

Review Article

Mechanisms of repeat-associated non-AUG translation in neurological microsatellite expansion disorders

Lydia M. Castelli¹, Wan-Ping Huang², Ya-Hui Lin¹, Kung-Yao Chang² and  Guillaume M. Hautbergue^{1,3}

¹Sheffield Institute for Translational Neuroscience (SITraN), Department of Neuroscience, University of Sheffield, Sheffield S10 2HQ, U.K.; ²Institute of Biochemistry, Life Science Building, National Chung-Hsing University, Taichung City 402, Taiwan; ³Neuroscience Institute, University of Sheffield, Western Bank, Sheffield S10 2TN, U.K.

Correspondence: Lydia M Castelli (l.m.castelli@sheffield.ac.uk) or Kung-Yao Chang (kychang@dragon.nchu.edu.tw) or Guillaume M Hautbergue (g.hautbergue@sheffield.ac.uk)



Repeat-associated non-AUG (RAN) translation was discovered in 2011 in spinocerebellar ataxia type 8 (SCA8) and myotonic dystrophy type 1 (DM1). This non-canonical form of translation occurs in all reading frames from both coding and non-coding regions of sense and antisense transcripts carrying expansions of trinucleotide to hexanucleotide repeat sequences. RAN translation has since been reported in 7 of the 53 known microsatellite expansion disorders which mainly present with neurodegenerative features. RAN translation leads to the biosynthesis of low-complexity polymeric repeat proteins with aggregating and cytotoxic properties. However, the molecular mechanisms and protein factors involved in assembling functional ribosomes in absence of canonical AUG start codons remain poorly characterised while secondary repeat RNA structures play key roles in initiating RAN translation. Here, we briefly review the repeat expansion disorders, their complex pathogenesis and the mechanisms of physiological translation initiation together with the known factors involved in RAN translation. Finally, we discuss research challenges surrounding the understanding of pathogenesis and future directions that may provide opportunities for the development of novel therapeutic approaches for this group of incurable neurodegenerative diseases.

Introduction

Microsatellite expansions have been characterised in a large number of incurable neurodegenerative diseases subdivided into polyglutamine (polyQ) and non-polyglutamine (non-polyQ) disorders [1]. Autosomal-dominant glutamine-encoding CAG repeat expansions in the Huntingtin gene (*HTT*) cause Huntington's disease (HD) [2,3] while CAG repeats in the coding regions of various unrelated ataxin genes lead to spinocerebellar ataxias (SCA) [4,5]. Non-polyQ expansion disorders are caused by various lengths of trinucleotide to hexanucleotide repeat sequences mostly contained within non-coding regions of genes (5'-/3'-untranslated regions (UTR) and introns). These most commonly include: CGG repeats in fragile X mental retardation 1 (*FMRI*) gene in Fragile X-associated syndromes [6]; thousands of CTG/CCTG repeats in the myotonic dystrophies (DM1 and DM2) [7,8] and trinucleotide, pentanucleotide or hexanucleotide repeats in non-polyQ SCAs [9]; GAA repeat expansions in Friedreich's ataxia [10]; thousands of GGGGCC repeats in chromosome 9 open reading frame 72 (*C9ORF72*) in the most common genetic forms of amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) [11,12]. Altogether, we compiled a list of 53 expansion disorders that mainly present with neurodegenerative conditions (Table 1).

Received: 20 November 2020
 Revised: 20 February 2021
 Accepted: 23 February 2021

Version of Record published:
 17 March 2021

Table 1 Microsatellite repeat expansion disorders

Disorder	Gene	Sense repeat	Antisense repeat	Disease length	Location in gene	RAN translated proteins	References
PolyQ microsatellite repeat expansion disorders:							
Dentatorubropallidolusian Atrophy (DRPLA)	<i>ATN1/DRPLA</i>	CAG	Unknown	49–88	Exon 5	Unknown	[127]
Schizophrenia/migraines	<i>KCNN3</i>	CAG	Unknown	>28	Exon 1	Unknown	[128]
Prostate/breast Cancer	<i>AIB/SRC-3</i>	CAG/CAA	Unknown	>23	Exon 20	Unknown	[129]
Huntington's Disease (HD)	<i>HTT</i>	CAG	CTG	36–250	Exon 1	polyS, polyA, polyC, polyL in human brains & <i>in vitro</i> ³	[2,52]
Spinal and Bulbar Muscular Atrophy (SBMA)	<i>AR</i>	CAG	Unknown	38–62	Exon 1	Unknown	[130]
Spinocerebellar Ataxia Type 1 (SCA1)	<i>ATXN1</i>	CAG	Unknown	49–88	Exon 8	Unknown	[131]
Spinocerebellar Ataxia Type 2 (SCA2)	<i>ATXN2</i>	CAG	CTG	33–77	Exon 1	polyQ, polyA, polyS <i>in vitro</i> ³	[132,133]
Spinocerebellar Ataxia Type 3 (SCA3) or Machado-Joseph Disease (MJD)	<i>ATXN3/MJD</i>	CAG	CTG	55–86	Exon 10	polyQ, polyA, polyS <i>in vitro</i> ³	[45,108,134]
Spinocerebellar Ataxia Type 6 (SCA6)	<i>CACNA1A</i>	CAG	Unknown	21–30	Exon 47	Unknown	[135]
Spinocerebellar Ataxia Type 7 (SCA7)	<i>ATXN7</i>	CAG	Unknown	28–120	Exon 3	Unknown	[136]
Spinocerebellar Ataxia Type 17 (SCA17)	<i>TBP</i>	CAG/CAA	Unknown	47–63	Exon 3	Unknown	[137]
Non-polyQ microsatellite repeat expansion disorders:							
Amyotrophic lateral sclerosis (ALS)/ Frontotemporal Dementia (FTD)	<i>C9ORF72</i>	GGGGCC	CCCCGG	30–4400	Intron 1	polyGA, polyGP, polyGR, polyPA, polyPR in human brains & <i>in vitro</i> ³	[11,12,49–51,120]
Baratela-Scott Syndrome	<i>XYLT1</i>	GGC	Untranscribed expansion	>100	Promoter	Unknown	[138]
Blepharophimosis-Ptosis-Epicanthus Inversus Syndactyl	<i>FOXL2</i>	GCG	Unknown	22–24	Exon 1	Unknown	[139]
Cerebellar Ataxia, Neuropathy, Vestibular Areflexia Syndrome (CANVAS)	<i>RFC1</i>	AAGGG	Unknown	400–2000	Intron 2	Unknown	[140]
Cleidocranial Dysplasia	<i>RUNX2/CBFA1</i>	GCG	Unknown	>20	Exon 1	Unknown	[141]
Congenital Central Hypoventilation/Haddad Syndrome	<i>PHOX2B</i>	GCG	Unknown	5–13	Exon 3	Unknown	[142]
Familial adult myoclonic epilepsy (FAME1/BAFME1)	<i>SAMD12</i>	TTTCA/TTTTA	Unknown	440–3680	Intron 4	Unknown	[143]
Fragile X syndrome (FRAXA/FXS)	<i>FMR1</i>	CGG	CCG	>230	5'-UTR	Unknown	[6]
Fragile X-associated tremor/ataxia syndrome (FXTAS)	<i>FMR1</i>	CGG	CCG	55–200	5'-UTR	polyG, polyP, polyA, polyR <i>in vitro</i> ³ and <i>Drosophila</i>	[46,144,145]
Fragile X-associated Primary Ovary Insufficiency (FXPOI)	<i>FMR1</i>	CGG	Not found ¹	55–200	5'-UTR	polyG in biopsied human ovarian stromal cells	[146,147]
Fragile XE mental retardation (FRAXE)	<i>AFF2/FMR2</i>	CGG/CCG	Untranscribed expansion	>200	Promoter	Unknown	[17]

Continued

Table 1 Microsatellite repeat expansion disorders

Disorder	Gene	Sense repeat	Antisense repeat	Disease length	Location in gene	RAN translated proteins	References
Fragile XF syndrome (FRAXF)	<i>TMEN185A</i>	GCC	Unknown	300–500	5'-UTR	Unknown	[148]
FRA2A-associated mental retardation	<i>AFF3</i>	CGG	Unknown	>300	5'-UTR	Unknown	[19]
FRA7A-associated autism spectrum disorder	<i>ZNF713</i>	CGG	Unknown	>85	Intron 1	Unknown	[19]
FRA10A-associated mental retardation	<i>FRA10AC1</i>	CGG	Unknown	>200	5'-UTR	Unknown	[20]
FRA11A-associated mental retardation	<i>C11ORF80</i>	CGG	Unknown	500	5'UTR	Unknown	[149]
FRA12A-associated mental retardation	<i>DIP2B</i>	CGG	Unknown	>50	5'-UTR	Unknown	[21]
FRA16A-associated mental retardation	<i>LOC109617027</i>	CGG	Unknown	1000–1900	5'-UTR	Unknown	[150]
Friedreich's ataxia (FRDA)	<i>FXN/X25</i>	GAA	TTC	>100	Intron 1	Unknown	[10]
Fuchs' Endothelial Corneal Dystrophy (FECD)	<i>TCF4</i>	CTG	CAG	>50	Intron 3	polyC in human corneal endothelium + polyA, polyQ, polyS <i>in vitro</i> ³	[151,152]
Hand-Foot-Genital Syndrome	<i>HOXA13</i>	GCG	Unknown	24–26	Exon 1	Unknown	[153]
Holoprosencephaly	<i>ZIC2</i>	GCG	Unknown	>25	Exon 3	Unknown	[154]
Huntington Disease-Like 2 (HDL2)	<i>JPH3</i>	CAG	CTG	>41	3'-terminal exon	polyQ, polyA, polyS <i>in vitro</i> ³	[45,155]
Jacobsen Syndrome	<i>FRA11B/CBL2</i>	CGG	Not found ¹	100–1000	5'-UTR	Unknown	[156]
Myoclonus Epilepsy of the Unverricht-Lundborg Type	<i>CYSTB</i>	CCCCGCCCGCG	Untranscribed expansion	12–13	Promoter	Unknown	[18]
Congenital Myotonic Dystrophy (CDM)/Steinert's Disease	<i>DMPK</i>	CTG	CAG	50–10000	3'-UTR	Unknown	[157]
Myotonic dystrophy (DM1)	<i>DMPK</i>	CTG	CAG	50–10000	3'-UTR	polyQ in human muscle/blood + polyA, polyS <i>in vitro</i> ³	[7,8,45,157]
Myotonic dystrophy type 2 (DM2)	<i>ZNF9</i>	CCTG	GGAC	75–1100	Intron 1	polyQAGRpolyLPAC in human brains & <i>in vitro</i> ³	[106,158]
Neuronal Intranuclear Inclusion Disease (NIID) & Amyotrophic lateral Sclerosis (ALS)	<i>NOTCH2NLC</i>	GGC	Unknown	>71	5'-UTR	Unknown	[159,160]
Oculopharyngeal Musclar Dystrophy	<i>PABPN1/PABP2</i>	GCG	Unknown	12–17	Exon 1	Unknown	[161]
Pseudoachondroplasia and Multiple Epiphyseal Dysplasia (PSACH/MED)	<i>COMP</i>	GAC	Not found ¹	>6	Exon 13	Unknown	[162]
Spinocerebellar Ataxia Type 8 (SCA8)	<i>ATXN8OS</i> & <i>ATXN8</i>	CTG	CAG	110–250	3'-UTR <i>ATXN8OS</i> ; 5'UTR <i>ATXN8</i>	polyA in human brains + polyS, polyQ <i>in vitro</i> ³	[15,45]
Spinocerebellar ataxia Type 10 (SCA10)	<i>ATXN10</i>	ATTCT	Not found ¹	32–4000	Intron 9	Unknown	[163]
Spinocerebellar ataxia Type 12 (SCA12)	<i>PPP2R2B</i>	CAG	CTG	66–78	5'-UTR	Unknown	[164]
Spinocerebellar ataxia Type 31 (SCA31)	<i>BEAN1</i>	TGGAA	TTCCA	>110	Intron 1	polyWNGME? ² <i>in vitro</i> ³	[165]
Spinocerebellar ataxia Type 36 (SCA36)	<i>NOP56</i>	GGCCTG	Unknown	>100	Intron 1		[166,167]

Continued

Table 1 Microsatellite repeat expansion disorders

Disorder	Gene	Sense repeat	Antisense repeat	Disease length	Location in gene	RAN translated proteins	References
Spinocerebellar ataxia Type 37 (SCA37) ⁴	<i>DAB1</i>	ATTTC	GAAAT	31–75	5'-UTR	Unknown	[168]
Synpolydactyl Type II (SPD)	<i>HOXD13</i>	GCG	Unknown	22–29	Exon 1	Unknown	[169]
X-Linked Dystonia-Parkinsonism (XPD) ⁵	<i>TAF1</i>	CCCTCT	Unknown	35–52	Intron 32	Unknown	[170]
X-Linked Mental Retardation and Abnormal Genitalia (XLAG)	<i>ARX</i>	GCN	Unknown	20	Exon 2	Unknown	[171]
X-Linked Mental Retardation (XMLR)	<i>ARX</i>	GCN	Unknown	18–23	Exon 2	Unknown	[172]
X-linked Mental Retardation with Growth Hormone Deficiency (XLMRGHD)	<i>SOX3</i>	GCN	Unknown	15–26	Exon 1	Unknown	[173]

¹Not found indicates that antisense transcripts were not detected;

²The polypeptide polyWNGME is produced from the intronic repeat expansion, however it can not be confirmed as a RAN translation product due to the presence of an ATG sequence encoding a canonical AUG start codon within the repeat expansion;

³*In vitro* indicates that the RAN-translated proteins were detected from reporter constructs in transfected cell model of diseases;

⁴Not classical expansions but insertions due to replication/recombination/duplication events;

⁵Not classical expansion but insertion due to retrotransposon event.

Pathogenic mechanisms induced by microsatellite repeat expansions

The transcription of repeat expansions located in coding and non-coding regions of genes generates pathological transcripts with polymorphic RNA-repeat sequences. The microsatellite loci are moreover bi-directionally transcribed in HD, DM, C9ORF72-ALS/FTD and in some SCAs and Fragile X-associated syndromes leading to expression of both sense and antisense repeat transcripts. These are thought to cause neuronal injury through complex intertwined mechanisms involving: (i) translation of proteins with expanded stretches of glutamine in poly-Q disorders; (ii) protein gain-of-functions caused by repeat-associated non-AUG (RAN) translation of toxic repeat proteins; (iii) RNA toxic gain-of-functions through the sequestration of RNA-binding proteins within RNA foci and onto repeat transcripts; (iv) protein loss-of-functions via haploinsufficiency (reviewed in [13,14]).

Translation of protein with expanded poly-glutamine domains

Polymorphic CAG repeat expansions in HD and poly-Q SCAs encode long stretches of poly-glutamine and the translation of proteins with polyQ domains. These promote misfolding/ aggregation, inhibit interactions with physiological binding protein partners and generate abnormal interactions with other proteins, mediating thus both toxic protein loss- and gain-of-functions [4]. The non-polyQ disorder SCA8 was initially shown to express expanded CUG repeats in the 3'-UTR of the *ATXN8OS* (*ATXN8 Opposite Strand*) gene [15]. Later, bidirectional expression of CAG expansion transcripts from *ATXN8* were reported and shown to result in the expression and accumulation of a polyQ protein that forms neuronal inclusions [9].

Haploinsufficiency

Loss-of-function of the genes harbouring the repeat expansions can directly contribute to the pathophysiology of the microsatellite repeats. Over 200 CGG repeats in the 5'-UTR of *FMR1* cause Fragile X syndrome (FXS) [16], the most common inherited form of intellectual disability, due to transcriptional silencing induced by DNA methylation of the CGG trinucleotides and loss of the FMRP protein which has roles in synaptic plasticity. A contributory loss-of-function is the likely pathological cause of diseases where the repeat expansions are found in promoters, e.g. fragile-XE mental retardation (*FMR2* gene; [17]) and myoclonus epilepsy of the Unverricht-Lundborg type (*CYSTB* gene; [18]). Loss-of-function is also associated with folate sensitive fragile sites harbouring CGG repeats (*FRA7A*, *FRA10A* and *FRA12A*) through DNA methylation of the repeat expansions [19–21]. Hexanucleotide-repeat expansions in the 5'-UTR region of *C9ORF72* lead to decreased expression levels of *C9ORF72* mRNAs, encoding a protein involved in autophagy regulation [22–25], vesicle trafficking [26,27] and immune response in mice [28,29] in several *in vitro* and *in vivo* models and post-mortem brains [11,30–33]. However, the direct contribution of reduced levels of *C9ORF72* protein to disease pathogenesis is still debated.

Formation of RNA foci and RNA-repeat sequestration of proteins

RNA-mediated cellular toxicity results in either protein gain- and loss- of-functions via sequestration of RNA-processing proteins on repeat transcripts which may either be co-transcriptionally processed or aggregated into RNA foci. Protein loss-of-functions have been implicated in a wide range of expansion disorders via RNA-repeat sequestration of mRNA-binding proteins which may lose their normal cellular functions including: muscleblind-like splicing regulator (MBLN) and CUG-binding protein and ETR3-like factor (CELF) families of proteins in myotonic dystrophy [34–36]; MBLN and other RNA-binding proteins in polyQ disorders [37,38]; Sam68 [39], PUR-alpha, hnRNP A2/B1, CUGBP1 [40,41] in fragile X-associated tremor ataxia syndrome (FXTAS); PUR-alpha, heterogeneous nuclear ribonucleoproteins (hnRNPs) and SR-rich splicing factors (SRSFs) among others in C9ORF72-ALS/FTD [42,43]. On the other hand, toxic protein gain-of-function also occurs through RNA-repeat sequestration of SRSF1 which triggers the nuclear export and subsequent RAN translation of sense and antisense *C9ORF72*-repeat transcripts retaining pathological expansions in intron-1 [44].

RAN translation of toxic repeat proteins

In 2011, Laura Ranum's group demonstrated that CAG-repeat transcripts lacking canonical AUG start codons are remarkably translated into homo-polymeric proteins in all frames (poly-glutamine, poly-serine and poly-alanine) by repeat-associated non-AUG (RAN) translation [45]. RAN-translated poly-alanine proteins driven

Figure 1. Canonical and physiological translation initiation mechanisms.

Part 2 of 2

delivering Met-tRNA^{Met} to the P site of the 40S ribosomal subunit in a GTP-dependent manner, through interaction with both the canonical AUG start codon and near cognate start codons CUG and GCG. Both eIF2A and eIF2D are also able to initiate translation, however this can occur in a GTP-dependent or independent manner, binding either charged or uncharged tRNA^{Met}. eIF2A can additionally bind Leu-tRNA^{CUG} to initiate translation. eIF2A can initiate translation at AUG, CUG and UUG codons, while eIF2D can initiate at AUG, CUG, GCG and UUG codons.

Finally, near-cognate start codons (typically CUG, GUG and UUG), which differ from the AUG start codon by one nucleotide, initiate translation in mammalian cells at a much lower efficiency, using the non-AUG initiator tRNA^{Met} and methionine as the initiating amino acid [80] or an elongator Leu-tRNA^{Leu} at a CUG codon in the case of the major histocompatibility complex class I molecules [81,82]. Near-cognate initiation sites can be used by mismatch recognition of eIF2 α -bound Met-tRNA^{Met} [80], when fidelity of start codon usage is affected depending on secondary structures downstream of the initiation codon and expression levels of other initiation factors such as eIF1 or eIF5 which respectively increases or decreases the fidelity of AUG recognition. Two other initiation factors (eIF2A and eIF2D) can also be used at non-AUG codons to initiate translation in either a GTP-dependent manner through initiator tRNA^{Met} or a GTP-independent manner through Leu-tRNA^{Leu} [83] (Figure 1C).

Mechanisms of RAN translation

RAN translation involves the translation of short repeated RNA sequences in sense and/or antisense transcripts in an AUG-independent manner and in multiple frames. However, how RAN translation occurs and which sets of factors are required for initiation, elongation and potential regulatory controls remains largely unknown, although it is clearly emerging that some features are shared with canonical and/or IRES-mediated initiation [84].

The roles of initiation factors and RNA structures in RAN translation

An improved understanding of the mechanisms driving RAN translation has begun to emerge with a clear role for the general translation initiation factor eIF4A, a DEAD-box RNA helicase, identified in stimulating the canonical translation of mRNAs containing complex secondary structures such as G-quadruplexes in the 5'-UTR of oncogenes [85]. The identification of inhibitors of eIF4A highlighted that the RNA helicase activity of eIF4A plays an essential role in unwinding secondary structures during ribosome scanning [86–88]. G-quadruplex structures are formed in GGGGCC-repeat [89] and CGG-repeat [90] RNAs. The eIF4A inhibitor hippuristanol showed that eIF4A is required for the RAN translation of CGG-repeat expansions in *FMR1* in FXTAS [47] (Figure 2A) and GGGGCC-repeat transcripts in C9ORF72-ALS/FTD [91] (Figure 2B). The eIF4A inhibitor FL3 [92] further confirmed the role of eIF4A in the RAN translation of sense C9ORF72-repeat transcripts. However, the RAN translation mechanisms of antisense C9ORF72-repeats which form a double RNA helix [93] remain completely unknown.

The RNA helicase activity of eIF4A is significantly enhanced by two cofactors, eIF4B and eIF4H [94–96], and eIF4B is essential for the translation of mRNAs with long-structured 5-UTRs independently of eIF4A [97]. Interestingly, recent *Drosophila* screens involving sense C9ORF72-repeat [98] and FXTAS CGG-repeat [99] transcripts identified eIF4B and eIF4H as disease modifiers of the RAN translation with down-regulation of eIF4B or eIF4H leading to reduced RAN translation and associated toxicity, ameliorating *Drosophila* eye neurodegenerative phenotypes and life span (Figure 2A,B). Interestingly, sequestration of eIF4H by GGGGCC-repeat sequences was previously reported [43]. DDX3X, another RNA helicase which is required for the resolution of RNA–RNA structures in long GC-rich 5'-UTRs, is also implicated in the RAN translation of FXTAS CGG-repeats, with suppression of this helicase inhibiting RAN translation and rescuing associated toxicity in *Drosophila* and primary neurons [99]. However, the role of DDX3X in RAN translation appears sequence-specific since the depletion and overexpression of DDX3X respectively lead to increased and reduced DPR levels in C9ORF72-ALS lymphoblasts [100]. Interestingly, the ribosomal protein RPS25, involved in IRES translation [101], also behaves differently during RAN translation of FXTAS CGG-repeats [99] and sense C9ORF72-repeats [102]. Suppression of RPS25 in a FXTAS *Drosophila* model enhanced RAN-translated protein production and associated toxicity [99], while suppression of RPS25 reduced DPR production and rescued associated toxicity in yeast, *Drosophila* and human C9ORF72-ALS/FTD models [102].

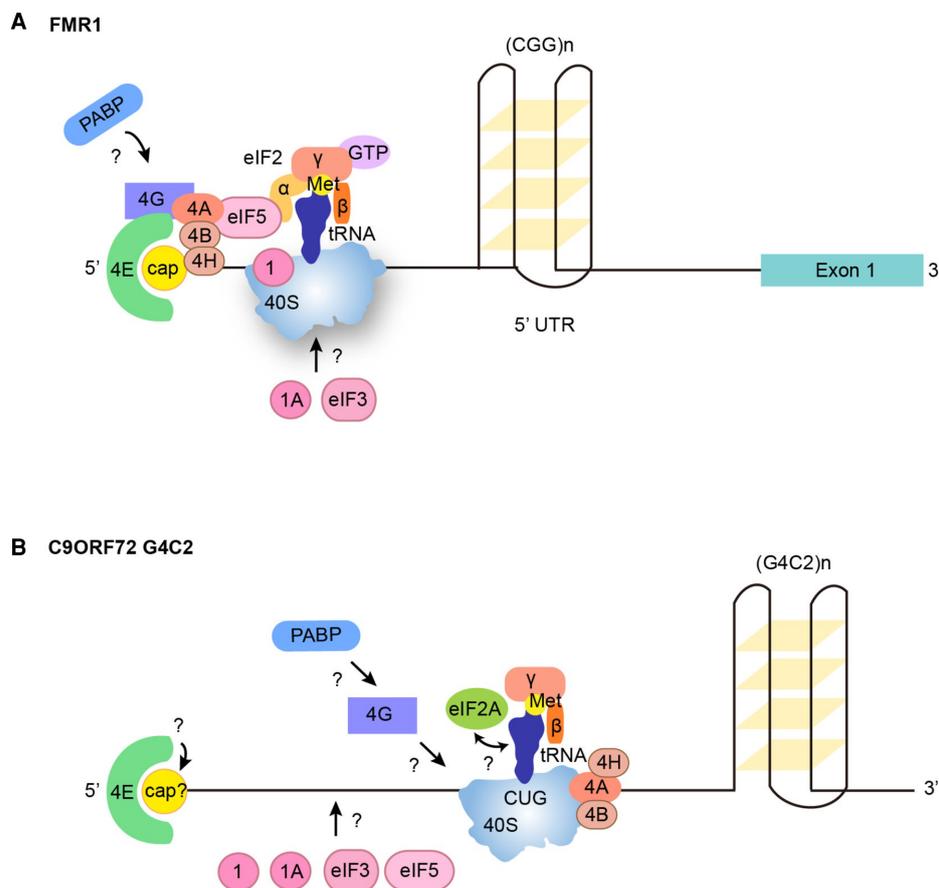


Figure 2. Known RNA structures and protein factors involved in RAN translation.

(A) RAN translation of *FMR1* occurs in a Cap-, eIF4E- and eIF4A-dependent manner along with eIF4A cofactors eIF4H and eIF4B, recruiting the 40S ribosome and eIF2 α -bound Met-tRNA^{Met} to the near-cognate ACG start codon upstream of the CGG repeat expansion. Regulation of start codon fidelity through eIF1 and eIF5 is important. Any potential role of the translation initiation factors PABP, eIF4G, eIF1A and eIF3 remain unknown. (B) RAN translation of the GGGGCC repeat expansion from *C9ORF72* occurs in an eIF4A-dependent manner to recruit the 40S ribosome subunit and eIF2A-bound Met-tRNA^{Met} to the near cognate CUG start codon upstream of the hexanucleotide repeat expansion. The eIF4A cofactors eIF4B and eIF4H have been shown to be disease modifiers and are involved. Contradictory evidence for the role of the m⁷cap and eIF4E factor indicates that further elucidation of their roles is required. Any potential role of the translation initiation factors PABP, eIF4G, eIF1, eIF1A, eIF3 and eIF5 still remain unknown. The mechanisms involved in the RAN translation of *C9ORF72* antisense CCGG-repeat transcripts have not yet been explored.

RAN translation of reporter constructs require both a m⁷G-cap and eIF4E for FXTAS CGG-repeats [47] and sense *C9ORF72*-repeats [91]. Accordingly, the eIF4E competitive inhibitor m⁷G-cap analogue (m⁷GpppG) prevented RAN translation of FXTAS CGG-repeats [47] and sense *C9ORF72*-repeats [91,92]. However, another study using a bicistronic reporter construct with all-frame stop codons prior to the initiating start codon reported that RAN translation of *C9ORF72*-repeat transcripts still occurred, suggesting recruitment of ribosomes in a cap-independent manner [103]. eIF4E was shown to be important for the RAN translation of sense *C9ORF72*-repeat transcripts using the 4EIRCat inhibitor [92] however, it was also reported that depletion of eIF4E does not result in a reduction in RAN translation [103] (Figure 2A,B). It thus clearly appears that the sequence-specific context surrounding repeat expansions regulate the mechanisms of RAN translation.

Additional translation initiation factors involved in start codon fidelity are implicated in RAN translation. eIF1 is important in increasing AUG start codon fidelity and overexpression reduces RAN translation and associated toxicity in FXTAS *Drosophila* [99]. eIF5, on the other hand, relaxes start codon fidelity and suppression

of eIF5 reduces RAN translation and associated toxicity in FXTAS *Drosophila* [99] (Figure 2A). SCA8 has been shown to require eIF3f during RAN translation of poly-serine and poly-alanine proteins [104]. eIF3f is a non-core component of eIF3 which plays a role in regulating both canonical and IRES-dependent translation [105]. Interestingly, the poly-S proteins produced by RAN translation in SCA8 [104] and HD [52] and poly-QAGR/LPAC in DM2 [106] accumulate in white matter brain regions, where eIF3f levels are elevated compared with grey matter [107], further supporting a potential role of eIF3f in the RAN translation and/or its regulation for some repeat expansions.

The roles of near-cognate and non-cognate initiation codons in non-AUG translation initiation

Complex secondary RNA-repeat structures such as G-quadruplexes appear to play direct initiating roles through altered ribosome scanning and/or recruitment of ribosomes at near-cognate initiation codons, which differ from AUG by only one nucleotide, and non-cognate initiation codons that differ by more than one nucleotide. FXTAS CGG-repeats within the 5'-UTR of *FMR1* initiate RAN translation of FMRpolyG at a near-cognate ACG codon embedded in a putative Kozak element 32 nucleotides upstream of the repeats in the +1 reading frame, while FMRpolyA initiates at a non-cognate GCG codon within the repeats in a +2 reading frame [47,48]. Mass spectrometry analysis also indicated that polyA RAN-translated proteins from the SCA8 CAG repeats are initiated at non-cognate GCA codons throughout the repeat tract [45] while non-cognate CUU and ACU codons are used to initiate RAN translation of polyQ upstream of the SCA3 CAG repeats and polyA proteins likely within the repeats [108] in transfected cell models. Sense *C9ORF72*-repeat transcripts initiate RAN translation with a Met-tRNA^{Met} through eIF2A at a Kozak-embedded CUG codon located 24 nucleotides upstream of the repeat sequence in the +1 reading frame of transfected reporter constructs [91,92,109].

The integrated stress response enhances non-canonical translation

Under non-stressed conditions, RAN translation of sense *C9ORF72*-repeats is strongly inhibited by an upstream open reading frame (uORF) of 55 nucleotides which is located in intron 1 and flanked by an AUG start codon and 2 downstream stop codons (UGA and UAA) [92] (Figure 3A). However, the integrated stress response (ISR), which is stimulated during disease progression, leads to phosphorylation of eIF2 α and down-regulation of canonical translation due to poor recruitment of the 60S subunit of the ribosome and inhibition of translation initiation. This results in read-through of the uORF and subsequent initiation at the downstream CUG codon responsible for RAN translation of the hexanucleotide-repeats [92] (Figure 3B). In neurons, eIF2 α phosphorylation is important for synaptic plasticity and rapid activity-dependent alterations of synaptic proteins [110,111]. Following cellular stress and activation of the ISR, a shift in translation occurs towards a subset of transcripts with 5'-UTRs containing uORFs, e.g. ATF4 and CHOP [112], cellular IRES e.g. HIAP2, HIF1 α and VEGF [113] and non-AUG start codons e.g. EPRS and major histocompatibility class I antigens [81,114], allowing these transcripts to escape eIF2 α -phosphorylated translational inhibition. Phosphorylation of eIF2 α increases non-canonical translation and increasing ISR occur concomitantly with increased RAN translation of both *FMR1* and *C9ORF72* repeats in a positive feedback loop [91,103,109,115]. It is noteworthy that ISR-resistant translation involves eIF2D, eIF2A and eIF5B [116–119] and as aforementioned, RAN translation of sense *C9ORF72*-repeats initiates at a CUG codon through eIF2A [109] (Figure 3B). Taken together, these studies strongly support a model in which increased ISR upon disease progression down-regulates canonical translation of the uORF to stimulate RAN translation in a positive disease-enhancing feedback loop.

Conclusions

The pathogenesis driven by nucleic acid repeat expansions and RAN-translated products is complex and still poorly understood. Multiple mechanisms of neuronal injury involve toxic RNA gain-of-functions, haploinsufficiency as well as protein gain-of-functions via canonical translation of proteins with extended polyQ domains and/or RAN translation of toxic repeat polypeptides which have been characterised *in vitro* in 13 reporter repeat expansion cell models and in patient bio-samples from seven diseases, including SCA8, DM1 [45,106], *C9ORF72*-ALS [49–51,120] and HD [52]. However, the molecular mechanisms involved in RAN translation remain poorly understood, hindering thus the development of therapeutic approaches for this incurable group of diseases. In addition, it has remained challenging to dissect how the life-long expression of repeat transcripts

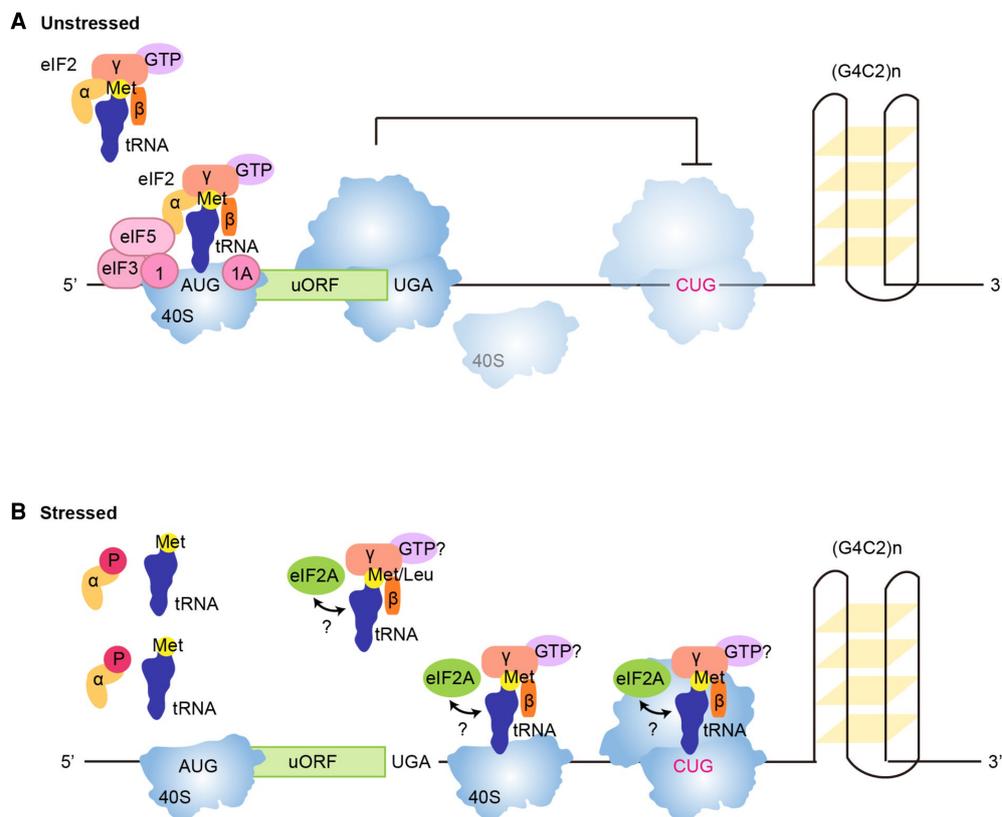


Figure 3. The role of a uORF in the translation of RAN proteins from pathological sense GGGGCC C9ORF72 repeat expansions.

(A) In unstressed cellular conditions, a uORF of 55 nucleotides in length within intron 1 of *C9ORF72* inhibits RAN translation of the downstream repeat expansion. The uORF is translated through the canonical translation machinery and ribosomes are unable to reassemble on the mRNA for initiation at the downstream CUG RAN initiation codon. (B) Following cellular stress, phosphorylation of eIF2 α prevents its binding of Met-tRNA^{Met} results in an inhibition of eIF2-driven canonical translation. Consequently, the alternative tRNA recruiting factor eIF2A is able to initiate non-canonical translation. The scanning 40S ribosomal complex scans through the uORF and eIF2A initiates RAN translation at the downstream CUG codon.

and RAN-translated proteins contribute to the pathogenesis of these progressive adult-onset diseases. In the past few years, growing evidence has implicated RAN translation as one of the main drivers of neurotoxicity in *C9ORF72*-ALS/FTD models. On the other hand, FXTAS was initially thought to be caused by intranuclear retention of transcripts and sequestration of splicing factors [40,41] however, the later discovery of RAN translation in the same model challenged this view [46]. Similarly in *C9ORF72*-ALS, increasing the number of intranuclear RNA foci does not appear to alter neuronal survival or global RNA processing while expression of DPRs drives the neurodegeneration process [44,54,121]. Sequestration of RNA-binding factors on repeat transcripts might not only necessarily implicate loss-of-function mechanisms, as often hypothesised, as cells can use compensatory mechanisms to up-regulate the transcription/translation of the sequestered proteins. So far, there has not been any evidence demonstrating that RNA foci and sequestration of proteins on RNA-repeats trigger protein loss-of-function mechanisms. In contrast, RNA-repeat sequestration of SRSF1 produces protein gain-of-function by driving the nuclear export of intron-retaining *C9ORF72*-repeat pre-mRNAs and subsequent RAN translation of cytotoxic DPRs in the cytoplasm [44].

If expression of polymeric repeat proteins can kill cells and animal models, it is difficult to evaluate which levels are RAN-translated in patients and which expression levels trigger toxicity depending on the nature of the repeat expansion, disease and cell type. Understanding the molecular mechanisms of RAN translation will be key to improving our understanding of pathogenesis. Additional mechanisms such as ribosome

frameshifting events also need to be explored. For example, frameshifting was suggested to occur during translation of CAG-repeat expanded transcripts in the -1 frame in SCA3 [122–124] as well as into the -1 [125] and $+1$ [126] directions in HD, however, the production of chimeric repeat proteins was not directly evidenced. RAN translation of three C9ORF72-related DPRs encoded from the three reading frames of sense repeat transcripts suggested that RAN translation initiates at multiple initiation sites [91,109], however, mutation of a near-cognate CUG start codon also inhibited the production of all sense DPRs, suggesting the occurrence of potential frameshifting mechanisms that remain to be demonstrated [92].

Perspectives

- Mechanisms of RAN translation and RAN-translated proteins/peptides still remain poorly characterised despite discovery in DM1/SCA8 patient samples in 2011, and later in C9ORF72-ALS and HD in 2013 and 2015, respectively.
- RAN-translation occurs in absence of the canonical methionine start codon, in all frames, and from coding and non-coding regions of transcripts encoding proteins of various functions. So far, it is known to involve RNA secondary structures formed by repeat sequences and general translation initiation factors, which exhibit/stimulate RNA-helicase activities, or play a role in the fidelity of start codon recognition.
- In the future, it will be fundamental to fully identify the RAN-translation machinery components and mechanisms in the context of the sequences flanking each microsatellite repeat expansions, as well as examine further the pathological contribution of RAN-translated products for the future development of much-needed disease treatments.

Competing Interests

The authors declare that there are no competing interests associated with the manuscript.

Funding

Royal Society International Exchanges 2017 Cost Share Grant IEC\R3\170103 (G.M.H. and K.Y.C.) and MOST-RS International Exchange Scheme: MOST 108-2911-I-005-504 (K.Y.C. and G.M.H.). G.M.H. also acknowledge support from the MRC (Medical Research Council) New Investigator grant MR/R024162/1 and Biotechnology and Biological Sciences Research Council (BBSRC) grant BB/S005277/1.

Open Access Statement

Open access for this article was enabled by the participation of University of Sheffield in an all-inclusive Read & Publish pilot with Portland Press and the Biochemical Society under a transformative agreement with JISC.

Author Contributions

L.M.C. and G.M.H. wrote the review. W.P.H. designed the figures. All authors contributed to the proofreading and editing of the manuscript.

Abbreviations

ALS, amyotrophic lateral sclerosis; C9ORF72, chromosome 9 open reading frame 72; DM1, dystrophy type 1; DPRs, dipeptide-repeat proteins; FMR1, fragile X mental retardation 1; FTD, frontotemporal dementia; FXS, Fragile X syndrome; FXTAS, fragile X-associated tremor ataxia syndrome; HD, Huntington's disease; IRES, internal ribosome entry site; ISR, integrated stress response; RAN, repeat-associated non-AUG; SCA8, spinocerebellar ataxia type 8; UTR, untranslated regions.

References

- 1 Paulson, H. (2018) Repeat expansion diseases. *Handb. Clin. Neurol.* **147**, 105–123 <https://doi.org/10.1016/B978-0-444-63233-3.00009-9>
- 2 MacDonald, M.E., Ambrose, C.M., Duyao, M.P., Myers, R.H., Lin, C., Srinidhi, L. et al. (1993) A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* **72**, 971–983 [https://doi.org/10.1016/0092-8674\(93\)90585-E](https://doi.org/10.1016/0092-8674(93)90585-E)
- 3 La Spada, A.R., Paulson, H.L. and Fischbeck, K.H. (1994) Trinucleotide repeat expansion in neurological disease. *Ann. Neurol.* **36**, 814–822 <https://doi.org/10.1002/ana.410360604>
- 4 Orr, H.T. and Zoghbi, H.Y. (2007) Trinucleotide repeat disorders. *Annu. Rev. Neurosci.* **30**, 575–621 <https://doi.org/10.1146/annurev.neuro.29.051605.113042>
- 5 Whaley, N.R., Fujioka, S. and Wszolek, Z.K. (2011) Autosomal dominant cerebellar ataxia type I: a review of the phenotypic and genotypic characteristics. *Orphanet J. Rare Dis.* **6**, 33 <https://doi.org/10.1186/1750-1172-6-33>
- 6 Verkerk, A.J., Pieretti, M., Sutcliffe, J.S., Fu, Y.H., Kuhl, D.P., Pizzuti, A. et al. (1991) Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* **65**, 905–914 [https://doi.org/10.1016/0092-8674\(91\)90397-H](https://doi.org/10.1016/0092-8674(91)90397-H)
- 7 Brook, J.D., McCurrach, M.E., Harley, H.G., Buckler, A.J., Church, D., Aburatani, H. et al. (1992) Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell* **68**, 799–808 [https://doi.org/10.1016/0092-8674\(92\)90154-5](https://doi.org/10.1016/0092-8674(92)90154-5)
- 8 Fu, Y.H., Pizzuti, A., Fenwick, R.G., King, J., Rajnarayan, S., Dunne, P.W. et al. (1992) An unstable triplet repeat in a gene related to myotonic muscular dystrophy. *Science* **255**, 1256–1258 <https://doi.org/10.1126/science.1546326>
- 9 Moseley, M.L., Zu, T., Ikeda, Y., Gao, W., Mosemiller, A.K., Daughters, R.S. et al. (2006) Bidirectional expression of CUG and CAG expansion transcripts and intranuclear polyglutamine inclusions in spinocerebellar ataxia type 8. *Nat. Genet.* **38**, 758–769 <https://doi.org/10.1038/ng1827>
- 10 Campuzano, V., Montermini, L., Molto, M.D., Pianese, L., Cossee, M., Cavalcanti, F. et al. (1996) Friedreich's ataxia: Autosomal recessive disease caused by an intronic GAA triplet repeat expansion. *Science* **271**, 1423–1427 <https://doi.org/10.1126/science.271.5254.1423>
- 11 DeJesus-Hernandez, M., Mackenzie, I.R., Boeve, B.F., Boxer, A.L., Baker, M., Rutherford, N.J. et al. (2011) Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-Linked FTD and ALS. *Neuron* **72**, 245–256 <https://doi.org/10.1016/j.neuron.2011.09.011>
- 12 Renton, A.E., Majounie, E., Waite, A., Simón-Sánchez, J., Rollinson, S., Gibbs, J.R. et al. (2011) A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* **72**, 257–268 <https://doi.org/10.1016/j.neuron.2011.09.010>
- 13 Rohilla, K.J. and Gagnon, K.T. (2017) RNA biology of disease-associated microsatellite repeat expansions. *Acta Neuropathol. Commun.* **5**, 63–22 <https://doi.org/10.1186/s40478-017-0468-y>
- 14 Sznajder, Ł.J. and Swanson, M.S. (2019) Short tandem repeat expansions and RNA-mediated pathogenesis in myotonic dystrophy. *Int. J. Mol. Sci.* **20**, 3365 <https://doi.org/10.3390/ijms20133365>
- 15 Koob, M.D., Moseley, M.L., Schut, L.J., Benzow, K.A., Bird, T.D., Day, J.W. et al. (1999) An untranslated CTG expansion causes a novel form of spinocerebellar ataxia (SCA8). *Nat. Genet.* **21**, 379–384 <https://doi.org/10.1038/7710>
- 16 McLennan, Y., Polussa, J., Tassone, F. and Hagerman, R. (2011) Fragile X syndrome. *Curr. Genomics* **12**, 216–224 <https://doi.org/10.2174/138920211795677886>
- 17 Geck, J., Gedeon, A.K., Sutherland, G.R. and Mulley, J.C. (1996) Identification of the gene FMR2, associated with FRA1E mental retardation. *Nat. Genet.* **13**, 105–108 <https://doi.org/10.1038/ng0596-105>
- 18 Lalioti, M.D., Scott, H.S., Buresi, C., Rossier, C., Bottani, A., Morris, M.A. et al. (1997) Dodecamer repeat expansion in cystatin B gene in progressive myoclonus epilepsy. *Nature* **386**, 847–851 <https://doi.org/10.1038/386847a0>
- 19 Metsu, S., Rainger, J.K., Debacker, K., Bernhard, B., Rooms, L., Grafodatskaya, D. et al. (2014) A CGG-repeat expansion mutation in ZNF713 causes FRA7A: association with autistic spectrum disorder in two families. *Hum. Mutat.* **35**, 1295–1300 <https://doi.org/10.1002/humu.22683>
- 20 Sarafidou, T., Kahl, C., Martinez-Garay, I., Mangelsdorf, M., Gesk, S., Baker, E. et al. (2004) Folate-sensitive fragile site FRA10A is due to an expansion of a CGG repeat in a novel gene, FRA10AC1, encoding a nuclear protein. *Genomics* **84**, 69–81 <https://doi.org/10.1016/j.ygeno.2003.12.017>
- 21 Winnepeninckx, B., Debacker, K., Ramsay, J., Smeets, D., Smits, A., FitzPatrick, D.R. et al. (2007) CGG-repeat expansion in the DIP2B gene is associated with the fragile site FRA12A on chromosome 12q13.1. *Am. J. Hum. Genet.* **80**, 221–231 <https://doi.org/10.1086/510800>
- 22 Sellier, C., Campanari, M.-L., Julie Corbier, C., Gaucherot, A., Kolb-Cheyne, I., Oulad-Abdelghani, M. et al. (2016) Loss of C9ORF72 impairs autophagy and synergizes with polyQ Ataxin-2 to induce motor neuron dysfunction and cell death. *EMBO J.* **35**, 1276–1297 <https://doi.org/10.15252/embj.201593350>
- 23 Sullivan, P.M., Zhou, X., Robins, A.M., Paushter, D.H., Kim, D., Smolka, M.B. et al. (2016) The ALS/FTLD associated protein C9orf72 associates with SMCR8 and WDR41 to regulate the autophagy-lysosome pathway. *Acta Neuropathol. Commun.* **4**, 51–16 <https://doi.org/10.1186/s40478-016-0324-5>
- 24 Webster, C.P., Smith, E.F., Bauer, C.S., Moller, A., Hautbergue, G.M., Ferraiuolo, L. et al. (2016) The C9orf72 protein interacts with Rab1a and the ULK1 complex to regulate initiation of autophagy. *EMBO J.* **35**, 1656–1676 <https://doi.org/10.15252/embj.201694401>
- 25 Yang, M., Liang, C., Swaminathan, K., Herrlinger, S., Lai, F., Shiekhhattar, R. et al. (2016) A C9ORF72/SMCR8-containing complex regulates ULK1 and plays a dual role in autophagy. *Sci Adv* **2**, e1601167 <https://doi.org/10.1126/sciadv.1601167>
- 26 Farg, M.A., Sundaramoorthy, V., Sultana, J.M., Yang, S., Atkinson, R.A.K., Levina, V. et al. (2014) C9ORF72, implicated in amyotrophic lateral sclerosis and frontotemporal dementia, regulates endosomal trafficking. *Hum. Mol. Genet.* **23**, 3579–3595 <https://doi.org/10.1093/hmg/ddu068>
- 27 Aoki, Y., Manzano, R., Lee, Y., Dafinca, R., Aoki, M., Douglas, A.G.L. et al. (2017) C9orf72 and RAB7L1 regulate vesicle trafficking in amyotrophic lateral sclerosis and frontotemporal dementia. *Brain* **140**, 887–897 <https://doi.org/10.1093/brain/awx024>
- 28 O'Rourke, J.G., Bogdanik, L., Yanez, A., Lall, D., Wolf, A.J., Muhammad, A.K.M.G. et al. (2016) C9orf72 is required for proper macrophage and microglial function in mice. *Science* **351**, 1324–1329 <https://doi.org/10.1126/science.aaf1064>
- 29 Atanasio, A., Decman, V., White, D., Ramos, M., Ikiz, B., Lee, H.-C. et al. (2016) C9orf72 ablation causes immune dysregulation characterized by leukocyte expansion, autoantibody production, and glomerulonephropathy in mice. *Sci. Rep.* **6**, 23204–23214 <https://doi.org/10.1038/srep23204>
- 30 Belzil, V.V., Bauer, P.O., Prudencio, M., Gendron, T.F., Stetler, C.T., Yan, I.K. et al. (2013) Reduced C9orf72 gene expression in c9FTD/ALS is caused by histone trimethylation, an epigenetic event detectable in blood. *Acta Neuropathol.* **126**, 895–905 <https://doi.org/10.1007/s00401-013-1199-1>

- 31 Ciura, S., Lattante, S., Le Ber, I., Latouche, M., Tostivint, H., Brice, A. et al. (2013) Loss of function of C9orf72 causes motor deficits in a zebrafish model of amyotrophic lateral sclerosis. *Ann. Neurol.* **74**, 180–187 <https://doi.org/10.1002/ana.23946>
- 32 Waite, A.J., Bäumer, D., East, S., Neal, J., Morris, H.R., Ansorge, O. et al. (2014) Reduced C9orf72 protein levels in frontal cortex of amyotrophic lateral sclerosis and frontotemporal degeneration brain with the C9ORF72 hexanucleotide repeat expansion. *Neurobiol. Aging* **35**, 1779.e5–1779.e13 <https://doi.org/10.1016/j.neurobiolaging.2014.01.016>
- 33 Shi, Y., Lin, S., Staats, K.A., Li, Y., Chang, W.-H., Hung, S.-T. et al. (2018) Haploinsufficiency leads to neurodegeneration in C9ORF72 ALS/FTD human induced motor neurons. *Nat. Med.* **24**, 313–325 <https://doi.org/10.1038/nm.4490>
- 34 Cooper, T.A., Wan, L. and Dreyfuss, G. (2009) RNA and disease. *Cell* **136**, 777–793 <https://doi.org/10.1016/j.cell.2009.02.011>
- 35 Batra, R., Charizanis, K., Manchanda, M., Mohan, A., Li, M., Finn, D.J. et al. (2014) Loss of MBNL leads to disruption of developmentally regulated alternative polyadenylation in RNA-mediated disease. *Mol. Cell* **56**, 311–322 <https://doi.org/10.1016/j.molcel.2014.08.027>
- 36 Lee, K.-Y., Chang, H.-C., Seah, C. and Lee, L.-J. (2019) Deprivation of muscleblind-like proteins causes deficits in cortical neuron distribution and morphological changes in dendritic spines and postsynaptic densities. *Front. Neuroanat.* **13**, 75 <https://doi.org/10.3389/fnana.2019.00075>
- 37 Li, L.-B., Yu, Z., Teng, X. and Bonini, N.M. (2008) RNA toxicity is a component of ataxin-3 degeneration in drosophila. *Nature* **453**, 1107–1111 <https://doi.org/10.1038/nature06909>
- 38 Tsoi, H. and Chan, H.Y.E. (2013) Expression of expanded CAG transcripts triggers nucleolar stress in huntington's disease. *Cerebellum* **12**, 310–312 <https://doi.org/10.1007/s12311-012-0447-6>
- 39 Sellier, C., Rau, F.E.D.E.R., Liu, Y., Tassone, F., Hukema, R.K., Gattoni, R. et al. (2010) Sam68 sequestration and partial loss of function are associated with splicing alterations in FXTAS patients. *EMBO J.* **29**, 1248–1261 <https://doi.org/10.1038/emboj.2010.21>
- 40 Jin, P., Duan, R., Qurashi, A., Qin, Y., Tian, D., Rosser, T.C. et al. (2007) Pur α binds to rCGG repeats and modulates repeat-mediated neurodegeneration in a drosophila model of fragile X tremor/Ataxia syndrome. *Neuron* **55**, 556–564 <https://doi.org/10.1016/j.neuron.2007.07.020>
- 41 Sofola, O.A., Jin, P., Qin, Y., Duan, R., Liu, H., de Haro, M. et al. (2007) RNA-Binding Proteins hnRNP A2/B1 and CUGBP1 suppress fragile X CGG premutation repeat-induced neurodegeneration in a drosophila model of FXTAS. *Neuron* **55**, 565–571 <https://doi.org/10.1016/j.neuron.2007.07.021>
- 42 Lee, Y.-B., Chen, H.-J., Peres, J.N., Gomez-Deza, J., Attig, J., Štálekár, M. et al. (2013) Hexanucleotide repeats in ALS/FTD form length-dependent RNA foci, sequester RNA binding proteins, and are neurotoxic. *Cell Rep.* **5**, 1178–1186 <https://doi.org/10.1016/j.celrep.2013.10.049>
- 43 Cooper-Knock, J., Walsh, M.J., Higginbottom, A., Robin Highley, J., Dickman, M.J., Edbauer, D. et al. (2014) Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. *Brain* **137**, 2040–2051 <https://doi.org/10.1093/brain/awu120>
- 44 Hautbergue, G.M., Castelli, L.M., Ferraiuolo, L., Sanchez-Martinez, A., Cooper-Knock, J., Higginbottom, A. et al. (2017) SRSF1-dependent nuclear export inhibition of C9ORF72 repeat transcripts prevents neurodegeneration and associated motor deficits. *Nat. Commun.* **8**, 16063 <https://doi.org/10.1038/ncomms16063>
- 45 Zu, T., Gibbens, B., Doty, N.S., Gomes-Pereira, M., Huguet, A., Stone, M.D. et al. (2011) Non-ATG-initiated translation directed by microsatellite expansions. *Proc. Natl Acad. Sci. U.S.A.* **108**, 260–265 <https://doi.org/10.1073/pnas.1013343108>
- 46 Todd, P.K., Oh, S.Y., Krans, A., He, F., Sellier, C., Frazer, M. et al. (2013) CGG Repeat-associated translation mediates neurodegeneration in fragile X tremor ataxia syndrome. *Neuron* **78**, 440–455 <https://doi.org/10.1016/j.neuron.2013.03.026>
- 47 Kearse, M.G., Green, K.M., Krans, A., Rodriguez, C.M., Linsalata, A.E., Goldstrohm, A.C. et al. (2016) CGG Repeat-Associated Non-AUG translation utilizes a Cap-Dependent scanning mechanism of initiation to produce toxic proteins. *Mol. Cell* **62**, 314–322 <https://doi.org/10.1016/j.molcel.2016.02.034>
- 48 Sellier, C., Buijss, R.A.M., He, F., Natla, S., Jung, L., Tropel, P. et al. (2017) Translation of expanded CGG repeats into FMRpolyG is pathogenic and may contribute to fragile X tremor ataxia syndrome. *Neuron* **93**, 331–347 <https://doi.org/10.1016/j.neuron.2016.12.016>
- 49 Mori, K., Weng, S.-M., Arzberger, T., May, S., Rentzsch, K., Kremmer, E. et al. (2013) The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTD/ALS. *Science* **339**, 1335–1338 <https://doi.org/10.1126/science.1232927>
- 50 Ash, P.E.A., Bieniek, K.F., Gendron, T.F., Caulfield, T., Lin, W.-L., DeJesus-Hernandez, M. et al. (2013) Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron* **77**, 639–646 <https://doi.org/10.1016/j.neuron.2013.02.004>
- 51 Zu, T., Liu, Y., Baez-Coronel, M., Reid, T., Pletnikova, O., Lewis, J. et al. (2013) RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. *Proc. Natl Acad. Sci. U.S.A.* **110**, E4968–E4977 <https://doi.org/10.1073/pnas.1315438110>
- 52 Bañez-Coronel, M., Ayhan, F., Tarabochia, A.D., Zu, T., Perez, B.A., Tusi, S.K. et al. (2015) RAN translation in huntington disease. *Neuron* **88**, 667–677 <https://doi.org/10.1016/j.neuron.2015.10.038>
- 53 Kwon, I., Xiang, S., Kato, M., Wu, L., Theodoropoulos, P., Wang, T. et al. (2014) Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science* **345**, 1139–1145 <https://doi.org/10.1126/science.1254917>
- 54 Mizielińska, S., Gronke, S., Niccoli, T., Ridler, C.E., Clayton, E.L., Devoy, A. et al. (2014) C9orf72 repeat expansions cause neurodegeneration in drosophila through arginine-rich proteins. *Science* **345**, 1192–1194 <https://doi.org/10.1126/science.1256800>
- 55 Boeynaems, S., Bogaert, E., Kovacs, D., Konijnenberg, A., Timmerman, A., Volkov, A. et al. (2017) Phase separation of C9orf72 dipeptide repeats perturbs stress granule dynamics. *Mol. Cell* **65**, 1044–1045 <https://doi.org/10.1016/j.molcel.2017.02.013>
- 56 Zhang, Y.-J., Gendron, T.F., Ebbert, M.T.W., O'Raw, A.D., Yue, M., Jansen-West, K. et al. (2018) Poly(GR) impairs protein translation and stress granule dynamics in C9orf72-associated frontotemporal dementia and amyotrophic lateral sclerosis. *Nat. Med.* **24**, 1136–1142 <https://doi.org/10.1038/s41591-018-0071-1>
- 57 Hao, Z., Liu, L., Tao, Z., Wang, R., Ren, H., Sun, H. et al. (2019) Motor dysfunction and neurodegeneration in a C9orf72 mouse line expressing poly-PR. *Nat. Commun.* **10**, 2906–2911 <https://doi.org/10.1038/s41467-019-10956-w>
- 58 Moens, T.G., Niccoli, T., Wilson, K.M., Atilano, M.L., Birsá, N., Gittings, L.M. et al. (2019) C9orf72 arginine-rich dipeptide proteins interact with ribosomal proteins in vivo to induce a toxic translational arrest that is rescued by eIF1A. *Acta Neuropathol.* **137**, 487–500 <https://doi.org/10.1007/s00401-018-1946-4>
- 59 Choi, S.Y., Lopez-Gonzalez, R., Krishnan, G., Phillips, H.L., Li, A.N., Seeley, W.W. et al. (2019) C9ORF72-ALS/FTD-associated poly(GR) binds Atp5a1 and compromises mitochondrial function in vivo. *Nat. Neurosci.* **22**, 851–862 <https://doi.org/10.1038/s41593-019-0397-0>
- 60 Zhang, Y.-J., Gendron, T.F., Grima, J.C., Sasaguri, H., Jansen-West, K., Xu, Y.-F. et al. (2016) C9ORF72 poly(GA) aggregates sequester and impair HR23 and nucleocytoplasmic transport proteins. *Nat. Neurosci.* **19**, 668–677 <https://doi.org/10.1038/nn.4272>

- 61 Lee, Y.-B., Baskaran, P., Gomez-Deza, J., Chen, H.-J., Nishimura, A.L., Smith, B.N. et al. (2017) C9orf72 poly GA RAN-translated protein plays a key role in amyotrophic lateral sclerosis via aggregation and toxicity. *Hum. Mol. Genet.* **26**, 4765–4777 <https://doi.org/10.1093/hmg/ddx350>
- 62 Schludi, M.H., Becker, L., Garrett, L., Gendron, T.F., Zhou, Q., Schreiber, F. et al. (2017) Spinal poly-GA inclusions in a C9orf72 mouse model trigger motor deficits and inflammation without neuron loss. *Acta Neuropathol.* **134**, 241–254 <https://doi.org/10.1007/s00401-017-1711-0>
- 63 Schludi, M.H., May, S., Grässer, F.A., Rentzsch, K., Kremmer, E., Küpper, C. et al. (2015) Distribution of dipeptide repeat proteins in cellular models and C9orf72 mutation cases suggests link to transcriptional silencing. *Acta Neuropathol.* **130**, 537–555 <https://doi.org/10.1007/s00401-015-1450-z>
- 64 Lin, Y., Mori, E., Kato, M., Xiang, S., Wu, L., Kwon, I. et al. (2016) Toxic PR poly-Dipeptides encoded by the C9orf72 repeat expansion target LC domain polymers. *Cell* **167**, 789–802.e12 <https://doi.org/10.1016/j.cell.2016.10.003>
- 65 Yin, S., Lopez-Gonzalez, R., Kunz, R.C., Gangopadhyay, J., Borufka, C., Gygi, S.P. et al. (2017) Evidence that C9ORF72 dipeptide repeat proteins associate with U2 snRNP to cause mis-splicing in ALS/FTD patients. *Cell Rep.* **19**, 2244–2256 <https://doi.org/10.1016/j.celrep.2017.05.056>
- 66 Zhang, K., Donnelly, C.J., Haeusler, A.R., Grima, J.C., Machamer, J.B., Steinwald, P. et al. (2015) The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. *Nature* **525**, 56–61 <https://doi.org/10.1038/nature14973>
- 67 Boeynaems, S., Bogaert, E., Michiels, E., Gijssels, I., Sieben, A., Jovičić, A. et al. (2016) Drosophila screen connects nuclear transport genes to DPR pathology in c9ALS/FTD. *Sci. Rep.* **6**, 20877–8 <https://doi.org/10.1038/srep20877>
- 68 Walker, C., Herranz-Martin, S., Karyka, E., Liao, C., Lewis, K., Elsayed, W. et al. (2017) C9orf72 expansion disrupts ATM-mediated chromosomal break repair. *Nat. Neurosci.* **20**, 1225–1235 <https://doi.org/10.1038/nn.4604>
- 69 Lopez-Gonzalez, R., Lu, Y., Gendron, T.F., Karydas, A., Tran, H., Yang, D. et al. (2016) Poly(GR) in C9ORF72-Related ALS/FTD compromises mitochondrial function and increases oxidative stress and DNA damage in iPSC-Derived motor neurons. *Neuron* **92**, 383–391 <https://doi.org/10.1016/j.neuron.2016.09.015>
- 70 Kanekura, K., Yagi, T., Cammack, A.J., Mahadevan, J., Kuroda, M., Harms, M.B. et al. (2016) Poly-dipeptides encoded by the C9ORF72 repeats block global protein translation. *Hum. Mol. Genet.* **25**, 1803–1813 <https://doi.org/10.1093/hmg/ddw052>
- 71 Hartmann, H., Hornburg, D., Czuppa, M., Bader, J., Michaelsen, M., Farny, D. et al. (2018) Proteomics and C9orf72 neuropathology identify ribosomes as poly-GR/PR interactors driving toxicity. *Life Sci. Alliance* **1**, e201800070 <https://doi.org/10.26508/lsa.201800070>
- 72 Gupta, R., Lan, M., Mojsilovic-Petrovic, J., Choi, W.H., Safren, N., Barmada, S. et al. (2017) The proline/Arginine dipeptide from hexanucleotide repeat expanded C9ORF72 inhibits the proteasome. *eNeuro* **4**, ENEURO.0249–16.2017 <https://doi.org/10.1523/ENEURO.0249-16.2017>
- 73 Guo, Q., Lehmer, C., Martínez-Sánchez, A., Rudack, T., Beck, F., Hartmann, H. et al. (2018) In situ structure of neuronal C9orf72 poly-GA aggregates reveals proteasome recruitment. *Cell* **172**, 696–705.e12 <https://doi.org/10.1016/j.cell.2017.12.030>
- 74 Hinnebusch, A.G. (2017) Structural insights into the mechanism of scanning and start codon recognition in Eukaryotic translation initiation. *Trends Biochem. Sci.* **42**, 589–611 <https://doi.org/10.1016/j.tibs.2017.03.004>
- 75 Costello, J., Castelli, L.M., Rowe, W., Kershaw, C.J., Talavera, D., Mohammad-Qureshi, S.S. et al. (2015) Global mRNA selection mechanisms for translation initiation. *Genome Biol.* **16**, 10 <https://doi.org/10.1186/s13059-014-0559-z>
- 76 Kozak, M. (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44**, 283–292 [https://doi.org/10.1016/0092-8674\(86\)90762-2](https://doi.org/10.1016/0092-8674(86)90762-2)
- 77 Komar, A.A. and Hatzoglou, M. (2014) Cellular IRES-mediated translation. *Cell Cycle* **10**, 229–240 <https://doi.org/10.4161/cc.10.2.14472>
- 78 Filbin, M.E. and Kieft, J.S. (2009) Toward a structural understanding of IRES RNA function. *Curr. Opin. Struct. Biol.* **19**, 267–276 <https://doi.org/10.1016/j.sbi.2009.03.005>
- 79 Jackson, R.J., Hellen, C.U.T. and Pestova, T.V. (2010) The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat. Rev. Mol. Cell Biol.* **11**, 113–127 <https://doi.org/10.1038/nrm2838>
- 80 Peabody, D.S. (1989) Translation initiation at non-AUG triplets in mammalian cells. *J. Biol. Chem.* **264**, 5031–5035 [https://doi.org/10.1016/S0021-9258\(18\)83694-8](https://doi.org/10.1016/S0021-9258(18)83694-8)
- 81 Schwab, S.R., Shugart, J.A., Horng, T., Malarkannan, S. and Shastri, N. (2004) Unanticipated antigens: translation initiation at CUG with leucine. *PLoS Biol.* **2**, e366 <https://doi.org/10.1371/journal.pbio.0020366>
- 82 Starck, S.R., Jiang, V., Pavon-Eterod, M., Prasad, S., McCarthy, B., Pan, T. et al. (2012) Leucine-tRNA initiates at CUG start codons for protein synthesis and presentation by MHC class I. *Science* **336**, 1719–1723 <https://doi.org/10.1126/science.1220270>
- 83 Kearse, M.G. and Wilusz, J.E. (2017) Non-AUG translation: a new start for protein synthesis in eukaryotes. *Genes Dev.* **31**, 1717–1731 <https://doi.org/10.1101/gad.305250.117>
- 84 Nguyen, L., Cleary, J.D. and Ranum, L.P.W. (2019) Repeat-associated Non-ATG translation: Molecular mechanisms and contribution to neurological disease. *Annu. Rev. Neurosci.* **42**, 227–247 <https://doi.org/10.1146/annurev-neuro-070918-050405>
- 85 Wolfe, A.L., Singh, K., Zhong, Y., Drewe, P., Rajasekhar, V.K., Sanghvi, V.R. et al. (2017) RNA G-quadruplexes cause eIF4A-dependent oncogene translation in cancer. *Nature* **513**, 65–70 <https://doi.org/10.1038/nature13485>
- 86 Bordeleau, M.-E., Cencic, R., Lindqvist, L., Oberer, M., Northcote, P., Wagner, G. et al. (2006) RNA-mediated sequestration of the RNA helicase eIF4A by pateamine A inhibits translation initiation. *Chem. Biol.* **13**, 1287–1295 <https://doi.org/10.1016/j.chembiol.2006.10.005>
- 87 Cruz-Migoni, A., Hautbergue, G.M., Artymiuk, P.J., Baker, P.J., Bokori-Brown, M., Chang, C.-T. et al. (2011) A *Burkholderia pseudomallei* toxin inhibits helicase activity of translation factor eIF4A. *Science* **334**, 821–824 <https://doi.org/10.1126/science.1211915>
- 88 Boussemart, L., Malka-Mahieu, H., Girault, I., Allard, D., Hemmingsson, O., Tomasic, G. et al. (2014) eIF4F is a nexus of resistance to anti-BRAF and anti-MEK cancer therapies. *Nature* **513**, 105–109 <https://doi.org/10.1038/nature13572>
- 89 Reddy, K., Zamiri, B., Stanley, S.Y.R., Macgregor, Jr, R.B. and Pearson, C.E. (2013) The disease-associated r(GGGGCC) nRepeat from the C9orf72 Gene forms tract length-dependent Uni- and multimolecular RNA G-quadruplex structures. *J. Biol. Chem.* **288**, 9860–9866 <https://doi.org/10.1074/jbc.C113.452532>
- 90 Blice-Baum, A.C. and Mihailescu, M.-R. (2014) Biophysical characterization of G-quadruplex forming FMR1 mRNA and of its interactions with different fragile X mental retardation protein isoforms. *RNA* **20**, 103–114 <https://doi.org/10.1261/rna.041442.113>
- 91 Green, K.M., Gloneburg, M.R., Kearse, M.G., Flores, B.N., Linsalata, A.E., Fedak, S.J. et al. (2017) RAN translation at C9orf72-associated repeat expansions is selectively enhanced by the integrated stress response. *Nat. Commun.* **8**, 2005 <https://doi.org/10.1038/s41467-017-02200-0>

- 92 Tabet, R., Schaeffer, L., Freyermuth, F., Jambeau, M., Workman, M., Lee, C.-Z. et al. (2018) CUG initiation and frameshifting enable production of dipeptide repeat proteins from ALS/FTD C9ORF72 transcripts. *Nat. Commun.* **9**, 152–114 <https://doi.org/10.1038/s41467-017-02643-5>
- 93 Dodd, D.W., Tomchick, D.R., Corey, D.R. and Gagnon, K.T. (2016) Pathogenic C9ORF72 antisense repeat RNA forms a double helix with tandem C:C mismatches. *Biochemistry* **55**, 1283–1286 <https://doi.org/10.1021/acs.biochem.6b00136>
- 94 Sun, Y., Atas, E., Lindqvist, L., Sonenberg, N., Pelletier, J. and Meller, A. (2012) The eukaryotic initiation factor eIF4H facilitates loop-binding, repetitive RNA unwinding by the eIF4A DEAD-box helicase. *Nucleic Acids Res.* **40**, 6199–6207 <https://doi.org/10.1093/nar/gks278>
- 95 Harms, U., Andreou, A.Z., Gubaev, A. and Klostermeier, D. (2014) eIF4B, eIF4G and RNA regulate eIF4A activity in translation initiation by modulating the eIF4A conformational cycle. *Nucleic Acids Res.* **42**, 7911–7922 <https://doi.org/10.1093/nar/gku440>
- 96 García-García, C., Frieda, K.L., Feoktistova, K., Fraser, C.S. and Block, S.M. (2015) RNA.BIOCHEMISTRY. Factor-dependent processivity in human eIF4A DEAD-box helicase. *Science* **348**, 1486–1488 <https://doi.org/10.1126/science.aaa5089>
- 97 Sen, N.D., Zhou, F., Harris, M.S., Ingolia, N.T. and Hinnebusch, A.G. (2016) eIF4B stimulates translation of long mRNAs with structured 5' UTRs and low closed-loop potential but weak dependence on eIF4G. *Proc. Natl Acad. Sci. U.S.A.* **113**, 10464–10472 <https://doi.org/10.1073/pnas.1612398113>
- 98 Goodman, L.D., Prudencio, M., Srinivasan, A.R., Rifai, O.M., Lee, V.M.Y., Petrucelli, L. et al. (2019) eIF4B and eIF4H mediate GR production from expanded G4C2 in a drosophila model for C9orf72-associated ALS. *Acta Neuropathol. Commun.* **7**, 62 <https://doi.org/10.1186/s40478-019-0711-9>
- 99 Linsalata, A.E., He, F., Malik, A.M., Glineburg, M.R., Green, K.M., Natla, S. et al. (2019) DDX3X and specific initiation factors modulate FMR1 repeat-associated non-AUG-initiated translation. *EMBO Rep.* **20**, e47498 <https://doi.org/10.15252/embr.201847498>
- 100 Cheng, W., Wang, S., Zhang, Z., Morgens, D.W., Hayes, L.R., Lee, S. et al. (2019) CRISPR-Cas9 Screens identify the RNA helicase DDX3X as a repressor of C9ORF72 (GGGGCC)_n repeat-Associated Non-AUG translation. *Neuron* **104**, 885–888 <https://doi.org/10.1016/j.neuron.2019.09.003>
- 101 Hertz, M.I., Landry, D.M., Willis, A.E., Luo, G. and Thompson, S.R. (2013) Ribosomal protein S25 dependency reveals a common mechanism for diverse internal ribosome entry sites and ribosome shunting. *Mol. Cell. Biol.* **33**, 1016–1026 <https://doi.org/10.1128/MCB.00879-12>
- 102 Yamada, S.B., Gendron, T.F., Niccoli, T., Genuth, N.R., Grosely, R., Shi, Y. et al. (2019) RPS25 is required for efficient RAN translation of C9orf72 and other neurodegenerative disease-associated nucleotide repeats. *Nat. Neurosci.* **22**, 1383–1388 <https://doi.org/10.1038/s41593-019-0455-7>
- 103 Cheng, W., Wang, S., Mestre, A.A., Fu, C., Makarem, A., Xian, F. et al. (2018) C9ORF72 GGGGCC repeat-associated non-AUG translation is upregulated by stress through eIF2 α phosphorylation. *Nat. Commun.* **9**, 51–12 <https://doi.org/10.1038/s41467-017-02495-z>
- 104 Ayhan, F., Perez, B.A., Shorrock, H.K., Zu, T., Bañez-Coronel, M., Reid, T. et al. (2018) SCA8 RAN polySer protein preferentially accumulates in white matter regions and is regulated by eIF3F. *EMBO J.* **37**, 1–15 <https://doi.org/10.15252/emboj.201899023>
- 105 Hashem, Y., Georges des, A., Dhote, V., Langlois, R., Liao, H.Y., Grassucci, R.A. et al. (2013) Hepatitis-C-virus-like internal ribosome entry sites displace eIF3 to gain access to the 40S subunit. *Nature* **503**, 539–543 <https://doi.org/10.1038/nature12658>
- 106 Zu, T., Cleary, J.D., Liu, Y., Bañez-Coronel, M., Bubenik, J.L., Ayhan, F. et al. (2017) RAN translation regulated by muscleblind proteins in myotonic dystrophy type 2. *Neuron* **95**, 1292–1295 <https://doi.org/10.1016/j.neuron.2017.08.039>
- 107 Mills, J.D., Kavanagh, T., Kim, W.S., Chen, B.J., Kawahara, Y., Halliday, G.M. 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 <https://doi.org/10.1371/journal.pone.0078480>
- 108 Jazurek-Ciesiolka, M., Ciesiolka, A., Komur, A.A., Urbanek-Trzeciak, M.O., Krzyzosiak, W.J. and Fiszer, A. (2020) RAN translation of the expanded CAG repeats in the SCA3 disease context. *J. Mol. Biol.* **432**, 166699 <https://doi.org/10.1016/j.jmb.2020.10.033>
- 109 Sonobe, Y., Ghadge, G., Masaki, K., Sendoel, A., Fuchs, E. and Roos, R.P. (2018) Translation of dipeptide repeat proteins from the C9ORF72 expanded repeat is associated with cellular stress. *Neurobiol. Dis.* **116**, 155–165 <https://doi.org/10.1016/j.nbd.2018.05.009>
- 110 Costa-Mattioli, M., Sossin, W.S., Klann, E. and Sonenberg, N. (2009) Translational control of long-lasting synaptic plasticity and memory. *Neuron* **61**, 10–26 <https://doi.org/10.1016/j.neuron.2008.10.055>
- 111 Bellato, H.M. and Hajj, G.N.M. (2016) Translational control by eIF2 α in neurons: beyond the stress response. *Cytoskeleton (Hoboken)* **73**, 551–565 <https://doi.org/10.1002/cm.21294>
- 112 Barbosa, C., Peixeiro, I. and Romão, L. (2013) Gene expression regulation by upstream open reading frames and human disease. *PLoS Genet.* **9**, e1003529 <https://doi.org/10.1371/journal.pgen.1003529>
- 113 Holcik, M. and Sonenberg, N. (2005) Translational control in stress and apoptosis. *Nat. Rev. Mol. Cell Biol.* **6**, 318–327 <https://doi.org/10.1038/nrm1618>
- 114 Young, S.K., Baird, T.D. and Wek, R.C. (2016) Translation regulation of the glutamyl-prolyl-tRNA synthetase gene EPRS through bypass of upstream open reading frames with noncanonical initiation codons. *J. Biol. Chem.* **291**, 10824–10835 <https://doi.org/10.1074/jbc.M116.722256>
- 115 Westergard, T., McAvoy, K., Russell, K., Wen, X., Pang, Y., Morris, B. et al. (2019) Repeat-associated non-AUG translation in C9orf72-ALS/FTD is driven by neuronal excitation and stress. *EMBO Mol. Med.* **11**, 1–14 <https://doi.org/10.15252/emmm.201809423>
- 116 Pestova, T.V., de Breyne, S., Pisarev, A.V., Abaeva, I.S. and Hellen, C.U.T. (2008) eIF2-dependent and eIF2-independent modes of initiation on the CSFV IRES: a common role of domain II. *EMBO J.* **27**, 1060–1072 <https://doi.org/10.1038/emboj.2008.49>
- 117 Dmitriev, S.E., Terenin, I.M., Andreev, D.E., Ivanov, P.A., Dunaevsky, J.E., Merrick, W.C. et al. (2010) GTP-independent tRNA delivery to the ribosomal P-site by a novel eukaryotic translation factor. *J. Biol. Chem.* **285**, 26779–26787 <https://doi.org/10.1074/jbc.M110.119693>
- 118 Skabkin, M.A., Skabkina, O.V., Dhote, V., Komar, A.A., Hellen, C.U.T. and Pestova, T.V. (2010) Activities of ligatin and MCT-1/DENR in eukaryotic translation initiation and ribosomal recycling. *Genes Dev.* **24**, 1787–1801 <https://doi.org/10.1101/gad.1957510>
- 119 Starck, S.R., Tsai, J.C., Chen, K., Shodiya, M., Wang, L., Yahiro, K. et al. (2016) Translation from the 5' untranslated region shapes the integrated stress response. *Science* **351**, aad3867 <https://doi.org/10.1126/science.aad3867>
- 120 Gendron, T.F., Bieniek, K.F., Zhang, Y.-J., Jansen-West, K., Ash, P.E.A., Caulfield, T. et al. (2013) Antisense transcripts of the expanded C9ORF72 hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS. *Acta Neuropathol.* **126**, 829–844 <https://doi.org/10.1007/s00401-013-1192-8>
- 121 Tran, H., Almeida, S., Moore, J., Gendron, T.F., Chalasani, U., Lu, Y. et al. (2015) Differential toxicity of nuclear RNA foci versus dipeptide repeat proteins in a drosophila model of C9ORF72 FTD/ALS. *Neuron* **87**, 1207–1214 <https://doi.org/10.1016/j.neuron.2015.09.015>
- 122 Gaspar, C., Jannatipour, M., Dion, P., Laganière, J., Sequeiros, J., Brais, B. et al. (2000) CAG tract of MJD-1 may be prone to frameshifts causing polyalanine accumulation. *Hum. Mol. Genet.* **9**, 1957–1966 <https://doi.org/10.1093/hmg/9.13.1957>

- 123 Toulouse, A., Au-Yeung, F., Gaspar, C., Roussel, J., Dion, P. and Rouleau, G.A. (2005) Ribosomal frameshifting on MJD-1 transcripts with long CAG tracts. *Hum. Mol. Genet.* **14**, 2649–2660 <https://doi.org/10.1093/hmg/ddi299>
- 124 Stochmanski, S.J., Therrien, M., Laganière, J., Rochefort, D., Laurent, S., Karemera, L. et al. (2012) Expanded ATXN3 frameshifting events are toxic in drosophila and mammalian neuron models. *Hum. Mol. Genet.* **21**, 2211–2218 <https://doi.org/10.1093/hmg/dds036>
- 125 Girstmair, H., Saffert, P., Rode, S., Czech, A., Holland, G., Bannert, N. et al. (2013) Depletion of cognate charged transfer RNA causes translational frameshifting within the expanded CAG stretch in huntingtin. *Cell Rep.* **3**, 148–159 <https://doi.org/10.1016/j.celrep.2012.12.019>
- 126 Saffert, P., Adaml, F., Schieweck, R., Atkins, J.F. and Ignatova, Z. (2016) An expanded CAG repeat in huntingtin causes +1 frameshifting. *J. Biol. Chem.* **291**, 18505–18513 <https://doi.org/10.1074/jbc.M116.744326>
- 127 Koide, R., Ikeuchi, T., Onodera, O., Tanaka, H., Igarashi, S., Endo, K. et al. (1994) Unstable expansion of CAG repeat in hereditary dentatorubral-pallidolusian atrophy (DRPLA). *Nat. Genet.* **6**, 9–13 <https://doi.org/10.1038/ng0194-9>
- 128 Chandy, K.G., Fantino, E., Wittekindt, O., Kalman, K., Tong, L.L., Ho, T.H. et al. (1998) Isolation of a novel potassium channel gene hSKCa3 containing a polymorphic CAG repeat: a candidate for schizophrenia and bipolar disorder? *Mol. Psychiatry* **3**, 32–37 <https://doi.org/10.1038/sj.mp.4000353>
- 129 Hsing, A.W., Chokkalingam, A.P., Gao, Y.-T., Wu, G., Wang, X., Deng, J. et al. (2002) Polymorphic CAG/CAA repeat length in the AIB1/SRC-3 gene and prostate cancer risk: a population-based case-control study. *Cancer Epidemiol. Biomarkers Prev.* **11**, 337–341 PMID:11927493
- 130 La Spada, A.R., Wilson, E.M., Lubahn, D.B., Harding, A.E. and Fischbeck, K.H. (1991) Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* **352**, 77–79 <https://doi.org/10.1038/352077a0>
- 131 Orr, H.T., Chung, M.Y., Banfi, S., Kwiatkowski, T.J., Servadio, A., Beaudet, A.L. et al. (1993) Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1. *Nat. Genet.* **4**, 221–226 <https://doi.org/10.1038/ng0793-221>
- 132 Pulst, S.M., Nechiporuk, A., Nechiporuk, T., Gispert, S., Chen, X.N., Lopes-Cendes, I. et al. (1996) Moderate expansion of a normally biallelic trinucleotide repeat in spinocerebellar ataxia type 2. *Nat. Genet.* **14**, 269–276 <https://doi.org/10.1038/ng1196-269>
- 133 Scoles, D.R., Ho, M.H.T., Dansithong, W., Pflieger, L.T., Petersen, L.W., Thai, K.K. et al. (2015) Repeat associated Non-AUG translation (RAN translation) dependent on sequence downstream of the ATXN2 CAG repeat. *PLoS ONE* **10**, e0128769 <https://doi.org/10.1371/journal.pone.0128769>
- 134 Kawaguchi, Y., Okamoto, T., Taniwaki, M., Aizawa, M., Inoue, M., Katayama, S. et al. (1994) CAG expansions in a novel gene for Machado-Joseph disease at chromosome 14q32.1. *Nat. Genet.* **8**, 221–228 <https://doi.org/10.1038/ng1194-221>
- 135 Zhuchenko, O., Bailey, J., Bonnen, P., Ashizawa, T., Stockton, D.W., Amos, C. et al. (1997) Autosomal dominant cerebellar ataxia (SCA6) associated with small polyglutamine expansions in the alpha 1A-voltage-dependent calcium channel. *Nat. Genet.* **15**, 62–69 <https://doi.org/10.1038/ng0197-62>
- 136 Lindblad, K., Savontaus, M.L., Stevanin, G., Holmberg, M., Digre, K., Zander, C. et al. (1996) An expanded CAG repeat sequence in spinocerebellar ataxia type 7. *Genome Res.* **6**, 965–971 <https://doi.org/10.1101/gr.6.10.965>
- 137 Zühlke, C.H., Spranger, M., Spranger, S., Voigt, R., Lanz, M., Gehlken, U. et al. (2003) SCA17 caused by homozygous repeat expansion in TBP due to partial isodisomy 6. *Eur. J. Hum. Genet.* **11**, 629–632 <https://doi.org/10.1038/sj.ejhg.5201018>
- 138 LaCroix, A.J., Stabley, D., Sahraroui, R., Adam, M.P., Mehaffey, M., Kernan, K. et al. (2019) GGC repeat expansion and exon 1 methylation of XYL1T1 is a common pathogenic variant in Baratela-Scott syndrome. *Am. J. Hum. Genet.* **104**, 35–44 <https://doi.org/10.1016/j.ajhg.2018.11.005>
- 139 Crisponi, L., Deiana, M., Loi, A., Chiappe, F., Uda, M., Amati, P. et al. (2001) The putative forkhead transcription factor FOXL2 is mutated in blepharophimosis/ptosis/epicanthus inversus syndrome. *Nat. Genet.* **27**, 159–166 <https://doi.org/10.1038/84781>
- 140 Cortese, A., Simone, R., Sullivan, R., Vandrovcova, J., Tariq, H., Yau, W.Y. et al. (2019) Biallelic expansion of an intronic repeat in RFC1 is a common cause of late-onset ataxia. *Nat. Genet.* **51**, 649–658 <https://doi.org/10.1038/s41588-019-0372-4>
- 141 Mundlos, S., Otto, F., Mundlos, C., Mulliken, J.B., Aylsworth, A.S., Albrigst, S. et al. (1997) Mutations involving the transcription factor CBFA1 cause cleidocranial dysplasia. *Cell* **89**, 773–779 [https://doi.org/10.1016/S0092-8674\(00\)80260-3](https://doi.org/10.1016/S0092-8674(00)80260-3)
- 142 Amiel, J., Laudier, B., Attié-Bitach, T., Trang, H., de Pontual, L., Gener, B. et al. (2003) Polyalanine expansion and frameshift mutations of the paired-like homeobox gene PHOX2B in congenital central hypoventilation syndrome. *Nat. Genet.* **33**, 459–461 <https://doi.org/10.1038/ng1130>
- 143 Ishiura, H., Doi, K., Mitsui, J., Yoshimura, J., Matsukawa, M.K., Fujiyama, A. et al. (2018) Expansions of intronic TTTC A and TTTTA repeats in benign adult familial myoclonic epilepsy. *Nat. Genet.* **50**, 581–590 <https://doi.org/10.1038/s41588-018-0067-2>
- 144 Hagerman, R.J., Leehey, M., Heinrichs, W., Tassone, F., Wilson, R., Hills, J. et al. (2001) Intention tremor, parkinsonism, and generalized brain atrophy in male carriers of fragile X. *Neurology* **57**, 127–130 <https://doi.org/10.1212/WNL.57.1.127>
- 145 Hagerman, P.J. and Hagerman, R.J. (2015) Fragile X-associated tremor/ataxia syndrome. *Ann. N.Y. Acad. Sci.* **1338**, 58–70 <https://doi.org/10.1111/nyas.12693>
- 146 Sherman, S.L. (2000) Premature ovarian failure in the fragile X syndrome. *Am. J. Med. Genet.* **97**, 189–194 [https://doi.org/10.1002/1096-8628\(200023\)97:3<189::AID-AJMG1036>3.0.CO;2-J](https://doi.org/10.1002/1096-8628(200023)97:3<189::AID-AJMG1036>3.0.CO;2-J)
- 147 Buijssen, R.A.M., Visser, J.A., Kramer, P., Severijnen, E.A.W.F.M., Gearing, M., Charlet-Berguerand, N. et al. (2016) Presence of inclusions positive for polyglycine containing protein, FMRpolyG, indicates that repeat-associated non-AUG translation plays a role in fragile X-associated primary ovarian insufficiency. *Hum. Reprod.* **31**, 158–168 <https://doi.org/10.1093/humrep/dev280>
- 148 Parrish, J.E., Oostra, B.A., Verkerk, A.J., Richards, C.S., Reynolds, J., Spikes, A.S. et al. (1994) Isolation of a GCC repeat showing expansion in FRAXF, a fragile site distal to FRAXA and FRAXE. *Nat. Genet.* **8**, 229–235 <https://doi.org/10.1038/ng1194-229>
- 149 Debacker, K., Winnepenninckx, B., Longman, C., Colgan, J., Tolmie, J., Murray, R. et al. (2007) The molecular basis of the folate-sensitive fragile site FRA11A at 11q13. *Cytogenet. Genome Res.* **119**, 9–14 <https://doi.org/10.1159/000109612>
- 150 Nancarrow, J.K., Kremer, E., Holman, K., Eyre, H., Doggett, N.A., Le Paslier, D. et al. (1994) Implications of FRA16A structure for the mechanism of chromosomal fragile site genesis. *Science* **264**, 1938–1941 <https://doi.org/10.1126/science.8009225>
- 151 Wieben, E.D., Aleff, R.A., Tosakulwong, N., Butz, M.L., Highsmith, W.E., Edwards, A.O. et al. (2012) A common trinucleotide repeat expansion within the transcription factor 4 (TCF4, E2-2) gene predicts fuchs corneal dystrophy. *PLoS ONE* **7**, e49083 <https://doi.org/10.1371/journal.pone.0049083>
- 152 Soragni, E., Petrosyan, L., Rinkoski, T.A., Wieben, E.D., Baratz, K.H., Fautsch, M.P. et al. (2018) Repeat-Associated Non-ATG (RAN) translation in fuchs' endothelial corneal dystrophy. *Invest. Ophthalmol. Vis. Sci.* **59**, 1888–1896 <https://doi.org/10.1167/iows.17-23265>
- 153 Goodman, F.R., Bacchelli, C., Brady, A.F., Brueeton, L.A., Fryns, J.P., Mortlock, D.P. et al. (2000) Novel HOXA13 mutations and the phenotypic spectrum of hand-foot-genital syndrome. *Am. J. Hum. Genet.* **67**, 197–202 <https://doi.org/10.1086/302961>

- 154 Brown, S.A., Warburton, D., Brown, L.Y., Yu, C.Y., Roeder, E.R., Stengel-Rutkowski, S. et al. (1998) Holoprosencephaly due to mutations in ZIC2, a homologue of Drosophila odd-paired. *Nat. Genet.* **20**, 180–183 <https://doi.org/10.1038/2484>
- 155 Holmes, S.E., O'Hearn, E., Rosenblatt, A., Callahan, C., Hwang, H.S., Ingersoll-Ashworth, R.G. et al. (2001) A repeat expansion in the gene encoding junctophilin-3 is associated with Huntington disease-like 2. *Nat. Genet.* **29**, 377–378 <https://doi.org/10.1038/ng760>
- 156 Jones, C., Penny, L., Mattina, T., Yu, S., Baker, E., Voullaire, L. et al. (1995) Association of a chromosome deletion syndrome with a fragile site within the proto-oncogene CBL2. *Nature* **376**, 145–149 <https://doi.org/10.1038/376145a0>
- 157 Mahadevan, M., Tsilifidis, C., Sabourin, L., Shutler, G., Amemiya, C., Jansen, G. et al. (1992) Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of the gene. *Science* **255**, 1253–1255 <https://doi.org/10.1126/science.1546325>
- 158 Liquori, C.L., Ricker, K., Moseley, M.L., Jacobsen, J.F., Kress, W., Naylor, S.L. et al. (2001) Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. *Science* **293**, 864–867 <https://doi.org/10.1126/science.1062125>
- 159 Sone, J., Mitsunashi, S., Fujita, A., Mizuguchi, T., Hamanaka, K., Mori, K. et al. (2019) Long-read sequencing identifies GGC repeat expansions in NOTCH2NL associated with neuronal intranuclear inclusion disease. *Nat. Genet.* **51**, 1215–1221 <https://doi.org/10.1038/s41588-019-0459-y>
- 160 Yuan, Y., Liu, Z., Hou, X., Li, W., Ni, J., Huang, L. et al. (2020) Identification of GGC repeat expansion in the NOTCH2NL gene in amyotrophic lateral sclerosis. *Neurology* **95**, e3394–e3405 <https://doi.org/10.1212/WNL.00000000000010945>
- 161 Brais, B., Bouchard, J.P., Xie, Y.G., Rochefort, D.L., Chrétien, N., Tomé, F.M. et al. (1998) Short GCG expansions in the PABP2 gene cause oculopharyngeal muscular dystrophy. *Nat. Genet.* **18**, 164–167 <https://doi.org/10.1038/ng0298-164>
- 162 Délot, E., King, L.M., Briggs, M.D., Wilcox, W.R. and Cohn, D.H. (1999) Trinucleotide expansion mutations in the cartilage oligomeric matrix protein (COMP) gene. *Hum. Mol. Genet.* **8**, 123–128 <https://doi.org/10.1093/hmg/8.1.123>
- 163 Matsuura, T., Yamagata, T., Burgess, D.L., Rasmussen, A., Grewal, R.P., Watase, K. et al. (2000) Large expansion of the ATTCT pentanucleotide repeat in spinocerebellar ataxia type 10. *Nat. Genet.* **26**, 191–194 <https://doi.org/10.1038/79911>
- 164 Holmes, S.E., O'Hearn, E.E., McInnis, M.G., Gorelick-Feldman, D.A., Kleiderlein, J.J., Callahan, C. et al. (1999) Expansion of a novel CAG trinucleotide repeat in the 5' region of PPP2R2B is associated with SCA12. *Nat. Genet.* **23**, 391–392 <https://doi.org/10.1038/70493>
- 165 Sato, N., Amino, T., Kobayashi, K., Asakawa, S., Ishiguro, T., Tsunemi, T. et al. (2009) AR TICLESpinocerebellar ataxia type 31 is associated with 'Inserted' penta-nucleotide repeats containing (TGGAA). *Am. J. Hum. Genet.* **85**, 544–557 <https://doi.org/10.1016/j.ajhg.2009.09.019>
- 166 Kobayashi, H., Abe, K., Matsuura, T., Ikeda, Y., Hitomi, T., Akechi, Y. et al. (2011) REPOR t expansion of intronic GGCTG hexanucleotide repeat in NOP56 causes SCA36, a type of spinocerebellar ataxia accompanied by motor neuron involvement. *Am. J. Hum. Genet.* **89**, 121–130 <https://doi.org/10.1016/j.ajhg.2011.05.015>
- 167 Todd, T.W., McEachin, Z.T., Chew, J., Burch, A.R., Jansen-West, K., Tong, J. et al. (2020) Hexanucleotide repeat expansions in c9FTD/ALS and SCA36 confer selective patterns of neurodegeneration In vivo. *Cell Rep.* **31**, 107616 <https://doi.org/10.1016/j.celrep.2020.107616>
- 168 Seixas, A.I., Loureiro, J.R., Costa, C., Ordóñez-Ugalde, A., Marcelino, H., Oliveira, C.L. et al. (2017) A pentanucleotide ATTC repeat insertion in the non-coding region of DAB1, mapping to SCA37, causes spinocerebellar ataxia. *Am. J. Hum. Genet.* **101**, 87–103 <https://doi.org/10.1016/j.ajhg.2017.06.007>
- 169 Akarsu, A.N., Stoilov, I., Yilmaz, E., Sayli, B.S. and Sarfarazi, M. (1996) Genomic structure of HOXD13 gene: a nine polyalanine duplication causes synpolydactyly in two unrelated families. *Hum. Mol. Genet.* **5**, 945–952 <https://doi.org/10.1093/hmg/5.7.945>
- 170 Bragg, D.C., Mangalaphiban, K., Vaine, C.A., Kulkarni, N.J., Shin, D., Yadav, R. et al. (2017) Disease onset in X-linked dystonia-parkinsonism correlates with expansion of a hexameric repeat within an SVA retrotransposon in TAF1. *Proc. Natl Acad. Sci. U.S.A.* **114**, E11020–E11028 <https://doi.org/10.1073/pnas.1712526114>
- 171 Kato, M., Das, S., Petras, K., Kitamura, K., Morohashi, K.-I., Abuelo, D.N. et al. (2004) Mutations of ARX are associated with striking pleiotropy and consistent genotype-phenotype correlation. *Hum. Mutat.* **23**, 147–159 <https://doi.org/10.1002/humu.10310>
- 172 Strømme, P., Mangelsdorf, M.E., Shaw, M.A., Lower, K.M., Lewis, S.M.E., Bruyere, H. et al. (2002) Mutations in the human ortholog of aristaless cause X-linked mental retardation and epilepsy. *Nat. Genet.* **30**, 441–445 <https://doi.org/10.1038/ng862>
- 173 Laumonier, F., Ronce, N., Hamel, B.C.J., Thomas, P., Lespinasse, J., Raynaud, M. et al. (2002) Transcription factor SOX3 is involved in X-linked mental retardation with growth hormone deficiency. *Am. J. Hum. Genet.* **71**, 1450–1455 <https://doi.org/10.1086/344661>