, *C9orf72* ALS/FTD (Cleary et al 2018).

1.3. RAN translation

The discovery of repeat associated non-ATG translation added further complexity to possible mechanisms involved in microsatellite expansion disorders. Repeat associated non-AUG (RAN) translation was first discovered while studying the CAG·CTG expansion mutations in

spinocerebellar ataxia 8 (SCA8) and myotonic dystrophy type 1 (DM1). In 2011, Zu et al., (Zu et al 2011) showed in transfected cells that both CAG and CUG expansion transcripts can produce unexpected homopolymeric proteins in all three reading frames in the absence of an AUG start codon with polyGln, polyAla and polySer expressed from CAG expansions and polyLeu, polyAla and polyCys proteins expressed across CUG expansion transcripts. Additionally these authors provided evidence for the accumulation of a novel polyAla and polyGln RAN proteins in SCA8 and DM1, respectively (Zu et al 2011). RAN proteins are now known to accumulate in eight distinct neurodegenerative and neuromuscular diseases and addition to SCA8 and DM1, have been reported in fragile-X tremor ataxia syndrome (FXTAS) (Todd et al 2013), Huntington disease (HD) (Bañez-Coronel et al 2015), *C9orf72* amyotrophic lateral sclerosis and frontotemporal dementia (C9 ALS/FTD) (Ash et al 2013, Mori et al 2013, Zu et al 2013), and recently in myotonic dystrophy type 2 (DM2) (Zu et al 2017), Fuchs' endothelial corneal dystrophy (FECD) (Soragni et al 2018), and spinocerebellar ataxia type 31 (SCA31) (Ishiguro et al 2017) (Table 1).

Taken together, bidirectional transcription and RAN translation blur the previous mechanistic lines drawn between protein-coding and non-coding classifications of microsatellite expansion mutations. Mutations previously considered to be non-coding and suggested to involve an RNA GOF mechanism (O'Rourke & Swanson 2009, Ranum & Cooper 2006, Todd & Paulson 2010) can express up to six mutant proteins e.g. *C9orf72* ALS/FTD (Cleary et al 2018, Cleary & Ranum 2013, Cleary & Ranum 2014). Similarly, expansion mutations located in ORFs can express up to five additional sense and antisense RAN proteins (Bañez-Coronel et al 2015). Additionally, many of the polyGln expansion disorders are known to express both sense CAG and antisense CUG expansion transcripts with could trigger RNA GOF features similar to those previously reported DM1 and SCA8. Now that we know that many of repeat expansion mutations can undergo this type of promiscuous expression it is critical to determine which expansion disorders express RAN proteins, when and where they accumulate and to understand their contribution to disease.

2. WHEN AND WHERE RAN PROTEINS ACCUMULATE: CLUES FROM THE SCENE OF THE CRIME

Comparisons across the growing number of microsatellite expansion disorders now known to express RAN proteins has revealed common themes. First, RAN proteins have been primarily reported to accumulate in the CNS (Cleary et al 2018, Cleary & Ranum 2014). Within the brain, selected RAN proteins have shown to have distinct patterns of accumulation. For example, in DM2, the LPAC and QAGR RAN proteins are expressed in grey and white matter, respectively (Zu et al 2017). Similarly, in HD, the ATG-initiated polyGln protein and four different HD RAN proteins are found in the caudate and putamen, but the HD RAN proteins show additional accumulation in the white matter bundle regions of the caudate and putamen (Bañez-Coronel et al 2015). Additionally, Ayhan et al. recently described SCA8 polySer RAN protein that preferentially accumulates in cerebellar and cortical white matter regions (Ayhan et al 2018). Curiously, polySer accumulates primarily in oligodendrocyte-like cells in the white matter regions while the SCA8 polyGln and polyAla are found in Purkinje cells and other neurons. These patterns of accumulation are

not explained by differences in transcript expression, as each of these proteins is expressed from the same CAG expansion transcript. These data suggest expression of RAN proteins may be differentially regulated and that cell specific factors and/or frame specific differences in flanking sequence may modulate RAN translation. In SCA8, eIF3F is elevated in white matter regions that show high levels of SCA8 RAN polySer accumulation and knocking down eIF3F expression decreases polySer accumulation. In contrast, eIF3F knockdown did not decrease translation of proteins with an upstream ATG or ATG-like sequence. This or similar differential effects that depend on sequence context may explain why some RAN proteins preferentially accumulate in white matter regions.

An additional feature of RAN protein diseases is that the patterns of accumulation are highly variable and often involve focal sites of accumulation and pathology. The patchy accumulation is evident in C9orf72 ALS/FTD, DM2, HD and SCA8 (Ash et al 2013, Ayhan et al 2018, Bañez-Coronel et al 2015, Mori et al 2013, Moseley et al 2006, Zu et al 2017, Zu et al 2013). For example, in HD, RAN proteins can be hard to detect in some cerebellar regions and highly abundant in other regions (Bañez-Coronel et al 2015). This variation is seen between patients as well as when comparing different sections of fixed tissue from the same brain regions in the same patient. In DM2 and SCA8, sites of DM2 QAGR and SCA8 polySer RAN protein accumulation occur in regions with neurodegeneration and white matter abnormalities, respectively(Ayhan et al 2018, Zu et al 2017). While it is not yet clear what the triggers are for the focal accumulation of RAN proteins, it is possible that focal regions of stress (Cheng et al 2018, Green et al 2017, Tabet et al 2018), a reported modulator of RAN translation, may lead to focal increases of RAN protein expression.

In *C9orf72* ALS/FTD, six different RAN proteins are expressed and accumulate in C9 human autopsy brains (Ash et al 2013, Mori et al 2013, Zu et al 2013). Although histological studies frequently report GA dipeptide protein aggregates as the most abundant, differences in antibodies, antibody affinities and antigen retrieval methods required for individual C9 RAN proteins. More quantitative and objective methods will be required to definitively measure the expression levels of individual C9 RAN proteins and assess the impact of therapeutic interventions on their levels. A particular challenge for the field has been to detect antisense C9 RAN proteins, especially in mice. In addition to the brain, GP proteins are also found in cerebrospinal fluid (CSF) and peripheral blood mononuclear cells (PBMCs). In human patients GP levels in the CSF are relatively stable (Gendron et al 2017, Lehmer et al 2017) and in mice have been shown to correlate with aggregates in the brain and to be reduced when treated with ASOs (Gendron et al 2017), making GP proteins an important pharmacodynamic biomarker that can be used to assess therapeutic efficacy (Gendron et al 2017, Westergard et al 2016). More work is needed to have a complete picture of when and where C9 RAN proteins accumulate.

Another important question is when are RAN proteins first expressed and how long it takes before they accumulate and contribute to disease. For many diseases, increases in repeat length result in increased severity and decreased ages of onset. While RAN proteins can be expressed from relatively short (e.g. HD, 40~100), and much longer (e.g. DM2 75-11,000) expansion mutations, in general RAN translation increases with repeat length (Ash et al 2013, Mori et al 2013, Sellier et al 2017, Todd et al 2013, Zu et al 2011, Zu et al 2013).

These data suggest that patients with longer repeat length may express RAN proteins earlier and at higher levels. Consistent with this relationship, cases of HD with longer repeat lengths (~100) that are associated with juvenile onset, shows much higher levels of RAN proteins in human autopsy samples than the adult onset cases with shorter repeats (Bañez-Coronel et al 2015). Additionally, in both SCA8 and C9-ALS/FTD mouse models, RAN protein aggregates are detected prior to overt onset of symptoms and their accumulation increases with age and disease severity (Ayhan et al 2018, Liu et al 2016). Furthermore, RAN proteins but not TDP-43 pathology are observed prior to the onset of symptom in several *C9orf72* ALS/FTD patients (Vatsavayai et al 2016) supporting the pre-symptomatic expression of RAN proteins. While these data are consistent with a role for RAN proteins as a driver of disease, the interplay between expansion RNAs and RAN translation is likely to also influence when and where RAN proteins will first be expressed. This relationship was recently explored in DM2, where RAN translation of CCUG expansions was shown to be modulated by MBNL1 (Zu et al 2017). The authors proposed a two-stage model of disease, the Nuclear Sequestration Failure Model, which posits that initial nuclear retention of expansion RNAs and RBP depletion is followed by a later disease phase where expansion RNAs are exported into the cytoplasm and undergo RAN translation, worsening disease. Understanding when RAN proteins first appear in microsatellite diseases and if and at what stage their accumulation contributes to individual diseases will be important for the development of effective therapies and the timing of their application in patients.

3. RAN PROTEIN TOXICITY

Since their initial discovery, there has been an enormous effort to understand the role of RAN proteins in disease. Studies in which expanded repeats are overexpressed in human cell lines (e.g. HEK293T, T98) have shown that the expression of several types of repeats are toxic to cells but do not distinguish the contribution of RNA vs. protein toxicity (Bañez-Coronel et al 2015, Soragni et al 2018, Stopford et al 2017, Zu et al 2017, Zu et al 2011). Driving higher levels of expression of specific RAN proteins in a particular reading frames with ATG-initiation codon has also been shown to increase cell death and decrease cell viability in comparison to non-ATG or vector controls (Todd et al 2013, Zu et al 2011, Zu et al 2013). Experiments showing that overexpression of individual RAN proteins in cells are toxic independent of the expansion RNA has been confirmed using non-hairpin forming alternative codon sequences (May et al 2014, Wen et al 2014, Zu et al 2017, Zu et al 2013). This approach has demonstrated that the RAN proteins from multiple repeat expansions (CAG, CUG, CCUG, CAGG, G4C2, G2C4) are toxic independent of the toxic effects of the expansion RNAs (Bañez-Coronel et al 2015, May et al 2014, Zu et al 2017, Zu et al 2011, Zu et al 2013) when overexpressed in cells. While RAN protein are toxic in these simple toxicity and viability assays using cell-based models, both RNA transcript and protein expression levels are often artificially high, making it critical to study toxicity in complex systems under more physiological conditions. An examination of toxicity studies for FXTAS, HD and C9 ALS/FTD provide examples of the complexity of examining RAN protein toxicity.

3.1. Toxicity of FXTAS RAN proteins

The toxicity of the FMR-polyGly in FXTAS has been explored across several models systems where phenotypes are strongly linked with neurodegeneration, such as rough-eye phenotypes for *Drosophia* models or gait abnormalities, anxiety, cage behavior and neurodegenerative changes for mouse models. To demonstrate toxicity of the FMR polyGly, Todd and co-workers generated *Drosophila* models overexpressing repeats in different sequence contexts capable of producing CGG expansion RNAs and/or FMR-polyGly protein (Todd et al 2013). Toxicity, as assessed by rough-eye phenotype, increased with increased production of FMR-polyGly and was suppressed when protein production was prevented. More recently, Sellier and co-workers generated two mouse lines: one expressing both expanded CGG transcripts and FMR polyGly, and another line containing a mutant 5' UTR that only expresses CGG transcripts (Sellier et al 2017). The mice expressing both RNA and FMR polyGly but not the one expressing just RNA, developed locomotor deficits including decreased mobility and obesity and died at 10 months of age. Using a doxycycline inducible mouse model, Castro and co-workers showed that expression of 90 CGG repeats outside of the context of the *FMR1* gene was sufficient to recapitulate FXTAS-like behavioral phenotypes (e.g. seizures, anxiety, and motor deficits) (Castro et al 2017). Anxiety-like phenotypes but not motor defects in mice were reversed immediately after stopping the expression of this transgene. These changes were accompanied by a reduction of ubiquitinated FMR polyGly inclusions in hippocampal regions but not in cerebellum. These results suggest a contribution of FMR polyGly to neurodegeneration in this overexpression model, however the contributions of other RAN protein species and the expansion RNAs were not excluded in this model. The toxicity of the FMR-polyGly, and to a lesser extent FMR-polyAla, has been linked with ubiquitin-proteasome system (UPS) impairment (Oh et al 2015). FMR polyGly was also found to recruit and alter the cellular distribution of Lap2β, which plays a role in organizing the nuclear lamina architecture, in FTXAS patients and iPSC-derived neurons (Sellier et al 2017). The antisense FMR polyPro and polyAla were also shown to accumulates in FXTAS brain tissue (Krans et al 2016) and likely contribute to protein toxicity in FXTAS. Given the number of FXTAS RAN proteins identified in patient brain tissue, understanding the toxicity and contribution of individual RAN proteins will be important for understanding the etiology of the disease but also for the development of effective therapeutic treatment strategies.

3.2. Huntington disease and polyGlutamine toxicity

The toxicity of the polyGln expansion protein, mutant huntingtin, has been a major focus in Huntington and other polyGln disorders, and has been reviewed extensively elsewhere (Bates et al 2015, Figiel et al 2012, Switonski et al 2012, Underwood & Rubinsztein 2008, Walsh et al 2005). PolyGln adopts amyloid-like fibril structure *in vitro* and in mammalian cells, and this structure is necessary for its toxicity (Bevivino & Loll 2001, Poirier et al 2005, Schaffar et al 2004, Scherzinger et al 1999). Recent findings on the structure of polyGln in expanded Huntingtin exon 1 in intact neurons showed that amyloid-like polyGln fibrils interacts with the endomembrane (Bauerlein et al 2017). Interactions of polyGln fibrils with endoplasmic reticulum (ER) alters ER organization and dynamics, which may play a role in the increased ER stress previously described in cellular and mouse models of HD (Duennwald & Lindquist 2008, Jiang et al 2016b, Kouroku et al 2002, Leitman et al

2013). While it is clear that the mutant huntingtin protein is toxic, four additional HD RAN proteins were recently shown to also accumulate in human autopsy samples. These HD RAN proteins are toxic to cells and found in brain regions with neurodegenerative changes suggesting that they may also contribute to disease (Bañez-Coronel et al 2015). A more complete understanding of the expression, distribution and contribution of sense and antisense RAN proteins and RNAs to HD will be required to understand the pathophysiology of the disease. Additionally, the possibility that RAN proteins may also be expressed in the other eight "polyGln" disorders needs to be investigated.

3.3. Toxicity and C9ORF72-ALS/FTD

A tremendous amount of work has been done to study the toxicity and contribution of individual dipeptide repeat (DPR) proteins, produced from the *C9orf72* G4C2 repeats, to ALS/FTD phenotypes. Toxicity studies have shown marked differences between the five DPR repeat motifs. GR and PR proteins have been described as highly toxic species due to their arginine-rich and shared biophysical properties. These properties favor binding to low complexity domains (LCDs) including LCDs found in hnRNPA1/2 and directly interacting with TIA-1, a stress granule marker) (Kwon et al 2014, Lee et al 2016, Lin et al 2016). Exposure of human cell lines to synthetic PR_{20} and GR_{20} resulted in significant decreases in cell viability (Kwon et al 2014). The toxicity of synthetic PR_{20} and GR_{20} was explained by cellular uptake of these peptides followed by their migration to nucleoli and disruption in pre-mRNA splicing and RNA biogenesis possibly through the binding to the LCD of hnRNPA2 (Kwon et al 2014, Lin et al 2016). In addition, overexpression of $(PR)_{50}$ but not $(GA)_{50}$ caused substantial increase in death of primary cortical, motor neuron cultures, and iPSC-derived neurons (ref needed). Overexpression of $(PR)_{50}$ was also shown to lead to its accumulation in nucleoli and correlate with stress granule formation (Wen et al 2014). To dissect RNA GOF and RAN toxicity in C9 ALS/FTD, Mizielinska and co-workers developed several *Drosophila* lines that expressed pure G_4C_2 repeats or stop-codon interrupted sequences that expressed RNA but not RAN proteins. Although the stop codon interruptions did not alter the G-quadruplex structures of RNA transcripts and foci formation, eye degeneration and lethality were observed only for lines expressing the pure repeats (Mizielinska et al 2014a). In addition, the work from several research groups showed that *Drosophila* lines expressing transcript containing alternative codons for arginine containing PR and GR proteins but not GP, GA or PA RAN proteins showed neurodegenerative eye phenotypes and a significant increase in lethality (Mizielinska et al 2014a, Wen et al 2014). More recently, Zhang and co-workers (Zhang et al 2018) showed that GFP-GR_{100} mice develop age-dependent neurodegeneration, brain atrophy, and motor and memory deficits. In this AAV mouse model, GFP-GR_{100} protein was found to co-localize with ribosomal subunit and initiation factor eIF3η, suggesting polyGR could impair ribosome function and protein translation. The mechanism of polyGR toxicity may be related to its ability to induce stress granule formation and impaired stress granule disassembly (Kwon et al 2014, Lee et al 2016, Zhang et al 2018). Toxicity may also be related to post-translation modifications of arginine-containing C9 RAN proteins as poly-GR colocalizes with asymmetrical dimethylarginine (aDMA) immunoreactivity in regions of neurodegeneration (Sakae et al 2018). Dimethylarginine modifications are associated with elevated levels of protein degradation found in neurodegenerative disorders. C9 patients also

accumulate symmetric arginine dimethylated proteins which colocalize with p62 suggestive of possible defects in stress granule degradation in these patients (Chitiprolu et al 2018). Additionally, $(GR)_{80}$ has been shown to dysregulate mitochondrial function, possibly by binding to mitochondrial ribosomal proteins which causes increased oxidative stress and DNA damage in iPSC-derived spinal motor neurons (Lopez-Gonzalez et al 2016). While GR and PR may be highly toxic, it is important not to overlook the contribution of other C9 RAN proteins.

3.3.1 PolyGA toxicity—The C9 polyGA dipeptide repeat (DPR) protein has been shown to have only moderate toxicity in various cell culture, zebrafish and mouse models compared to polyGR and polyPR (Kwon et al 2014, Lee et al 2016, Lin et al 2016, May et al 2014, Ohki et al 2017, Zhang et al 2016, Zhang et al 2014). Nevertheless, AAV delivery and overexpression of GFP-$(GA)_{50}$ in the CNS in mice produces a number of neurological phenotypes including brain atrophy, neuronal loss, astrogliosis and behavioral deficits (including hyperactivity, anxiety-like behavior, motor and cognitive defects) (Zhang et al 2016). Unlike other C9 DPRs, polyGA forms cytoplasmic filamentous aggregates in C9 patient tissue (Zhang et al 2014) that have recently been shown to be composed of highly packed twisted ribbons (Guo et al 2018). *In vitro,* short GA peptides of 6 or 15 repeats adopt amyloid-like structure that are toxic to neurons (Chang et al 2016, Flores et al 2016). Disruption of these amyloid structures in cell culture by inclusion of a proline residue after every 5 GA repeats in a GFP-$(GA)_{50}$ construct resulted in a more diffuse signal and decreased its toxicity. Toxicity of the polyGA protein may be related to the interaction and recruitment of other proteins and structures within cells, including proteasome subunits (e.g. PSMB6, PSMC4), UPS-related proteins HR23A/B, the trafficking protein Unc119 and other nucleocytoplasmic transport proteins (Guo et al 2018, Khosravi et al 2017, May et al 2014, Zhang et al 2016, Zhang et al 2014). In cultured mouse neurons, polyGA expression results in caspase-3 activation, decreased proteasome activity and increased levels of binding immunoglobulin protein (BIP), phosphorylated protein kinase RNA (PKR)-like ER kinase (PERK), and transcription factor C/EBP homologous protein (CHOP) (Zhang et al 2014). The activation of endoplasmic reticulum (ER) stress by polyGA is further supported by the fact that ER stress inhibitors (salubrinal and TUDCA) provide protection against poly(GA)-induced toxicity. In addition, C9 polyGA is believed to act as a seeding template for other C9 RAN protein aggregation. For example, $(GR)_{80}$ was recruited to cytoplasmic $(GA)_{80}$ aggregates in HeLa cells, iPSC-derived neurons and flies, which partially suppressed the Notch pathway-associated toxicity of $(GR)_{80}$ in flies (Yang et al 2015). In addition, co-expression of polyGA induces the aggregation of polyGP and polyPA, which normally show diffuse signals. Additionally poly GP and polyGA colocalize as aggregates in human autopsy brains (Lee et al 2017), further supporting a role for polyGA in the recruitment of other proteins in the human disease.

3.3.2 Toxicity of other DPR proteins—Toxicity studies using cellular and animal models indicate C9 polyGP and polyPA are likely the least toxic C9 DPR species. These two proteins are uncharged, have a diffuse cellular localization in cell culture and have been shown to have a compact flexible coil structure (Lee et al 2016). The GP DPR does localize to discrete cytoplasmic and nuclear foci in some Drosophila models (Freibaum et al 2015,

Lee et al 2016, Mizielinska et al 2014b, Wen et al 2014) suggesting limited aggregation activity. This DPR also appears to affect the functionality of the ubiquitin-proteasome system and increases cell death in the presence of the proteasome inhibitor MG-132 (Yamakawa et al 2015). The limited toxicity of the GP DPRs may explain why its levels do not correlate with disease onset, clinical and disease outcome measures (Gendron et al 2017), although GP has been proposed as pharmacokinetic biomarker for *C9orf72* expression and mutation carrier status. Despite limited toxicity data, polyPA was reported to modulate the GA-dependent toxicity in cells and chick spinal cord (Lee et al 2017), suggesting a complicated relationship between C9 DRPs, protein aggregation and toxicity. Given that overwhelming the cellular protein degradation system has been proposed as a disease mechanism for C9-ALS/FTD, it is possible that all C9 DPR may contribute to disease, apart from their individual inherent toxicity or biophysical properties, simply through their unintended expression and accumulation. Additionally, it will be important to understand if toxicity of specific DPRs or RAN proteins is specific to a particular cell types (e.g. neurons versus glia cells) in order to understand if the neurodegeneration seen in the disease is cell autonomous or involves neighboring glia cells.

3.4. Modifiers of C9 DPR toxicity—Identifying modifiers of C9 DPR toxicity is likely to provide insight into the underlying disease mechanisms and for the development of therapeutic strategies. One of the first C9 RAN protein toxicity modifier screens was performed in yeast, a system that recapitulates the toxicity of polyPR and polyGR (Jovicic et al 2015). Two unbiased screens for modifiers of $(PR)_{50}$ toxicity identified dozens of suppressors and enhancers, including a number of genes encoding proteins involved in nucleocytoplasmic transport (e.g. MTR10 an import receptor, NDC1 a key component of the nuclear pore complex). Similar screens for polyGR toxicity modifiers revealed 133 gene candidates suppressors, also including genes involved in nucleocytoplasmic transport (Chai & Gitler 2018). However the majority of polyGR regulators were genes involved in rRNA processing and ribosome synthesis, supporting distinct toxicity mechanisms for polyPR and polyGR RAN protiens. Similar modifiers were generated by a $(PR)_{20}$ CRISPR-Cas9 screen in human cells and primary mouse neurons, which included genes related nucleocytoplasmic transport, ER, proteasome, RNA processing pathway, and chromatin modification (Kramer et al 2018). This group further investigated two potent modifiers, *RAB7A*, an endolysosomal trafficking gene and *Tmx2*, an ER-resident transmembrane thioredoxin protein. While knockdown of *RAB7A* in HeLa cells altered PR_{20} subcellular localization, *Tmx2* knockdown resulted in upregulation of pro-survival unfolded-protein-response-pathway genes suggesting *Tmx2* mitigates toxicity by modulating ER-stress response. Interestingly, the C9ORF72 protein may itself be a modifier of polyGR and polyPR toxicity. Studies on *C9ORF72*$^{-/+}$ and *C9ORF72*$^{-/-}$ iPSC-derived motor neurons (iMNs) show that they degenerated faster than controls and overexpression of the A or B isoforms of C9ORF72 protein induced the clearance of $(PR)_{50}$ aggregates in iMNs (Shi et al 2018). The interplay between C9 RAN proteins, modifiers and other cellular components has been proposed to form a feedforward loop that may exacerbate disease. Based on work in *Drosophila*, expression of C9 RAN proteins but not expansion RNAs was shown to cause the cytoplasmic accumulation of TDP-43, which in turn triggers karyopherin-α2/ cytoplasmic mislocalization and nucleocytoplasmic defects that further increase RAN protein levels

(Solomon et al 2018). Overall, protein transport and clearance pathway genes have been identified from modifier screens for other microsatellite expansion diseases (Todd et al 2010, VoSSfeldt et al 2012) suggesting targeting these pathways may generate therapeutics that work across multiple diseases.

4. MECHANISM OF RAN TRANSLATION

While there has been tremendous interest in the role of RAN proteins in disease, there is also a growing interest in understanding how RAN translation occurs. Canonical translation initiation is a complex, highly regulated, step-wise process that typically includes: the recognition of a 5' methyl-7-guanosine (m^7G) cap structure; the binding of a number of eukaryotic initiation factors (eIFs), the scanning for an AUG or a near-cognate AUG codon by preinitiation complex (43S) (PIC) containing small (40S) ribosomal subunit and Met-tRNAi, the irreversible hydrolysis of eIF2-GTP to release eIF2-GDP and other eIFs and the recruitment of the large (60S) subunit to form 80S ribosome where elongation begins (Harding et al 2000, Sonenberg & Hinnebusch 2009). Alternative initiation pathways have also been described such as internal ribosome entry sites (IRESs), in which structured RNAs directly recruit PIC to the start codon to circumvent the scanning process, especially under conditions where eIFs are inhibited (Komar & Hatzoglou 2011). Near-cognate start codons with one nucleotide different from the AUG are also used in mammalian cells for translation initiation (Hinnebusch 2014). The fact that many microsatellite expansion mutations are GC-rich sequences that form secondary structures similar to IRES sites, suggests that RAN translation may behave in an IRES-like manner. Additionally since RAN products across multiple reading frames in multiple diseases have been detected in patient tissues, it is possible that RAN translation uses both canonical scanning and initiation and alternative initiation pathways depending on the reading frame.

Clues on the mechanisms of RAN translation initiation have been garnered primarily from work on SCA8 CAG, FXTAS CGG and *C9orf72* G_4C_2 expansions. In the original paper, Zu et al. (Zu et al 2011), showed RAN translation was repeat length dependent and dependent on structured RNAs. Additionally, Zu et al., ruled out RNA editing and frame shifting mechanisms that could explain the expression of these proteins using canonical rules. Mass spectrometry of polyAla showed RAN proteins expressed in cell culture can initiate in that reading frame throughout the repeat track, whereas RAN proteins expressed from the polySer or polyGln reading frame began close to or at the beginning of the repeat tract. Additionally, Zu et al., showed that sequence context affected the expression of RAN proteins in individual reading frames and that stop codons placed immediately in front of the CAG expansion blocked the expression of the RAN polyGln but not RAN polyAla or RAN polySer proteins. Using an *in vitro* rabbit reticulocyte lysates, Zu et al., showed RAN translation only occurred across transcripts and in reading frames with close cognate initiation codons and in these cases used a Met-$tRNA_i^{Met}$ was used. Translation in this *in vitro* system was much more limited than the promiscuous translation observed in a variety of sequence contexts in mammalian cells raising the possibility that modeling RAN translation in cell free *in vitro* systems may not mimic what is occurring in disease.

Using a nanoluciferase reporter to monitor RAN translation, Kearse and co-workers showed that RAN translation of FXTAS CGG repeats is also repeat-length dependent. Additionally, they showed translation initiation in the polyGly and polyAla reading frames and utilizes cap-dependent scanning mechanism to initiate near-cognate start codons upstream (e.g. ACG or GUG) or within the repeat (Kearse et al 2016). The cap-dependent mechanism of polyGly across CGG repeats was supported by multiple lines of evidence including: significant reductions in polyGly translation products when using non-function A cap in the *FMR1* mRNA; reduced translation efficiency upon titration of m^7G-cap analog indicating the involvement of eIF4E protein and; inhibition of translation in two frames of CGG repeat by an eIF4A inhibitor showing the requirement of eIF4A helicase. Interestingly, insertion of three stop codons upstream of CGG repeats significantly reduced polyGly but not polyAla expression in transfected cells. Furthermore FMR-polyGly expression occurs independent of the repeat itself, in contrast to expression of sense FMR-polyAla and antisense proteins (polyPro, polyArg and polyAla). Taken together these data provide additional evidence for reading frame specific differences in FXTAS.

Several research groups have recently investigated translational mechanisms for the *C9orf72* expansion RNAs. First, several groups showed that RAN translation across both G4C2 and G2C4 repeats is repeat length dependent (Mori et al 2013, Zu et al 2013). Recently, two groups have reported that translation of *C9orf72* sense G4C2 expansion RNAs is 5' cap and eIF4E dependent and uses an upstream near-cognate CUG codon (Green et al 2017, Tabet et al 2018). In addition, Tabet and co-workers showed that the presence of a uORF on misspliced transcripts and a single CUG codon influence the expression of RAN proteins in all three reading frames in transfected cells. , upstream of the G4C2 repeat mRNA on intron retained transcripts, inhibited the RAN translation, possibly through preventing its interaction with PIC. In this model, initiation occurs at a close cognate CUG and frameshifting is proposed to lead to expression in the alternative reading frames. In contrast, Cheng et al. used a bicistronic reporter system with multiple stop codons inserted just upstream of the second ORF $(G4C2)_{70}$-nLuc and detected significant expression, supporting a cap-independent mechanism (Cheng et al 2018). Translational efficiency also varies between the C9 reading frames with GA frame translated more efficiently than GP and GR frames in cell-free system. C9 RAN protein expression levels also appear affected by repeat length, although these findings are still somewhat controversial. Similar to the frame-specific differences in FXTAS, C9 polyGP DPR production was remarkably influenced by repeat length while no striking difference in expression levels was observed for polyGA and polyGR between 30 to 66 repeats (Tabet et al 2018). In contrast, Green and co-workers reported expression of polyGA increased with repeat length in their system (Green et al 2017). Thus, whether *C9* G4C2 repeat mRNA utilizes cap-dependent or IRES-like mechanism or both mechanism *in vivo* and if both, which one is dominant, remain to be investigated.

Clues to possible translation initiation factors involved in RAN translation have arisen from two recent publications. First, a novel SCA8 polySer protein was recently shown to accumulate as aggregates in the white matter regions of SCA8 mouse and human autopsy breains (Ayhan et al 2018). White matter accumulation of polySer and QAGR has also been reported for HD and DM2 (Bañez-Coronel et al 2015, Zu et al 2017) suggesting a specific

factor in white matter regions may favor RAN translation. Because eIF3F was reported to be elevated in white matter regions (Mills et al 2013), Ayhan and co-workers tested this factors in RAN translation and showed that knockdown of eIF3F influenced SCA8 RAN translation (Ayhan et al 2018). The mammalian eIF3 complex, of which eIF3F is a “non-core” subunit, is essential for most forms of cap-dependent and cap-independent translation. Interestingly eIF3F knockdown reduced steady-state levels of polySer and polyAla expressed from constructs lacking ATG start codons but not from constructs with ATG start codon or ATG-like codon. Similarly, a decrease of polyGP levels produced from G4C2 repeats was found in eIF3F knockdown cells, suggesting eIF3F may act as RAN trans-activating factor (RAN-TAF) of RAN translation in C9 ALS/FTD and SCA8 (Ayhan et al 2018). Second, Sonobe and co-workers (Sonobe et al 2018) recently showed that eIF2A is used, at least partially, for translation of the *C9orf72* polyGA DPRs. While knockdown of eIF2A decreased translation of nanoLuciferase-polyGA in chick embryo neural cells, its knockout did not completely abolish polyGA translation, suggesting other translation initiation factors are also involved. When canonical cap-dependent translation is limited by eIF2α-phosphorylation, eIF2A becomes increasingly involved in translation, especially for mRNAs with an uORF. Understanding additional factors regulating RAN translation may help explain the regional distribution of specific RAN proteins and their contribution to neurodegeneration and disease.

4.1. Stress and RAN translation

Integrated stress response (ISR) is activated by cells in response to a diverse set of stress stimuli and leads to decreased global protein synthesis and expression of select genes to promote cellular recovery (Kroemer et al 2010, Pakos-Zebrucka et al 2016). The critical step in this pathway is the phosphorylation of eukaryotic initiation factor 2 alpha (eIF2α) (Clemens 2001, Harding et al 2000, Holcik & Sonenberg 2005), which leads to the attenuation of cap-dependent translation and activation of non-canonical translation initiation, such as those driven by IRESs. Given the uncertain mechanistic underpinnings of RAN translation, there has been considerable recent focus on the effects of the ISR on RAN translation.

Introduction of ER stress via thapsigargin (an ER calcium pump inhibitor), tunicamycin (a global translation inhibitor) or oxidative stress (sodium arsenite) significantly increase expression of polyGly and polyAla from *FMR1* CGG repeats (Green et al 2017). Expression of C9-RAN proteins were also unregulated in primary rat cortical neurons under ER stress induced by thapsigargin (Green et al 2017). Cap-independent RAN translation of $(G4C2)_{70}$ and intronic repeat mRNA was also recently shown to be upregulated by various stress stimuli, including oxidative stress (sodium arsenite) or unfolded protein stress (MG132-induced) (Cheng et al 2018). Additionally, recent results show that C9 polyGA translation utilizes eIF2A, a non-canonical translation initiation factor, and is associated with ISR induction that can lead to attenuation of conventional cap-dependent translation (Sonobe et al 2018). Taken together these data suggest a feedforward mechanism in which ISR induction in neural cells triggers eIF2α phosphorylation and use of eIF2A which in turn leads to increased DPR synthesis. Since C9 RAN proteins are aggregation prone and their expression induces stress response, this pathway may feeds back on itself and lead to further

RAN protein production and the acceleration of disease progression. Understanding the precipitating event in this cycle may provide a valuable future target for therapeutics.

5. THERAPEUTIC APPROACHES AND RAN TRANSLATION

For repeat expansion diseases, there are a wealth of potential therapeutic targets but the complex nature of these diseases has hampered the development of effective therapies. Potential targets range from the repeats in the genomic DNA, the sense and antisense expansion RNAs, a plethora of RAN proteins and affected downstream cellular pathways.

The expansion RNAs, the current primary focus of much research, have been targeted by various techniques including antisense oligonucleotide (ASOs), small molecules, drugs and alterations of upstream processes (Cleary et al 2018, Connelly et al 2016, Gao & Cooper 2013, Taylor et al 2016). ASO strategies are being tested and/or developed for numerous expansion disorders, including DM1, HD, SCA2, SCA3 and *C9orf72* ALS/FTD (Gao & Cooper 2013, Jiang et al 2016a, Moore et al 2017, Sah & Aronin 2011, Scoles et al 2017, Wheeler et al 2012). In C9 ALS/FTD, a single dose of an ASO targeting the sense-transcript to a C9 BAC mouse model decreased expanded *C9ORF72* transcript levels, sense foci and RAN proteins levels (poly-GA & poly-GP) (Jiang et al 2016a). Even 6 months following ASO injection, these mice still demonstrated beneficial effects, such as lowered polyGP and polyGA levels, despite expansion RNA levels return to pre-injection levels. An important consideration for any of the ASO-based approaches will be the bidirectional nature of expression from the expanded repeat, especially since most ASOs target either sense or antisense transcripts but not both transcripts. One interesting approach is targeting transcription of the repeat rather than expansion RNAs, either directly by deactivated Cas9 (Pinto et al 2017) or by upstream cellular processes (Cheng et al 2015, Kramer et al 2016). The use of a deactivated Cas9 and guide RNAs against the repeat itself was recently shown to impede transcription of expanded microsatellite repeats, rescue molecular and cellular phenotypes inpatient-derived cells as well as muscle phenotypes in a mouse model of myotonic dystrophy (Pinto et al 2017). Targeting transcription was also attempted using knockdown of SUPT4H1/SUPT5H transcriptional elongation factor complex, which reduces transcription of genes with long stretches of expanded repeats (Cheng et al 2015, Jiang & Cleveland 2016, Kramer et al 2016). Genetic deletion or ASO knockdown reduction of SUPT4H1 levels in 2Q175 or R6/2 mice decreased mutant expanded HTT mRNA, Htt aggregates and resulted in phenotypic recovery (Cheng et al 2015). A similar approach in C9 ALS fibroblasts and C9 iPSC-derived cortical neurons reduced both sense and antisense RNA foci as well as poly-GP protein (Kramer et al 2016). Small molecule approaches to targeting transcription have also applied to expansion diseases including the use of Actinomycin D and furamidine to reduce transcription of expansion RNA in myotonic dystrophy model systems (Jenquin et al 2018, Siboni et al 2015). While these drugs may primarily reduce transcription of the repeat, they may also act on other events such as the expression of RBPs or disruption of RBP-expansion RNA interaction (Jenquin et al 2018). Alternative approaches such as inhibiting nuclear export of C9 expansion RNAs by targeting SRSF1. This prevents the interaction of expansion RNAs with nuclear export receptors and has been shown to prevent neurodegeneration and locomotor deficits in flies (Hautbergue et al 2017). It is important to note that due to the multiple pathogenic mechanisms involved in

repeat expansion diseases, targeting one process may have unintended consequences on others. For example, inhibiting sequestration of expansion RNAs by RBPs may allow the transcript to be exported to the cytoplasm where they can undergo RAN translation and exacerbate disease. Ultimately understanding the nature, extent and cellular consequences of on and off-target effects of each of these approaches will be critical for the development of effective clinical treatments for expansion diseases.

Surprisingly, given the wealth of knowledge on RAN proteins, especially in C9ORF72 ALS/FTD, there are limited therapeutic strategies aimed directly at the proteins. Treatment with anti-GA antibodies has been shown to inhibit intracellular poly-GA aggregation in cell culture and block seeding activity in brain extracts (Zhou et al 2017). However there are more strategies aimed at the downstream consequences of RAN proteins. For example targeting expansion-associated nucleocytoplasmic transport defects are neuroprotective in both C9ORF72-ALS/FTD mouse models (Hautbergue et al 2017, Zhang et al 2015) and HD (Grima et al 2017). Overexpression of the small heat shock protein B8 (HSPB8), that modulates autophagy-mediated disposal of misfolded aggregation-prone proteins, was also recently shown to decrease the accumulation of most C9-RAN proteins (Cristofani et al 2018). Salubrinal, an ubiquitin-proteasome-system inhibitor, and other compounds that reduces ER stress enhances cell survival in neurons expressing poly(GA) (Zhang et al 2014). Given the increasing wealth of knowledge regarding RAN proteins and their downstream consequences, a large number of potential therapeutic targets are likely to come to light and advance to pre-clinical testing in short order. Clearly, a better understanding of the mechanisms of RAN translation and the role of individual RAN proteins in disease will increase the pace and breadth of future therapeutic opportunities.

ACKNOWLEDGEMENTS

The authors thank the families that have participated in research and have made research into the mechanisms of microsatellite diseases and the discovery of RAN translation possible. This work was supported by the NIH (P01NS058901, R37NS040389, R01NS098819), the ALS Association, Target ALS, Muscular Dystrophy Association, the Robert Packard Center for ALS Research, the Myotonic Dystrophy Foundation, donors to the Center for NeuroGenetics and the University of Florida.

Literature cited

Ash PE, Bieniek KF, Gendron TF, Caulfield T, Lin WL, et al. 2013 Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. Neuron 77: 639–46 [PubMed: 23415312]

Ayhan F, Perez BA, Shorrock HK, Zu T, Banez-Coronel M, et al. 2018 SCA8 RAN polySer protein preferentially accumulates in white matter regions and is regulated by eIF3F. EMBO J

Bañez-Coronel M, Ayhan F, Tarabochia AD, Zu T, Perez BA, et al. 2015 RAN Translation in Huntington Disease. Neuron 88: 667–77 [PubMed: 26590344]

Bates GP, Dorsey R, Gusella JF, Hayden MR, Kay C, et al. 2015 Huntington disease. Nature Reviews Disease Primers 1

Bäuerlein FJB, Saha I, Mishra A, Kalemanov M, Martinez-Sanchez A, et al. 2017 In Situ Architecture and Cellular Interactions of PolyQ Inclusions. Cell 171: 179–87 [PubMed: 28890085]

Bevivino AE, Loll PJ. 2001 An expanded glutamine repeat destabilizes native ataxin-3 structure and mediates parallel beta-fibrils. P Natl Acad Sci USA 98: 11955–60

Bidichandani SI, Ashizawa T, Patel PI. 1998 The GAA triplet-repeat expansion in Friedreich ataxia interferes with transcription and may be associated with an unusual DNA structure. Am J Hum Genet 62: 111–21 [PubMed: 9443873]

Campuzano V, Montermini L, Moltó MD, Pianese L, Cossée M, et al. 1996 Friedreich's ataxia: autosomal recessive disease caused by an intronic GAA triplet repeat expansion. Science 271: 1423–7 [PubMed: 8596916]

Castro H, Kul E, Buijsen RAM, Severijnen L, Willemsen R, et al. 2017 Selective rescue of heightened anxiety but not gait ataxia in a premutation 90CGG mouse model of Fragile X-associated tremor/ataxia syndrome. Hum Mol Genet 26: 2133–45 [PubMed: 28369393]

Chai N, Gitler AD. 2018 Yeast screen for modifiers of C9orf72 poly(glycine-arginine) dipeptide repeat toxicity. FEMS Yeast Res 18

Chang YJ, Jeng US, Chiang YL, Hwang IS, Chen YR. 2016 The Glycine-Alanine Dipeptide Repeat from C9orf72 Hexanucleotide Expansions Forms Toxic Amyloids Possessing Cell-to-Cell Transmission Properties. Journal of Biological Chemistry 291: 4903–11 [PubMed: 26769963]

Cheng HM, Chern Y, Chen IH, Liu CR, Li SH, et al. 2015 Effects on murine behavior and lifespan of selectively decreasing expression of mutant huntingtin allele by supt4h knockdown. PLoS Genet 11: e1005043 [PubMed: 25760041]

Cheng W, Wang S, Mestre AA, Fu C, Makarem A, et al. 2018 C9ORF72 GGGGCC repeat- associated non-AUG translation is upregulated by stress through eIF2alpha phosphorylation. Nat Commun 9: 51 [PubMed: 29302060]

Chitiprolu M, Jagow C, Tremblay V, Bondy-Chorney E, Paris G, et al. 2018 A complex of C9ORF72 and p62 uses arginine methylation to eliminate stress granules by autophagy. Nat Commun 9: 2794 [PubMed: 30022074]

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: 483–9 [PubMed: 16285929]

Cleary JD, LaSpada AR, Pearson CE. 2006 DNA Replication, Repeat Instability, and Human Disease In DNA Replication and Human Disease, ed. DePamphilis ML. Cold Spring Harbour: Cold Spring Harbor Laboratory Press

Cleary JD, Pattamatta A, Ranum LPW. 2018 Repeat associated non-ATG (RAN) translation. J Biol Chem

Cleary JD, Ranum LP. 2013 Repeat-associated non-ATG (RAN) translation in neurological disease. Hum Mol Genet 22: R45–51 [PubMed: 23918658]

Cleary JD, Ranum LP. 2014 Repeat associated non-ATG (RAN) translation: new starts in microsatellite expansion disorders. Curr Opin Genet Dev 26: 6–15 [PubMed: 24852074]

Clemens MJ. 2001 Initiation factor eIF2 alpha phosphorylation in stress responses and apoptosis. Prog Mol Subcell Biol 27: 57–89 [PubMed: 11575161]

Connelly CM, Moon MH, Schneekloth JS. 2016 The Emerging Role of RNA as a Therapeutic Target for Small Molecules. Cell Chem Biol 23: 1077–90 [PubMed: 27593111]

Cristofani R, Crippa V, Vezzoli G, Rusmini P, Galbiati M, et al. 2018 The small heat shock protein B8 (HSPB8) efficiently removes aggregating species of dipeptides produced in C9ORF72-related neurodegenerative diseases. Cell Stress Chaperon 23: 1–12

Duennwald ML, Lindquist S. 2008 Impaired ERAD and ER stress are early and specific events in polyglutamine toxicity. Gene Dev 22: 3308–19 [PubMed: 19015277]

Figiel M, Szlachcic WJ, Switonski PM, Gabka A, Krzyzosiak WJ. 2012 Mouse Models of Polyglutamine Diseases: Review and Data Table. Part I. Mol Neurobiol 46: 393–429 [PubMed: 22956270]

Flores BN, Dulchaysky ME, Krans A, Sawaya MR, Paulson HL, et al. 2016 Distinct C9orf72-Associated Dipeptide Repeat Structures Correlate with Neuronal Toxicity. PLoS One 11

Freibaum BD, Lu Y, Lopez-Gonzalez R, Kim NC, Almeida S, et al. 2015 GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. Nature 525: 129–33 [PubMed: 26308899]

Gao Z, Cooper TA. 2013 Antisense oligonucleotides: rising stars in eliminating RNA toxicity in myotonic dystrophy. Hum Gene Ther 24: 499–507 [PubMed: 23252746]

Gendron TF, Chew J, Stankowski JN, Hayes LR, Zhang YJ, et al. 2017 Poly(GP) proteins are a useful pharmacodynamic marker for C9ORF72-associated amyotrophic lateral sclerosis. Science Translational Medicine 9

Gijselinck I, Van Mossevelde S, van der Zee J, Sieben A, Engelborghs S, et al. 2016 The C9orf72 repeat size correlates with onset age of disease, DNA methylation and transcriptional downregulation of the promoter. Mol Psychiatr 21: 1112–24

Green KM, Glineburg MR, Kearse MG, Flores BN, Linsalata AE, et al. 2017 RAN translation at C9orf72-associated repeat expansions is selectively enhanced by the integrated stress response. Nat Commun 8: 2005 [PubMed: 29222490]

Grima JC, Daigle JG, Arbez N, Cunningham KC, Zhang K, et al. 2017 Mutant Huntingtin Disrupts the Nuclear Pore Complex. Neuron 94: 93–107 e6 [PubMed: 28384479]

Guo Q, Lehmer C, Martinez-Sanchez A, Rudack T, Beck F, et al. 2018 In Situ Structure of Neuronal C9orf72 Poly-GA Aggregates Reveals Proteasome Recruitment. Cell 172: 696–705 **e12** [PubMed: 29398115] **e12**

Hagerman P 2013 Fragile X-associated tremor/ataxia syndrome (FXTAS): pathology and mechanisms. Acta Neuropathol 126: 1–19 [PubMed: 23793382]

Hagerman R, Hagerman P. 2013 Advances in clinical and molecular understanding of the FMR1 premutation and fragile X-associated tremor/ataxia syndrome. Lancet Neurol 12: 786–98 [PubMed: 23867198]

Harding HP, Novoa I, Zhang Y, Zeng H, Wek R, et al. 2000 Regulated translation initiation controls stress-induced gene expression in mammalian cells. Mol Cell 6: 1099–108 [PubMed: 11106749]

Harley HG, Brook JD, Rundle SA, Crow S, Reardon W, et al. 1992 Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy. Nature 355: 545–6. [PubMed: 1346923]

Hautbergue GM, Castelli LM, Ferraiuolo L, Sanchez-Martinez A, Cooper-Knock J, et al. 2017 SRSF1-dependent nuclear export inhibition of C9ORF72 repeat transcripts prevents neurodegeneration and associated motor deficits. Nat Commun 8: 16063 [PubMed: 28677678]

Hinnebusch AG. 2014 The scanning mechanism of eukaryotic translation initiation. Annu Rev Biochem 83: 779–812 [PubMed: 24499181]

Holcik M, Sonenberg N. 2005 Translational control in stress and apoptosis. Nat Rev Mol Cell Biol 6: 318–27 [PubMed: 15803138]

Ishiguro T, Sato N, Ueyama M, Fujikake N, Sellier C, et al. 2017 Regulatory Role of RNA Chaperone TDP-43 for RNA Misfolding and Repeat-Associated Translation in SCA31. Neuron

Jenquin JR, Coonrod LA, Silverglate QA, Pellitier NA, Hale MA, et al. 2018 Furamidine Rescues Myotonic Dystrophy Type I Associated Mis-Splicing through Multiple Mechanisms. ACS Chem Biol 13: 2708–18 [PubMed: 30118588]

Jiang J, Cleveland DW. 2016 Bidirectional Transcriptional Inhibition as Therapy for ALS/FTD Caused by Repeat Expansion in C9orf72. Neuron 92: 1160–63 [PubMed: 28009271]

Jiang J, Zhu Q, Gendron TF, Saberi S, McAlonis-Downes M, et al. 2016a Gain of Toxicity from ALS/FTD-Linked Repeat Expansions in C9ORF72 Is Alleviated by Antisense Oligonucleotides Targeting GGGGCC-Containing RNAs. Neuron 90: 535–50 [PubMed: 27112497]

Jiang YW, Chadwick SR, Lajoie P. 2016b Endoplasmic reticulum stress: The cause and solution to Huntington's disease? Brain Research 1648: 650–57 [PubMed: 27040914]

Jovicic A, Mertens J, Boeynaems S, Bogaert E, Chai N, et al. 2015 Modifiers of C9orf72 dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS. Nat Neurosci 18: 1226–9 [PubMed: 26308983]

Kanadia RN, Shin J, Yuan Y, Beattie SG, Wheeler TM, et al. 2006 Reversal of RNA missplicing and myotonia after muscleblind overexpression in a mouse poly(CUG) model for myotonic dystrophy. Proc Natl Acad Sci U S A 103: 11748–53 [PubMed: 16864772]

Kearse MG, Green KM, Krans A, Rodriguez CM, Linsalata AE, et al. 2016 CGG Repeat-Associated Non-AUG Translation Utilizes a Cap-Dependent Scanning Mechanism of Initiation to Produce Toxic Proteins. Mol Cell 62: 314–22 [PubMed: 27041225]

Khosravi B, Hartmann H, May S, Mohl C, Ederle H, et al. 2017 Cytoplasmic poly-GA aggregates impair nuclear import of TDP-43 in C9orf72 ALS/FTLD. Hum Mol Genet 26: 790–800 [PubMed: 28040728]

Komar AA, Hatzoglou M. 2011 Cellular IRES-mediated translation: the war of ITAFs in pathophysiological states. Cell Cycle 10: 229–40 [PubMed: 21220943]

Kouroku Y, Fujita E, Jimbo A, Kikuchi T, Yamagata T, et al. 2002 Polyglutamine aggregates stimulate ER stress signals and caspase-12 activation. Human Molecular Genetics 11: 1505–15 [PubMed: 12045204]

Kramer NJ, Carlomagno Y, Zhang YJ, Almeida S, Cook CN, et al. 2016 Spt4 selectively regulates the expression of C9orf72 sense and antisense mutant transcripts. Science 353: 708–12 [PubMed: 27516603]

Kramer NJ, Haney MS, Morgens DW, Jovicic A, Couthouis J, et al. 2018 CRISPR-Cas9 screens in human cells and primary neurons identify modifiers of C9ORF72 dipeptide-repeat-protein toxicity. Nat. Genet 50: 603–12 [PubMed: 29507424]

Krans A, Kearse MG, Todd PK. 2016 RAN translation from antisense CCG repeats in Fragile X Tremor/Ataxia Syndrome. Ann Neurol

Kroemer G, Marino G, Levine B. 2010 Autophagy and the Integrated Stress Response. Molecular Cell 40: 280–93 [PubMed: 20965422]

Kwon I, Xiang S, Kato M, Wu L, Theodoropoulos P, et al. 2014 Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. Science 345: 1139–45 [PubMed: 25081482]

Labbadia J, Morimoto RI. 2013 Huntington's disease: underlying molecular mechanisms and emerging concepts. Trends Biochem Sci 38: 378–85 [PubMed: 23768628]

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, et al. 2001 Initial sequencing and analysis of the human genome. Nature 409: 860–921 [PubMed: 11237011]

Lee KH, Zhang P, Kim HJ, Mitrea DM, Sarkar M, et al. 2016 C9orf72 Dipeptide Repeats Impair the Assembly, Dynamics, and Function of Membrane-Less Organelles. Cell 167: 774–88.e17 [PubMed: 27768896]

Lee YB, Baskaran P, Gomez-Deza J, Chen HJ, Nishimura AL, et al. 2017 C9orf72 poly GA RAN-translated protein plays a key role in amyotrophic lateral sclerosis via aggregation and toxicity. Human Molecular Genetics 26: 4765–77 [PubMed: 28973350]

Lehmer C, Oeckl P, Weishaupt JH, Volk AE, Diehl-Schmid J, et al. 2017 Poly-GP in cerebrospinal fluid links C9orf72-associated dipeptide repeat expression to the asymptomatic phase of ALS/FTD. EMBO Mol Med 9: 859–68 [PubMed: 28408402]

Leitman J, Hartl FU, Lederkremer GZ. 2013 Soluble forms of polyQ-expanded huntingtin rather than large aggregates cause endoplasmic reticulum stress. Nature Communications 4

Lin Y, Mori E, Kato M, Xiang S, Wu L, et al. 2016 Toxic PR Poly-Dipeptides Encoded by the C9orf72 Repeat Expansion Target LC Domain Polymers. Cell 167: 789–802.e12 [PubMed: 27768897]

Liu Y, Pattamatta A, Zu T, Reid T, Bardhi O, et al. 2016 C9orf72 BAC Mouse Model with Motor Deficits and Neurodegenerative Features of ALS/FTD. Neuron 90: 521–34 [PubMed: 27112499]

Lopez-Gonzalez R, Lu Y, Gendron TF, Karydas A, Tran H, et al. 2016 Poly(GR) in C9ORF72-Related ALS/FTD Compromises Mitochondrial Function and Increases Oxidative Stress and DNA Damage in iPSC-Derived Motor Neurons. Neuron 92: 383–91 [PubMed: 27720481]

Lopez Castel A, Cleary JD, Pearson CE. 2010 Repeat instability as the basis for human diseases and as a potential target for therapy. Nat Rev Mol Cell Biol 11: 165–70 [PubMed: 20177394]

May S, Hornburg D, Schludi MH, Arzberger T, Rentzsch K, et al. 2014 C9orf72 FTLD/ALS-associated Gly-Ala dipeptide repeat proteins cause neuronal toxicity and Unc119 sequestration. Acta Neuropathologica 128: 485–503 [PubMed: 25120191]

Miller JW, Urbinati CR, Teng-Umnuay P, Stenberg MG, Byrne BJ, et al. 2000 Recruitment of human muscleblind proteins to (CUG)(n) expansions associated with myotonic dystrophy. EMBO J 19: 4439–48 [PubMed: 10970838]

Mills JD, Kavanagh T, Kim WS, Chen BJ, Kawahara Y, et al. 2013 Unique transcriptome patterns of the white and grey matter corroborate structural and functional heterogeneity in the human frontal lobe. PLoS ONE 8: e78480 [PubMed: 24194939]

Mizielinska S, Grönke S, Niccoli T, Ridler CE, Clayton EL, et al. 2014a C9orf72 repeat expansions cause neurodegeneration in Drosophila through arginine-rich proteins. Science 345: 1192–4 [PubMed: 25103406]

Mizielinska S, Gronke S, Niccoli T, Ridler CE, Clayton EL, et al. 2014b C9orf72 repeat expansions cause neurodegeneration in Drosophila through arginine-rich proteins. Science 345: 1192–94 [PubMed: 25103406]

Mohan A, Goodwin M, Swanson MS. 2014 RNA-protein interactions in unstable microsatellite diseases. Brain Res 1584: 3–14 [PubMed: 24709120]

Moore LR, Rajpal G, Dillingham IT, Qutob M, Blumenstein KG, et al. 2017 Evaluation of Antisense Oligonucleotides Targeting ATXN3 in SCA3 Mouse Models. Mol Ther Nucleic Acids 7: 200–10 [PubMed: 28624196]

Mori K, Weng SM, Arzberger T, May S, Rentzsch K, et al. 2013 The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. Science 339: 1335–8 [PubMed: 23393093]

Moseley ML, Zu T, Ikeda Y, Gao W, Mosemiller AK, et al. 2006 Bidirectional expression of CUG and CAG expansion transcripts and intranuclear polyglutamine inclusions in spinocerebellar ataxia type 8. Nat Genet 38: 758–69 [PubMed: 16804541]

Nelson DL, Orr HT, Warren ST. 2013 The unstable repeats--three evolving faces of neurological disease. Neuron 77: 825–43 [PubMed: 23473314]

O'Rourke JR, Swanson MS. 2009 Mechanisms of RNA-mediated disease. J Biol Chem 284: 7419–23 [PubMed: 18957432]

Oh SY, He F, Krans A, Frazer M, Taylor JP, et al. 2015 RAN translation at CGG repeats induces ubiquitin proteasome system impairment in models of fragile X-associated tremor ataxia syndrome. Human Molecular Genetics 24: 4317–26 [PubMed: 25954027]

Ohki Y, Wenninger-Weinzierl A, Hruscha A, Asakawa K, Kawakami K, et al. 2017 Glycine-alanine dipeptide repeat protein contributes to toxicity in a zebrafish model of C9orf72 associated neurodegeneration. Mol Neurodegener 12: 6 [PubMed: 28088213]

Orr HT, Zoghbi HY. 2007 Trinucleotide repeat disorders. Annu Rev Neurosci 30: 575–621 [PubMed: 17417937]

Pakos-Zebrucka K, Koryga I, Mnich K, Ljujic M, Samali A, Gorman AM. 2016 The integrated stress response. EMBO Rep 17: 1374–95 [PubMed: 27629041]

Pinto BS, Saxena T, Oliveira R, Mendez-Gomez HR, Cleary JD, et al. 2017 Impeding Transcription of Expanded Microsatellite Repeats by Deactivated Cas9. Mol Cell 68: 479–90 e5 [PubMed: 29056323]

Poirier MA, Jiang H, Ross CA. 2005 A structure-based analysis of huntingtin mutant polyglutamine aggregation and toxicity: evidence for a compact beta-sheet structure. Human Molecular Genetics 14: 765–74 [PubMed: 15689354]

Ranum LPW, Cooper TA. 2006 RNA-mediated neuromuscular disorders. Annual Review of Neuroscience 29: 259–77

Sah DW, Aronin N. 2011 Oligonucleotide therapeutic approaches for Huntington disease. J Clin Invest 121: 500–7 [PubMed: 21285523]

Sakae N, Bieniek KF, Zhang YJ, Ross K, Gendron TF, et al. 2018 Poly-GR dipeptide repeat polymers correlate with neurodegeneration and Clinicopathological subtypes in C9ORF72-related brain disease. Acta Neuropathol Commun 6: 63 [PubMed: 30029693]

Schaffar G, Breuer P, Boteva R, Behrends C, Tzvetkov N, et al. 2004 Cellular toxicity of polyglutamine expansion proteins: Mechanism of transcription factor deactivation. Molecular Cell 15: 95–105 [PubMed: 15225551]

Scherzinger E, Sittler A, Schweiger K, Heiser V, Lurz R, et al. 1999 Self-assembly of polyglutamine-containing huntingtin fragments into amyloid-like fibrils: Implications for Huntington's disease pathology. P Natl Acad Sci USA 96: 4604–09

Scoles DR, Meera P, Schneider MD, Paul S, Dansithong W, et al. 2017 Antisense oligonucleotide therapy for spinocerebellar ataxia type 2. Nature 544: 362–66 [PubMed: 28405024]

Scotti MM, Swanson MS. 2016 RNA mis-splicing in disease. Nat Rev Genet 17: 19–32 [PubMed: 26593421]

Sellier C, Buijsen RAM, He F, Natla S, Jung L, et al. 2017 Translation of Expanded CGG Repeats into FMRpolyG Is Pathogenic and May Contribute to Fragile X Tremor Ataxia Syndrome. Neuron 93: 331–47 [PubMed: 28065649]

Shi YX, Lin SY, Staats KA, Li YC, Chang WH, et al. 2018 Haploinsufficiency leads to neurodegeneration in C9ORF72 ALS/FTD human induced motor neurons. Nature Medicine 24: 313–+

Siboni RB, Nakamori M, Wagner SD, Struck AJ, Coonrod LA, et al. 2015 Actinomycin D Specifically Reduces Expanded CUG Repeat RNA in Myotonic Dystrophy Models. Cell Rep 13: 2386–94 [PubMed: 26686629]

Solomon DA, Stepto A, Au WH, Adachi Y, Diaper DC, et al. 2018 A feedback loop between dipeptide-repeat protein, TDP-43 and karyopherin-alpha mediates C9orf72-related neurodegeneration. Brain 141: 2908–24 [PubMed: 30239641]

Sonenberg N, Hinnebusch AG. 2009 Regulation of translation initiation in eukaryotes: mechanisms and biological targets. Cell 136: 731–45 [PubMed: 19239892]

Sonobe Y, Ghadge G, Masaki K, Sendoel A, Fuchs E, Roos RP. 2018 Translation of dipeptide repeat proteins from the C9ORF72 expanded repeat is associated with cellular stress. Neurobiol Dis 116: 155–65 [PubMed: 29792928]

Soragni E, Petrosyan L, Rinkoski TA, Wieben ED, Baratz KH, et al. 2018 Repeat-Associated Non-ATG (RAN) Translation in Fuchs' Endothelial Corneal Dystrophy. Invest Ophth Vis Sci 59: 1888–96

Stopford MJ, Higginbottom A, Hautbergue GM, Cooper-Knock J, Mulcahy PJ, et al. 2017 C9ORF72 hexanucleotide repeat exerts toxicity in a stable, inducible motor neuronal cell model, which is rescued by partial depletion of Pten. Hum Mol Genet 26: 1133–45 [PubMed: 28158451]

Switonski PM, Szlachcic WJ, Gabka A, Krzyzosiak WJ, Figiel M. 2012 Mouse Models of Polyglutamine Diseases in Therapeutic Approaches: Review and Data Table. Part II. Mol Neurobiol 46: 430–66 [PubMed: 22944909]

Tabet R, Schaeffer L, Freyermuth F, Jambeau M, Workman M, et al. 2018 CUG initiation and frameshifting enable production of dipeptide repeat proteins from ALS/FTD C9ORF72 transcripts. Nat Commun 9: 152 [PubMed: 29323119]

Taylor JP, Brown RH, Cleveland DW. 2016 Decoding ALS: from genes to mechanism. Nature 539: 197–206 [PubMed: 27830784]

Todd PK, Oh SY, Krans A, He F, Sellier C, et al. 2013 CGG repeat-associated translation mediates neurodegeneration in fragile X tremor ataxia syndrome. Neuron 78: 440–55 [PubMed: 23602499]

Todd PK, Oh SY, Krans A, Pandey UB, Di Prospero NA, et al. 2010 Histone deacetylases suppress CGG repeat-induced neurodegeneration via transcriptional silencing in models of fragile X tremor ataxia syndrome. PLoS Genet 6: e1001240 [PubMed: 21170301]

Todd PK, Paulson HL. 2010 RNA-mediated neurodegeneration in repeat expansion disorders. Ann Neurol 67: 291–300 [PubMed: 20373340]

Tóth G, Gáspári Z, Jurka J. 2000 Microsatellites in different eukaryotic genomes: survey and analysis. Genome Res 10: 967–81 [PubMed: 10899146]

Trottier Y, Biancalana V, Mandel JL. 1994 Instability of CAG repeats in Huntington's disease: relation to parental transmission and age of onset. J Med Genet 31: 377–82. [PubMed: 8064815]

Underwood BR, Rubinsztein DC. 2008 Spinocerebellar ataxias caused by polyglutamine expansions: A review of therapeutic strategies. Cerebellum 7: 215–21 [PubMed: 18418676]

Vatsavayai SC, Yoon SJ, Gardner RC, Gendron TF, Vargas JN, et al. 2016 Timing and significance of pathological features in C9orf72 expansion-associated frontotemporal dementia. Brain 139: 3202–16 [PubMed: 27797809]

VoSSfeldt H, Butzlaff M, PrüSSing K, Ní Chárthaigh RA, Karsten P, et al. 2012 Large-scale screen for modifiers of ataxin-3-derived polyglutamine-induced toxicity in Drosophila. PLoS One 7: e47452 [PubMed: 23139745]

Walsh R, Storey E, Stefani D, Kelly L, Turnbull V. 2005 The roles of proteolysis and nuclear localisation in the toxicity of the polyglutamine diseases. A review. Neurotox Res 7: 43–57 [PubMed: 15639797]

Wen X, Tan W, Westergard T, Krishnamurthy K, Markandaiah SS, et al. 2014 Antisense proline-arginine RAN dipeptides linked to C9ORF72-ALS/FTD form toxic nuclear aggregates that initiate in vitro and in vivo neuronal death. Neuron 84: 1213–25 [PubMed: 25521377]

Westergard T, Jensen BK, Wen X, Cai J, Kropf E, et al. 2016 Cell-to-Cell Transmission of Dipeptide Repeat Proteins Linked to C9orf72-ALS/FTD. Cell Rep 17: 645–52 [PubMed: 27732842]

Wheeler TM, Leger AJ, Pandey SK, MacLeod AR, Nakamori M, et al. 2012 Targeting nuclear RNA for in vivo correction of myotonic dystrophy. Nature 488: 111–5 [PubMed: 22859208]

Wheeler TM, Thornton CA. 2007 Myotonic dystrophy: RNA-mediated muscle disease. Curr Opin Neurol 20: 572–6 [PubMed: 17885447]

Yamakawa M, Ito D, Honda T, Kubo K, Noda M, et al. 2015 Characterization of the dipeptide repeat protein in the molecular pathogenesis of c9FTD/ALS. Hum Mol Genet 24: 1630–45 [PubMed: 25398948]

Yang D, Abdallah A, Li Z, Lu Y, Almeida S, Gao FB. 2015 FTD/ALS-associated poly(GR) protein impairs the Notch pathway and is recruited by poly(GA) into cytoplasmic inclusions. Acta Neuropathol 130: 525–35 [PubMed: 26031661]

Zhang K, Donnelly CJ, Haeusler AR, Grima JC, Machamer JB, et al. 2015 The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. Nature 525: 56–61 [PubMed: 26308891]

Zhang Y-J, Gendron TF, Ebbert MTW, O'Raw AD, Yue M, et al. 2018 Poly(GR) impairs protein translation and stress granule dynamics in C9orf72-associated frontotemporal dementia and amyotrophic lateral sclerosis. Nat. Med 24: 1136–42 [PubMed: 29942091]

Zhang YJ, Gendron TF, Grima JC, Sasaguri H, Jansen-West K, et al. 2016 C9ORF72 poly(GA) aggregates sequester and impair HR23 and nucleocytoplasmic transport proteins. Nat Neurosci 19: 668–77 [PubMed: 26998601]

Zhang YJ, Jansen-West K, Xu YF, Gendron TF, Bieniek KF, et al. 2014 Aggregation-prone c9FTD/ALS poly(GA) RAN-translated proteins cause neurotoxicity by inducing ER stress. Acta Neuropathol 128: 505–24 [PubMed: 25173361]

Zhou Q, Lehmer C, Michaelsen M, Mori K, Alterauge D, et al. 2017 Antibodies inhibit transmission and aggregation of C9orf72 poly-GA dipeptide repeat proteins. EMBO Mol Med 9: 687–702 [PubMed: 28351931]

Zu T, Cleary JD, Liu Y, Banez-Coronel M, Bubenik JL, et al. 2017 RAN Translation Regulated by Muscleblind Proteins in Myotonic Dystrophy Type 2. Neuron 95: 1292–305 **e5** [PubMed: 28910618] **e5**

Zu T, Gibbens B, Doty NS, Gomes-Pereira M, Huguet A, et al. 2011 Non-ATG-initiated translation directed by microsatellite expansions. Proc Natl Acad Sci U S A 108: 260–5 [PubMed: 21173221]

Zu T, Liu Y, Banez-Coronel M, Reid T, Pletnikova O, et al. 2013 RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. Proc Natl Acad Sci U S A 110:E4968–77 [PubMed: 24248382]

SUMMARY POINTS

- Bidirectional transcription and RAN translation have expanded the traditional "coding vs non-coding" disease mechanism classification of microsatellite repeat diseases.
- RAN proteins with a variety of repeat motifs accumulate in the brain regions in multiple microsatellite repeat diseases.
- Individual RAN proteins can show patterns of accumulation that are highly variable and often involve focal sites of accumulation and pathology.
- Expression of RAN proteins can occur prior to the overt onset of symptoms and may represent a long-term stress on cellular system.
- Toxicity varies greatly between individual RAN proteins expressed from the same repeat expansion and may cause toxicity directly or through interactions with other proteins.
- Nucleocytoplasmic transport, ER and protein clearance pathways are affected by the expression of individual RAN proteins.
- RAN translation initiation can be variable and share characteristics of both canonical and non-canonical translation pathways depending upon the repeat.
- The stress pathway system and eIF2α are closely connected to the regulation of RAN translation.

FUTURE ISSUES

- Identifying the best tools to quantitatively and reliably examine RAN proteins in both model systems and patients.
- Identification of individual RAN proteins as biomarkers of disease and the best candidates for therapeutic interventions?
- When and where in humans are RAN proteins first expressed, begin to accumulate and ultimately contribute to disease.
- How many repeat expansion diseases are there and what are the common molecular mechanisms between these diseases?
- How frequent does RAN translation occurs in humans including healthy individuals? Does RAN translation play any role in translation regulation.

Table 1.

Repeat expansion diseases and RAN proteins.

Diseases	Gene	Expansion	RAN	Cells	Patient tissues
C9orf72 ALS/FTD	*C9orf72*	G_4C_2	polyGA$_S$	x	hippocampus, frontotemporal neocortex, cerebellum
			polyGP$_S$	x	hippocampus, frontotemporal neocortex, cerebellum, spinal cord
			polyGR$_S$	x	hippocampus, frontotemporal neocortex, cerebellum
	C9orf72$_{AS}$	G_2C_4	polyPR$_{AS}$	x	hippocampus, frontotemporal neocortex, cerebellum, spinal cord
			polyPA$_{AS}$,	x	hippocampus, frontotemporal neocortex, cerebellum
			polyGP$_{AS}$	x	hippocampus, frontotemporal neocortex, cerebellum, spinal cord
DM1	*DMPK*	CTG	polyL$_S$	x	
			polyC$_S$	x	
			polyA$_S$	x	
	DM1$_{AS}$	CAG	polyQ$_{AS}$	x	myoblast, skeletal muscle, leukocytes
			polyS$_{AS}$	x	
			polyA$_{AS}$	x	
DM2	*CNBP*	CCTG	polyLPAC	x	neurons, activated microglia in grey matter
	DM2$_{AS}$	CAGG	polyQAGR	x	oligodendrocytes in white matter
HD	*Htt1*	CAG	polyQ$_S$	x	caudate, grey matter of FCX
			polyS$_S$	x	WMB caudate and putamen, WM of FCX and CB
			polyA$_S$	x	WMB caudate and putamen, WM of FCX and CB
	HD$_{AS}$	CTG	polyL$_{AS}$	x	WMB caudate and putamen, WM of FCX and CB
			polyC$_{AS}$	x	WMB caudate and putamen, WM of FCX and CB
			polyA$_{AS}$	x	WMB caudate and putamen, WM of FCX and CB
FECD	*TCF4*	CTG	polyL$_S$		
			polyC$_S$	x	corneal endothelium
			polyA$_S$		
	TCF4 $_{AS}$	CAG	polyQ$_{AS}$		
			polyS$_{AS}$		
			polyA$_{AS}$		
FXTAS	*FMRl*	CGG	polyG$_S$	x	hippocampus, frontal cortex, cerebellum
			polyR$_S$		
			polyA$_S$	x	
	FMR1$_{AS}$	CCG	polyA$_{AS}$	x	hippocampus, frontal cortex, cerebellum, midbrain
			polyP$_{AS}$	x	hippocampus, frontal cortex, cerebellum, midbrain
			polyR$_{AS}$	x	
SCA8	*ATXN8*	CAG	polyQ$_S$	x	Purkinje cells
			polyS$_S$	x	cerebellar white matter regions, brain stem, frontal cortex) cells

Diseases	Gene	Expansion	RAN	Cells	Patient tissues
			$polyA_S$	x	Purkinje cells
	ATXN8OS	CTG	$polyL_{AS}$	x	
			$polyC_{AS}$	x	
			$polyA_{AS}$	x	Purkinje cells
SCA31	*BEAN*	TGGAA	$polyWNGME_S$		Purkinje cells
	$SCA31_{AS}$	TTCCA	$polyFHSIP_{AS}$		