Translation of Expanded CGG Repeats into FMRpolyG Is Pathogenic and May Contribute to Fragile X Tremor Ataxia Syndrome

Highlights

- CGG repeats in the 5′ UTR of FMR1 are translated through initiation to an ACG codon

- Translation of CGG repeats in the polyglycine protein, FMRpolyG, is toxic in mice

- FMRpolyG binds and disrupts protein of the nuclear lamina

Authors

Chantal Sellier, Ronald A.M. Buijsen, Fang He, ..., Cecile Martinat, Peter K. Todd, Nicolas Charlet-Berguerand

Correspondence

sellier@igbmc.fr (C.S.),
ncharlet@igbmc.fr (N.C.-B.)

In Brief

Sellier et al. show that translation of expanded CGG repeats located in the 5′ UTR of the FMR1 gene require an upstream ACG near-cognate initiation codon. Translation of CGG repeats into a short polyglycine-containing protein, FMRpolyG, is pathogenic in mouse models.

CrossMark

CellPress

Translation of Expanded CGG Repeats into FMRpolyG Is Pathogenic and May Contribute to Fragile X Tremor Ataxia Syndrome

Chantal Sellier,[1,*] Ronald A.M. Buijsen,[2] Fang He,[3,4,5] Sam Natla,[3] Laura Jung,[1] Philippe Tropel,[1] Angeline Gaucherot,[1] Hugues Jacobs,[1,6] Hamid Meziane,[1,6] Alexandre Vincent,[1] Marie-France Champy,[1,6] Tania Sorg,[1,6] Guillaume Pavlovic,[1,6] Marie Wattenhofer-Donze,[1,6] Marie-Christine Birling,[1,6] Mustapha Oulad-Abdelghani,[1] Pascal Eberling,[1] Frank Ruffenach,[1] Mathilde Joint,[1] Mathieu Anheim,[7] Veronica Martinez-Cerdeno,[8,9,10] Flora Tassone,[10] Rob Willemsen,[2] Renate K. Hukema,[2] Stéphane Viville,[1,11,12] Cecile Martinat,[13] Peter K. Todd,[3,4] and Nicolas Charlet-Berguerand[1,14,15,16,17,*]

[1]Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), INSERM U964, CNRS UMR7104, University of Strasbourg, 67400 Illkirch, France
[2]Department of Clinical Genetics, Erasmus MC, 3015 Rotterdam, the Netherlands
[3]Department of Neurology, University of Michigan, Ann Arbor, MI 48109, USA
[4]Veteran Association Health System, Ann Arbor, MI 48105, USA
[5]Department of Biological and Health Sciences, Texas A&M University – Kingsville, Kingsville, TX 78363, USA
[6]PHENOMIN, Institut Clinique de la Souris (ICS), INSERM U964, CNRS UMR7104, University of Strasbourg, 67400 Illkirch, France
[7]Department of Neurology, University Hospital of Strasbourg, Hôpital de Hautepierre, 67200 Strasbourg, France
[8]Department of Pathology and Laboratory Medicine, University of California, Davis, Sacramento, CA 95817, USA
[9]Institute for Pediatric Regenerative Medicine and Shriners Hospitals for Children Northern California, Sacramento, CA 95817, USA
[10]M.I.N.D. Institute, University of California, Davis, Health System, Sacramento, CA 95817, USA
[11]Laboratoire de Diagnostic Génétique, UF3472 – Infertilité, Nouvel Hôpital Civil, 1 place de l'Hôpital, 67091 Strasbourg, France
[12]IPPTS, 3 rue Koeberlé, 67000 Strasbourg, France
[13]INSERM/UEVE UMR 861, I-STEM, AFM, 91030 Evry, France
[14]Université de Strasbourg, 67000 Strasbourg, France
[15]Centre National de la Recherche Scientifique, UMR7104, 67400 Illkirch, France
[16]Institut National de la Santé et de la Recherche Médicale, U964, 67400 Illkirch, France
[17]Lead Contact
*Correspondence: sellier@igbmc.fr (C.S.), ncharlet@igbmc.fr (N.C.-B.)
http://dx.doi.org/10.1016/j.neuron.2016.12.016

SUMMARY

Fragile X-associated tremor/ataxia syndrome (FXTAS) is a neurodegenerative disorder caused by a limited expansion of CGG repeats in the 5′ UTR of *FMR1*. Two mechanisms are proposed to cause FXTAS: RNA gain-of-function, where CGG RNA sequesters specific proteins, and translation of CGG repeats into a polyglycine-containing protein, FMRpolyG. Here we developed transgenic mice expressing CGG repeat RNA with or without FMRpolyG. Expression of FMRpolyG is pathogenic, while the sole expression of CGG RNA is not. FMRpolyG interacts with the nuclear lamina protein LAP2β and disorganizes the nuclear lamina architecture in neurons differentiated from FXTAS iPS cells. Finally, expression of LAP2β rescues neuronal death induced by FMRpolyG. Overall, these results suggest that translation of expanded CGG repeats into FMRpolyG alters nuclear lamina architecture and drives pathogenesis in FXTAS.

INTRODUCTION

Fragile X-associated tremor/ataxia syndrome (FXTAS) is a neurodegenerative disorder caused by a pre-mutation of 55 to 200 CGG repeats in the 5′ untranslated region (UTR) of the *fragile X mental retardation 1* (*FMR1*) gene (Hagerman et al., 2001). The carrier prevalence of the CGG pre-mutation is approximately 1 of ∼200 females and ∼450 males, but due to incomplete penetrance, it is estimated that 1 in 3,000 men older than 50 years will develop FXTAS (Jacquemont et al., 2004; Tassone et al., 2012). The clinical features of FXTAS include progressive intention tremor and gait ataxia, frequently accompanied by progressive cognitive decline, parkinsonism, peripheral neuropathy, and autonomic dysfunctions (Jacquemont et al., 2003). Principal neuropathologies of FXTAS include mild brain atrophy and white matter lesions with the presence of ubiquitin-positive nuclear neuronal and astrocytic inclusions (Greco et al., 2006). In contrast to fragile X syndrome, where expanded full mutation alleles (>200 CGG repeats) result in hypermethylation and silencing of the *FMR1* gene, FXTAS carriers of pre-mutation expanded alleles (55–200 CGG repeats) present increased levels of *FMR1* mRNA but slightly reduced expression of the protein encoded by *FMR1*, FMRP (Tassone et al., 2000, 2007; Kenneson et al., 2001).

Because FXTAS is not observed in fragile X patients with fully silenced *FMR1* alleles, a pathogenic mechanism based on expression of mutant *FMR1* mRNAs containing expanded CGG repeats has been proposed (Hagerman and Hagerman, 2004). In support of this hypothesis, multiple studies have demonstrated adverse consequences of expressing RNA containing expanded CGG repeats in cell, fly, and mouse models (Willemsen et al., 2003; Jin et al., 2003; Arocena et al., 2005; Hashem et al., 2009; Entezam et al., 2007; Hukema et al., 2014, 2015). However, how *FMR1* mRNA containing expanded CGG repeats is pathogenic is unclear. A first proposed model is that FXTAS results from a toxic RNA gain-of-function mechanism in which mutant RNA containing expanded CGG repeats sequesters specific RNA binding proteins, resulting in neuronal cell dysfunctions (Iwahashi et al., 2006; Sofola et al., 2007; Jin et al., 2007; Sellier et al., 2010, 2013). A second proposed mechanism is that FXTAS is caused by repeat-associated non-AUG (RAN) translation of the expanded CGG repeats into polyalanine- and polyglycine-containing proteins named FMRpolyA and FMRpolyG (Todd et al., 2013). Ranum and colleagues originally demonstrated that expanded CAG repeats can be translated in all three frames in the absence of any AUG initiation codon (Zu et al., 2011). Subsequently, RAN translation was described and proposed as a causative mechanism in various inherited microsatellite neurodegenerative disorders (reviewed in Cleary and Ranum, 2014), including FXTAS (Todd et al., 2013; Kearse et al., 2016). However, the mechanism and the pathological consequences of translating expanded CGG repeats are not fully understood. Specifically, which pathological mechanism (RNA gain-of-function or RAN translation) drives FXTAS pathogenesis is a crucial question.

Here we find that translation of expanded CGG repeats occurs predominantly in the glycine frame through initiation at a near-cognate ACG codon located upstream of the expanded CGG repeats. Importantly, transgenic mice expressing both CGG RNA repeats and the polyglycine protein (FMRpolyG mouse), but not mice expressing only the mutant RNA containing expanded CGG repeats (CGG RNA mouse), exhibit inclusion formation, motor phenotypes, and reduced lifespan. FMRpolyG interacts with LAP2β, a protein essential to anchoring lamin proteins to the inner nuclear membrane, and overexpression of LAP2β rescues neuronal cell death induced by expression of FMRpolyG. Overall, these results suggest that translation of expanded CGG repeats into FMRpolyG, which alters nuclear lamina, may contribute to FXTAS.

RESULTS

Translation of Expanded CGG Repeats Initiates at an Upstream Near-Cognate Codon

To confirm previous observations of translation of expanded CGG repeats in the absence of an AUG start codon (Todd et al., 2013; Kearse et al., 2016), we cloned 100 CGG repeats embedded within the natural human 5′ UTR of *FMR1*, which was fused to the GFP deleted of its ATG and in all three possible frames (Figure S1A). These frames were named according to the polypeptide potentially encoded by the expanded CGG repeats—namely, glycine, alanine, and arginine. Cell transfection

and immunoblotting against GFP confirm previous data (Todd et al., 2013; Kearse et al., 2016) and demonstrate that the 5′ UTR of *FMR1* with expanded CGG repeats allows translation of a GFP protein with an ~12 kDa N-terminal extension corresponding to the expanded CGG repeats translated into the glycine frame (Figure 1A and Figure S1B). In contrast, we observed no translation when the *FMR1* sequence located upstream of the repeats was deleted (Figure 1A and Figure S1B). Serial deletions confirmed that a short *FMR1* sequence located upstream of the CGG repeats was required for FMRpolyG translation (Figure 1B). Treatment with Lysostaphin, a polyglycine endopeptidase, confirmed that expanded CGG repeats are translated into a polyglycine-containing protein (Figure S1C).

To further characterize the initiation site of expanded CGG repeat translation, we immunoprecipitated FMRpolyG and determined its N-terminal sequence by proteomic analysis after trypsin digestion. LC-MS/MS spectra revealed initiation to an ACG near-cognate codon located 32 nucleotides upstream of the CGG expansion (Figure 1C). To exclude any bias of digestion, we repeated this experiment with a construct containing a mutation in which a lysine (sensitive to the LysC enzyme) is present upstream of the expanded CGG repeats, but the remaining sequence, notably the ACG codon, is untouched. Proteomic analysis of FMRpolyG after LysC digestion confirmed that FMRpolyG translation initiates at the ACG near-cognate codon (Figure S1D). Of interest, proteomic analysis also revealed that the initial amino acid of FMRpolyG is a methionine, suggesting that the ACG codon is decoded by an initiator Met-tRNA despite imperfect match (Figures 1C and S1D). This ACG near-cognate codon is embedded in a putative Kozac consensus sequence that is conserved among multiple species (Figure S1E). Translation initiation at the ACG near-cognate codon is predicted to result in a small FMRpolyG protein composed of a short, 12-amino-acid N terminus, a central glycine stretch with a length that corresponds to the number of expanded CGG repeats, and a C terminus of 42 amino acids with no predicted structure or homology (Figure 1D). Thus, according to the number of CGG repeats, FMRpolyG may range from 7 kDa in control individuals with 30 CGG repeats to ~15 kDa in carriers of pre-mutation with 150 CGG repeats.

A Minimum of 60 to 70 Expanded CGG Repeats Is Required to Detect FMRpolyG

Upstream ORFs (uORFs) are short open reading frames that are located in the 5′ UTR of mRNAs upstream of the main ORF. Due to the 5′ to 3′ scanning ribosome process, and in absence of ribosome re-initiation, the translation of a uORF generally impairs translation of the downstream main ORF. To avoid complete translation inhibition, uORFs typically start by an AUG or a near-cognate codon (GUG, CUG, UUG, ACG, etc.) embedded in a poor Kozac consensus sequence that enables leaky ribosomal scanning and hence translation initiation at the downstream main ORF (reviewed in Sonenberg and Hinnebusch, 2009). To test the presence of a putative uORF encompassing the CGG repeats in *FMR1*, we cloned various lengths of CGG expansion within the human 5′ UTR of *FMR1* fused in the glycine frame with a small FLAG tag (eight amino acids, ~1 kDa). Immunoblotting revealed that expression of FMRpolyG-FLAG is
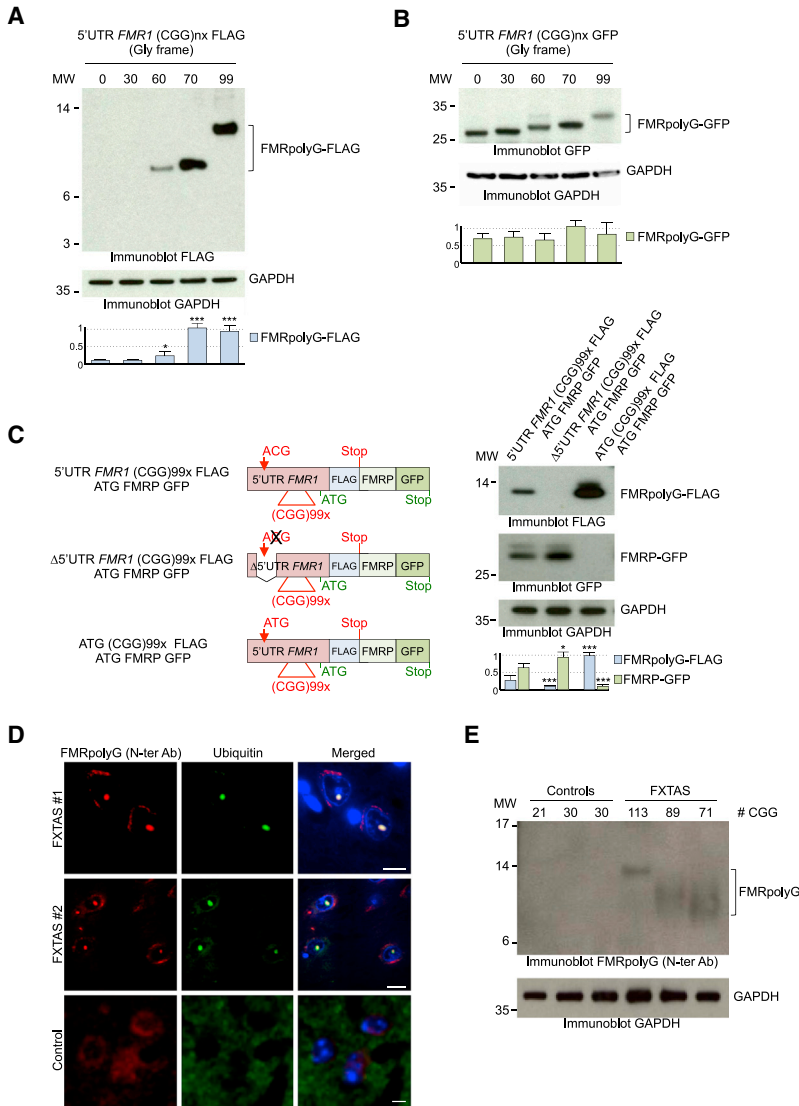
Figure 1. Translation of CGG Repeats Initiates at an Upstream Near-Cognate Codon

(A) Immunoblotting against GFP on the soluble lysate fraction of HeLa cells transfected for 24 hr with expanded CGG repeats embedded, or not, in the 5′ UTR of *FMR1* and fused in all three possible frames with the GFP deleted of its ATG.

(B) Upper, schemes of the *FMR1* 5′ UTR deletion constructs tested. Middle, immunoblotting against GFP on the soluble lysate fraction of HeLa cells transfected for 24 hr with mutants of *FMR1* 5′ UTR containing expanded CGG repeats fused to the GFP in the glycine frame. Lower, quantification of FMRpolyG-GFP expression reported to GAPDH.

(C) LC-MS/MS spectra of the N-terminal part of the immunoprecipitated and trypsin-digested protein translated from expanded CGG embedded in the 5′ UTR of *FMR1*.

(D) Scheme and partial nucleotide sequence of human *FMR1* exons 1 and 2. Amino acid sequence of *FMR1* uORF translated into FMRpolyG is indicated in red. Amino acid sequence of the beginning of the FMRP ORF is indicated in green.

Error bars indicate SEM of three independent transfections. Student's t test, *** indicates p < 0.001.

A 5'UTR *FMR1* (CGG)nx FLAG (Gly frame)

B 5'UTR *FMR1* (CGG)nx GFP (Gly frame)

C 5'UTR *FMR1* (CGG)99x FLAG ATG FMRP GFP

Δ5'UTR *FMR1* (CGG)99x FLAG ATG FMRP GFP

ATG (CGG)99x FLAG ATG FMRP GFP

D FMRpolyG (N-ter Ab)   Ubiquitin   Merged

E

detected only with expanded CGG repeats over 60 to 70 CGG repeats (Figure 2A), the threshold over which pre-mutation carriers are at risk of developing FXTAS. Fusion of expanded CGG repeats in the glycine frame to the GFP (25 kDa) confirmed that translation occurred with expanded CGG repeats of various lengths (70 to 100 repeats), characteristic of pre-mutation carriers (Figure 2B). However, FMRpolyG-GFP was also translated with short stretches of CGG repeats (30) found in control individuals or even without any CGG repeats (Figure 2B). These results suggest that translation initiation occurs at the ACG near-cognate codon independently of the CGG repeats but that expansion over 70 repeats is important for detection of FMRpolyG. This is characteristic of short upstream ORFs that are generally translated into small and most often undetectable peptides but which are detected when fused with large tags, resulting in stable and detectable proteins (Aspden et al., 2014).

Next, we noted that FMRpolyG and FMRP ORFs are in different frames, with the last 20 amino acids of FMRpolyG naturally overlapping the FMRP N terminus (Figure 1D). Thus, translation of the FMRpolyG uORF may potentially impair ribosomal re-initiation to the downstream FMRP ORF. To test that hypothesis, we fused the 5′ UTR of *FMR1* to a FLAG tag in the glycine frame and also fused the downstream FMRP ORF in frame with the GFP (Figure 2C). This construct expresses the FLAG-tagged FMRpolyG uORF or the downstream GFP-tagged FMRP main ORF (Figure 2C), suggesting that translation initiation occurs either at the FMRpolyG ACG near-cognate codon or at the FMRP ATG codon. Deletion of the 5′ UTR sequence containing the ACG near-cognate initiation codon abolished expression of the FLAG-tagged FMRpolyG uORF but enhanced translation of the downstream GFP-tagged FMRP ORF (Figure 2C). In contrast, mutation of the ACG near-cognate codon into a cognate AUG initiation codon predictably enhanced translation of the FLAG-tagged FMRpolyG uORF, but abolished expression of the downstream, GFP-tagged FMRP ORF (Figure 2C).

To confirm that FMRpolyG is encoded by an ORF initiating upstream of the CGG repeats, we developed an antibody directed against the 12 amino acids located upstream of the glycine repeats (Figure S2A). Immunofluorescence revealed the presence of FMRpolyG N terminus in brain sections of FXTAS patients, but not in age-matched control individuals (Figure 2D). FMRpolyG
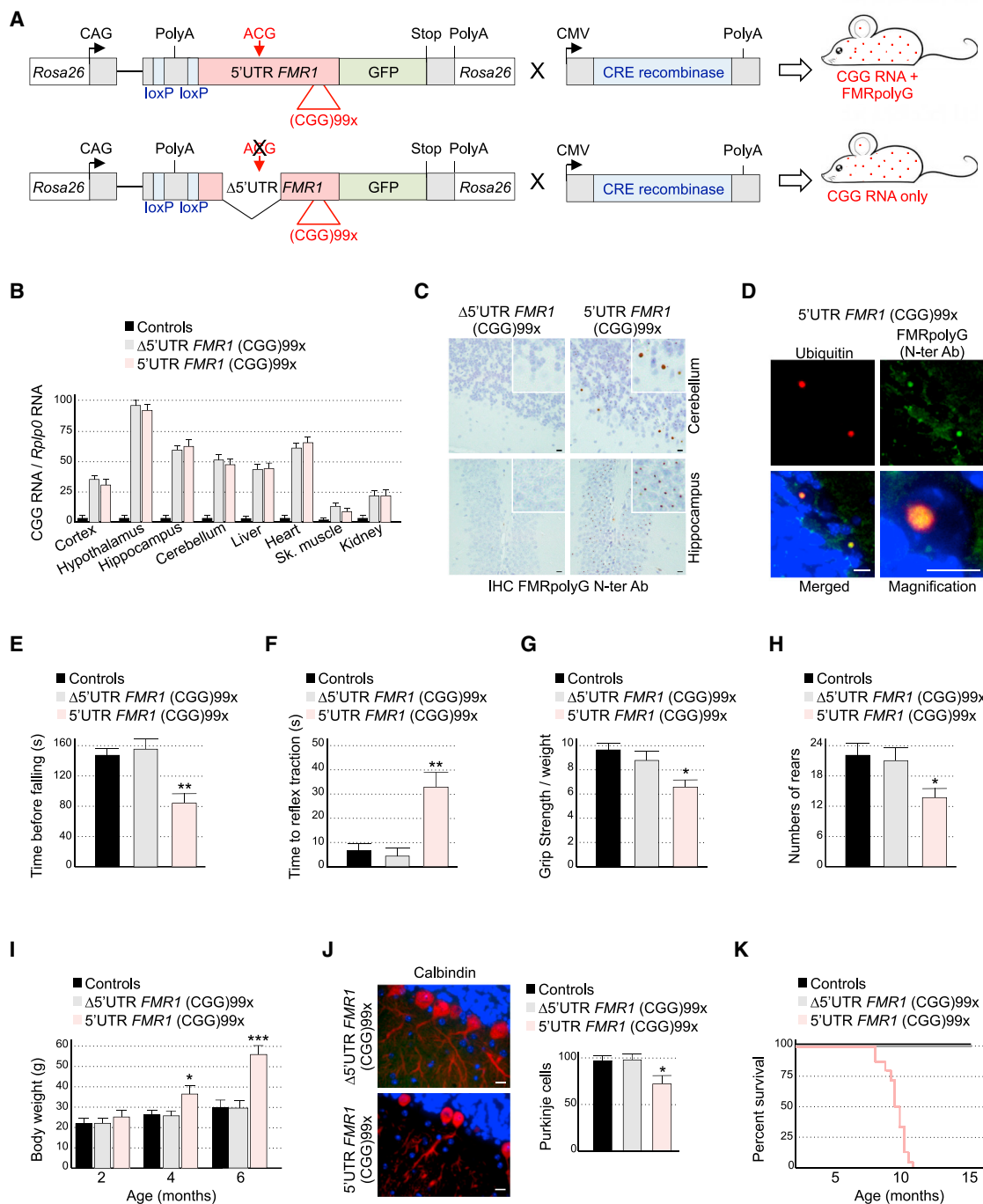
Figure 3. Expression of FMRpolyG Is Pathogenic in Mice

(A) Schemes of the mouse transgene constructs.

(B) Quantitative RT-PCR analysis of transgene expression relative to the *RplpO* mRNA in different brain areas and tissues of 6-month-old control (n = 3), bigenic CMV-cre/full-length (n = 3), or mutant (n = 3) *FMR1* 5′ UTR transgenic mice.

(C) Immunohistochemistry against FMRpolyG N terminus of cerebellum and hippocampus areas of 6-month-old bigenic CMV-cre/full-length or mutant *FMR1* 5′ UTR transgenic mice. Scale bars, 10 μm. Sections were counter-stained with Nissl staining.

(D) Immunofluorescence against FMRpolyG N terminus and ubiquitin on cerebellum areas of 6-month-old bigenic CMV-cre/full-length or mutant *FMR1* 5′ UTR transgenic mice. Scale bars, 10 μm. Nuclei were counter-stained with DAPI.

(E) Rotarod test: the time before falling from a rotating rod of 3-month-old control (n = 8), bigenic CMV-cre/full-length (n = 9), or mutant (n = 9) *FMR1* 5′ UTR transgenic male mice.

(legend continued on next page)

N terminus is detected as single nuclear inclusions that co-localize with ubiquitin, which is consistent with the known histopathological features of FXTAS (Greco et al., 2006). We confirmed these results by immunoblotting and found that, according to the size of the CGG expansion, FMRpolyG was detected as a 10- to 14-kDa protein in the insoluble fraction of brain lysate of FXTAS individuals (Figure 2E). The poor quality of the immunoblotting is probably due to the propensity of this glycine-rich protein to aggregate. To ensure that translation successfully passes across the expanded CGG repeats, we also developed an antibody against the amino acids located downstream of the glycine repeats (Figure S2A). Consistent with previous analyses (Todd et al., 2013; Buijsen et al., 2014), immunofluorescence confirmed presence of FMRpolyG C terminus in brain sections of FXTAS patients, but not in age-matched control individuals (Figure S2B). Finally, translation of expanded CGG repeats into a polyalanine-containing protein was reported previously (Todd et al., 2013; Kearse et al., 2016). We developed an antibody against the putative C-terminal part of this protein (Figure S2C), but observed no or very little FMRpolyA in individuals with FXTAS (Figure S2D).

Translation of Expanded CGG Repeats into FMRpolyG Is Pathogenic in Mice

The presence of both CGG RNA aggregates, which can titrate out RNA binding proteins, and FMRpolyG–positive inclusions in individuals with FXTAS raises the question of which pathogenic mechanism is driving neuronal degeneration in this disease. To differentiate between these two hypotheses, we developed two transgenic mouse models. The first one contains the full human 5′ UTR of FMR1 with expanded 99 CGG repeats that express both CGG RNA and FMRpolyG protein, while the second mouse model also expresses 99 CGG repeats, but the non-canonical ACG initiation codon and surrounding 5′ UTR sequence are deleted so it only expresses the CGG RNA (Figure 3A). Both constructs are driven by a chimeric CAG promoter inserted within the neutral Rosa26 locus and under repression of a loxP-polyadeneylation cassette (Figure 3A).

Deletion of the loxP cassette using a ubiquitously and embryonically expressed Cre recombinase led to high expression of transgene RNA throughout the brain, heart, and liver, with lower expression in skeletal muscle, kidneys, and other organs (Figure 3B). Importantly, transgene RNA expression was similar in bigenic CMV-cre mice with either the full or mutant FMR1 5′ UTR (Figure 3B). However, we found no or only very rare CGG RNA foci in brain sections of full or mutant FMR1 5′ UTR transgenic

mice (Figures S3A and S3B). This is consistent with the rare occurrence of CGG RNA aggregates in other transgenic mouse models expressing expanded CGG repeats (Sellier et al., 2010). Concerning the FMRpolyG protein, immunohistochemistry assays demonstrated expression and accumulation of nuclear aggregates of FMRpolyG in brain sections from the full 5′ UTR FMR1 transgenic mice, but not in the mutant 5′ UTR mice (Figures 3C and S3C). These results confirm in mouse models that translation of the CGG repeats in the glycine frame requires the presence of an upstream FMR1 sequence. As observed in brain samples of individuals with FXTAS, aggregates of FMRpolyG in mice co-localized with ubiquitinated inclusions (Figure 3D). Nuclear aggregates of FMRpolyG accumulated over time, with the largest burden of inclusions occurring within the hypothalamus, mirroring transgene RNA expression (Figures S3C and S3D). Some rare aggregates of FMRpolyG were found in other tissues than brain, which is reminiscent of reported FMRpolyG aggregates in non-CNS tissues in FXTAS patients (Buijsen et al., 2014). In contrast, we did not observe aggregates of FMRpolyA in the full or mutant 5′ UTR mice (Figure S3E).

To determine the consequences of FMRpolyG production in mice, we conducted a series of behavioral and locomotor assays on both mouse lines. Mice with the full 5′ UTR of FMR1 develop obesity at 6 months of age. Therefore, behavioral tests were performed at 3 months of age when weight is identical between full and mutant FMR1 5′ UTR transgenic mice. Importantly, we observed that only mice with both the full 5′ UTR sequence of FMR1 and expression of the FMRpolyG protein present locomotor deficiency (Movies S1 and S2), with increased falling from rotarod (Figure 3E), decreased ability of traction from the hind limbs (Figure 3F), decreased grip strength (Figure 3G), and a decreased number of rears in open field observation (Figure 3H). At 6 months of age, mice with the full 5′ UTR of FMR1 lose mobility and develop obesity, while mice with the mutant 5′ UTR that express only the CGG RNA remain normal (Figure 3I). We did not observe massive neuronal cell death, but found some loss of Purkinje cells in mice with the full 5′ UTR of FMR1 compared to mutant 5′ UTR mice (Figures 3J and S3F). Furthermore, Iba1 and Gfap staining were mildly increased in brain sections of mice with the full 5′ UTR of FMR1 compared to mutant 5′ UTR mice, suggesting some increased neuroinflammation in FMRpolyG-expressing animals (Figure S3G). Finally, expression of FMRpolyG is deleterious, as mice expressing the full 5′ UTR of FMR1 die at around 10 months, while mutant FMR1 5′ UTR mice exhibit normal longevity and are indistinguishable from control mice (Figure 3K).

(F) String test: the time to gain hindlimb traction for forelimb-hanging 3-month-old control (n = 8), bigenic CMV-cre/full-length (n = 9), or mutant (n = 9) FMR1 5′ UTR transgenic male mice.
(G) Grip test: the maximal force relative to mouse body weight exerted to release 3-month-old control (n = 8), bigenic CMV-cre/full-length (n = 9), or mutant (n = 9) FMR1 5′ UTR transgenic male mice holding a grid with their forepaws.
(H) Open field: number of rears during 5 min observation in open field of 3-month-old control (n = 8), bigenic CMV-cre/full-length (n = 9), or mutant (n = 9) FMR1 5′ UTR transgenic male mice.
(I) Body weight of 2-, 4-, and 6-month-old control (n = 6), bigenic CMV-cre/full-length (n = 6), or mutant (n = 6) FMR1 5′ UTR transgenic male mice.
(J) Left, immunofluorescence labeling of calbindin of cerebellum sections of 9-month-old bigenic CMV-cre/full-length or mutant FMR1 5′ UTR transgenic mice. Scale bars, 10 μm. Nuclei were counter-stained with DAPI. Right, quantification of Purkinje cells (n = 100) in cerebellum sections of 9-month-old bigenic CMV-cre/full-length (n = 3) or mutant FMR1 5′ UTR (n = 3) transgenic mice.
(K) Kaplan-Meier survival curve of control (n = 15), bigenic CMV-cre/full-length (n = 15), or mutant (n = 15) FMR1 5′ UTR male and female transgenic mice.
Error bars indicate SEM. Student's t test, * indicates p < 0.05, ** indicates p < 0.01 and *** indicates p < 0.001.
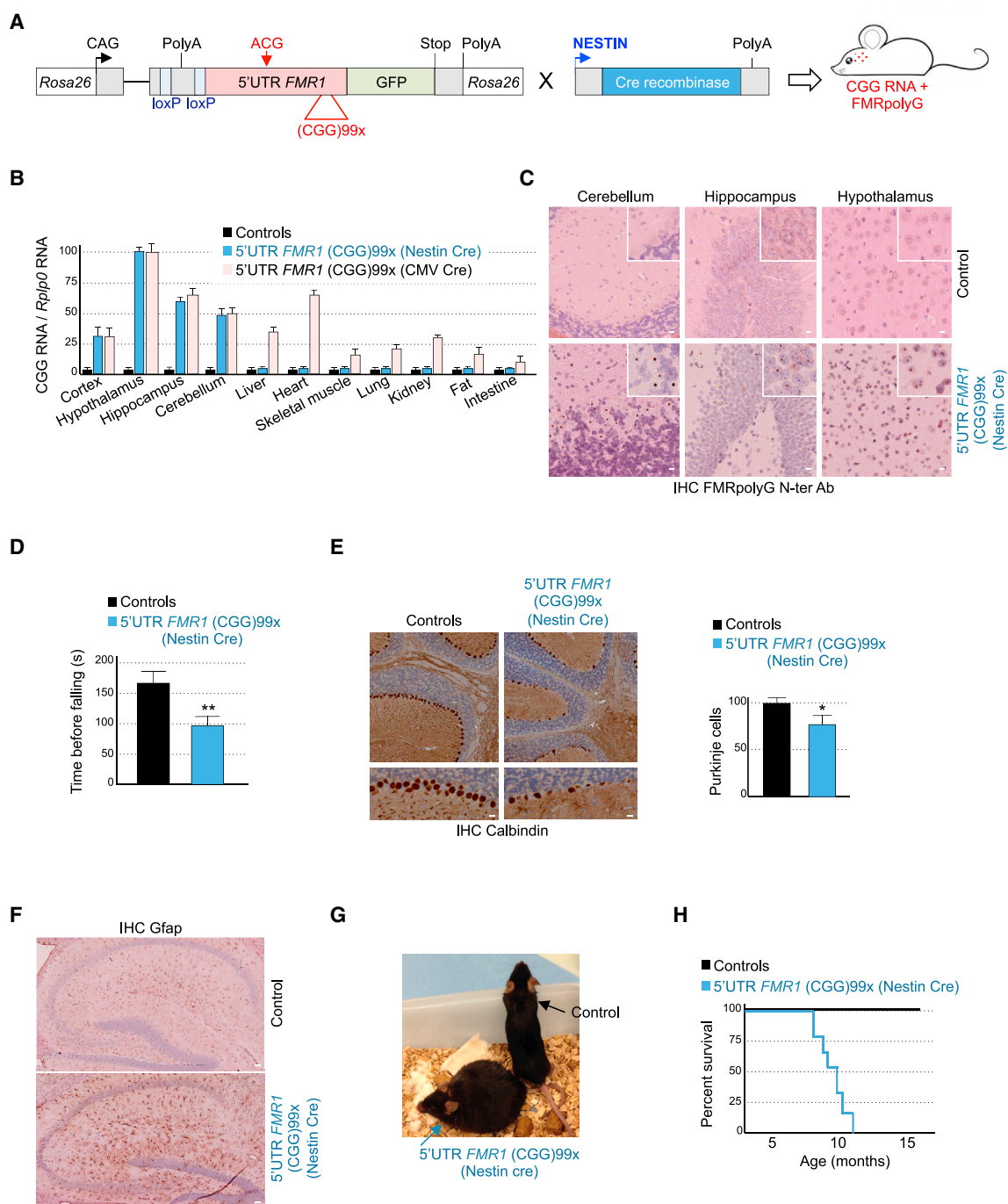
Figure 4. Neuronal Expression of FMRpolyG Is Pathogenic in Mice

(A) Schemes of the mouse transgene constructs.

(B) Quantitative RT-PCR analysis of transgene expression relative to the *RplpO* mRNA in different tissues of 6-month-old control (n = 3), bigenic CMV-cre/full-length *FMR1* 5′ UTR (n = 3), or bigenic Nestin-cre full-length *FMR1* 5′ UTR (n = 3) mice.

(C) Immunohistochemistry against FMRpolyG N terminus in the cerebellum, hippocampus, and hypothalamus of 6-month-old bigenic Nestin-cre/full-length *FMR1* 5′ UTR mice. Scale bars, 10 μm. Sections were counter-stained with H&E staining.

(D) Rotarod test: time before falling of 3-month-old control (n = 6) or bigenic Nestin-cre/full-length *FMR1* 5′ UTR (n = 6) male mice.

(E) Left, immunohistochemistry against calbindin of cerebellum sections of 10-month-old control or bigenic Nestin-cre/full-length *FMR1* 5′ UTR mice. Scale bars, 10 μm. Sections were counter-stained with H&E staining. Right, quantification of Purkinje cells (n = 50) in cerebellum sections of 10-month-old control (n = 3) or bigenic Nestin-cre/full-length *FMR1* 5′ UTR (n = 3) mice.

(F) Immunohistochemistry against Gfap of hippocampal sections of 10-month-old control or bigenic Nestin-cre/full-length *FMR1* 5′ UTR mice.

(legend continued on next page)

To assess the tissue origin of these phenotypes, we next analyzed offspring of full 5′ UTR transgenic mice crossed with Nestin-cre mice, which express the Cre recombinase in precursors of neurons and glia cells around E10.5 (Figure 4A). Quantitative RT-PCR revealed high expression of expanded CGG RNA only in mouse brain of the double transgenic Nestin-cre/full 5′ UTR *FMR1* mice, confirming the specificity of the Nestin promoter (Figure 4B). Immunohistochemistry indicated expression and accumulation of nuclear aggregates of FMRpolyG in brain sections of these bigenic mice, but not in control, non-transgenic animals (Figures 4C and S4A). Importantly, locomotor testing at 3 months of age revealed increased falling from rotarod compared to control, non-transgenic mice (Figure 4D). Histopathological analyzes revealed some Purkinje cell loss (Figure 4E) and evidence of neuroinflammation (Figure 4F) in 9- to 10-month-old Nestin-cre/full 5′ UTR *FMR1* mice compared to control animals. Interestingly, mice expressing FMRpolyG only in the brain developed obesity (Figure 4G) and exhibited reduced longevity (Figure 4H), similar to the CMV-cre/full 5′ UTR *FMR1* mice. The obesity developed by these mice may originate from dysfunction of the hypothalamus, as this area is the one expressing the most and the earliest FMRpolyG. In support of that hypothesis, RT-qPCR indicated that expression of the leptin receptor, which is essential to regulate energy balance by inhibiting hunger, was decreased in the hypothalamic area of these mice (Figure S4B).

The 5′ UTR Sequence of *FMR1* Impairs Formation of Expanded CGG-Repeat RNA Foci

The absence of RNA foci of expanded CGG repeats in full and mutant 5′ UTR *FMR1* mice is puzzling, as FXTAS is proposed to result from nuclear accumulation of mutant CGG RNA that would sequester various RNA binding proteins. Thus, we tested whether transfection of expanded CGG repeats embedded in their natural *FMR1* 5′ UTR sequence would form any RNA foci in mouse neuronal cell cultures. As positive control, we used a construct expressing expanded CGG repeats without the surrounding *FMR1* sequence that was previously shown to form RNA foci in transfected cells (Sellier et al., 2010, 2013; Figure S4C). RNA FISH assays indicated that expanded CGG repeats embedded in the 5′ UTR of *FMR1* formed rare RNA foci compared to the expanded CGG repeats without the surrounding *FMR1* sequence (Figure S4D). Furthermore, RT-PCR performed on nuclear and cytoplasmic fractions indicated that most of the RNAs containing expanded CGG repeats embedded within the 5′ UTR of *FMR1* are exported from the nucleus to the cytoplasm (Figure S4E). In contrast, expanded CGG-repeat RNAs without the *FMR1* sequence are largely retained within cell nuclei (Figure S4E). Both constructs present similar plasmid backbones with strong polyadenylation signals (Figure S4C). These results highlight the nuclear retention bias induced by using artificial constructs in which microsatellite repeats are separated from their natural sequence context. Importantly, these data also indicate that CGG expanded repeats embedded in

their natural *FMR1* sequence are exported into the cytoplasm and thus available for translation into FMRpolyG.

Expression of FMRpolyG Is Toxic for Neuronal Cells

Next, we investigated by which mechanisms the FMRpolyG protein may elicit neuronal cell dysfunction. Immunofluorescence assays indicated that FMRpolyG first accumulates in the cytoplasm, where it forms aggregates. These aggregates then migrate and form an inclusion within the cell nucleus in primary mouse-embryonic neuronal cultures (Figure 5A). Similar results were observed in HEK293 cells and when a smaller FLAG epitope tag was used instead of GFP (Figures S5A and S5B). Furthermore, immunoblotting indicated that FMRpolyG progressively accumulates in the insoluble fraction of transfected cell lysates, which is consistent with its propensity to form aggregates (Figure S5C).

To identify the sequence driving FMRpolyG aggregation, we cloned various deletion mutants of FMRpolyG expressing either its N terminus with the glycine repeats or its C terminus in isolation. To obtain comparable expression, all mutant constructs were driven by an ATG (Figure S5D). Expression of the full-length FMRpolyG protein in primary cultures of E18 mouse cortical neurons leads to nuclear aggregates associated with cell death (Figure 5B). Expression of the polyglycine stretch in isolation, with the C terminus of FMRpolyG deleted, was sufficient to elicit aggregation but was not overtly toxic. In contrast, expression of GFP fused to the 42 amino acids constituting the C terminus of FMRpolyG caused neuronal cell death without forming nuclear aggregates (Figure 5B; Movies S3, S4, S5, and S6). Similar results were observed in Neuro2A cells (Figure S5E) and when the GFP tag was replaced by the smaller FLAG tag (Figure S5F).
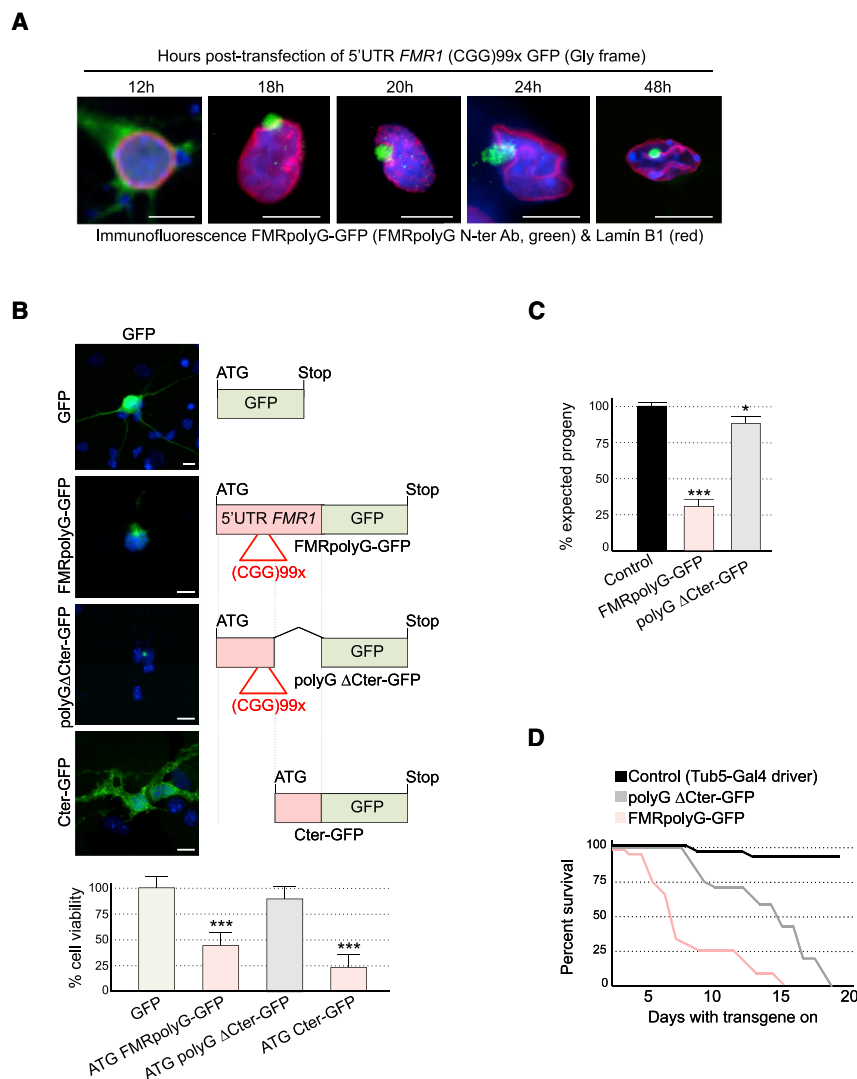
To confirm these results in an animal model and test the toxicity of FMRpolyG on a longer time period, we developed *Drosophila* transgenic lines expressing either the full FMRpolyG protein or the polyglycine stretch in isolation under a *UAS* promoter. Toxicity was assessed by two separate assays. First, *UAS* FMRpolyG-GFP and *UAS* polyG ΔCter-GFP flies, both expressed under an ATG initiator codon, were crossed with an *Act5c-Gal4* driver line, which leads to ubiquitous expression of the transgene during development. Total progeny carrying either the transgenes or a balancer chromosome were then quantified over three independent crosses. Importantly, expression of the full FMRpolyG was toxic and reduced progeny eclosion by half, while the eclosion rate was only slightly reduced in flies expressing the polyglycine stretch in isolation (Figure 5C). To exclude any potential bias due to random insertion effects, we generated and analyzed various independent lines expressing either FMRpolyG or its polyglycine stretch. As control, quantitative RT-PCR showed similar expression of the FMRpolyG or polyglycine transgenes (Figure S5G). Importantly, all lines expressing FMRpolyG-GFP showed reduced viability compared to GFP controls. In contrast, the lines expressing the glycine repeats in isolation consistently exhibited a milder phenotype (Figure S5H).

Scale bars, 50 μm. Sections were counter-stained with H&E staining.

(G) Representative image of 8-month-old control or bigenic Nestin-cre/full-length *FMR1* 5′ UTR mice.

(H) Kaplan-Meier survival curves of control (n = 10) or bigenic Nestin-cre/full-length *FMR1* 5′ UTR (n = 10) male and female mice.

Error bars indicate SEM. Student's t test, * indicates p < 0.05, ** indicates p < 0.01.

A

Hours post-transfection of 5'UTR *FMR1* (CGG)99x GFP (Gly frame)

| 12h | 18h | 20h | 24h | 48h |

Immunofluorescence FMRpolyG-GFP (FMRpolyG N-ter Ab, green) & Lamin B1 (red)

B

C

D

Figure 5. FMRpolyG Toxicity Is Influenced by Its Carboxyl Terminus

(A) Immunofluorescence against FMRpolyG N terminus and Lmnb1 in primary cultures of E18 mouse cortical neurons transfected for the indicated time period with expanded CGG repeats embedded within the 5′ UTR of *FMR1* and fused to the GFP in the glycine frame. Scale bars, 10 μm. Nuclei were counter-stained with DAPI.

(B) Left, representative images of primary cultures of E18 mouse cortical neurons transfected with GFP or ATG-driven FMRpolyG-GFP full-length or mutants. Scale bars, 10 μm. Nuclei were counter-stained with DAPI. Right, schemes of the mutant constructs of FMRpolyG-GFP. The Nter construct corresponds to the N-terminal part of FMRpolyG including its polyglycine repeats fused to the GFP. The Cter construct corresponds to the last 42 amino acids of FMRpolyG fused to the GFP. Lower, quantification of neuronal cell viability of GFP-positive (n = 100 cells, three independent transfections) transfected E18 mouse cortical neurons.

(C) Progeny eclosion ratio (n = 100, three independent crosses) of *Drosophila* ubiquitously expressing FMRpolyG, either full-length or deleted of its C terminus compared to control driver line (*Actin5C-Gal4/+*).

(D) Kaplan-Meier survival curve of *Drosophila* expressing FMRpolyG full-length or deleted of its C terminus compared to control driver line (*Tub5-Gal4/+*).
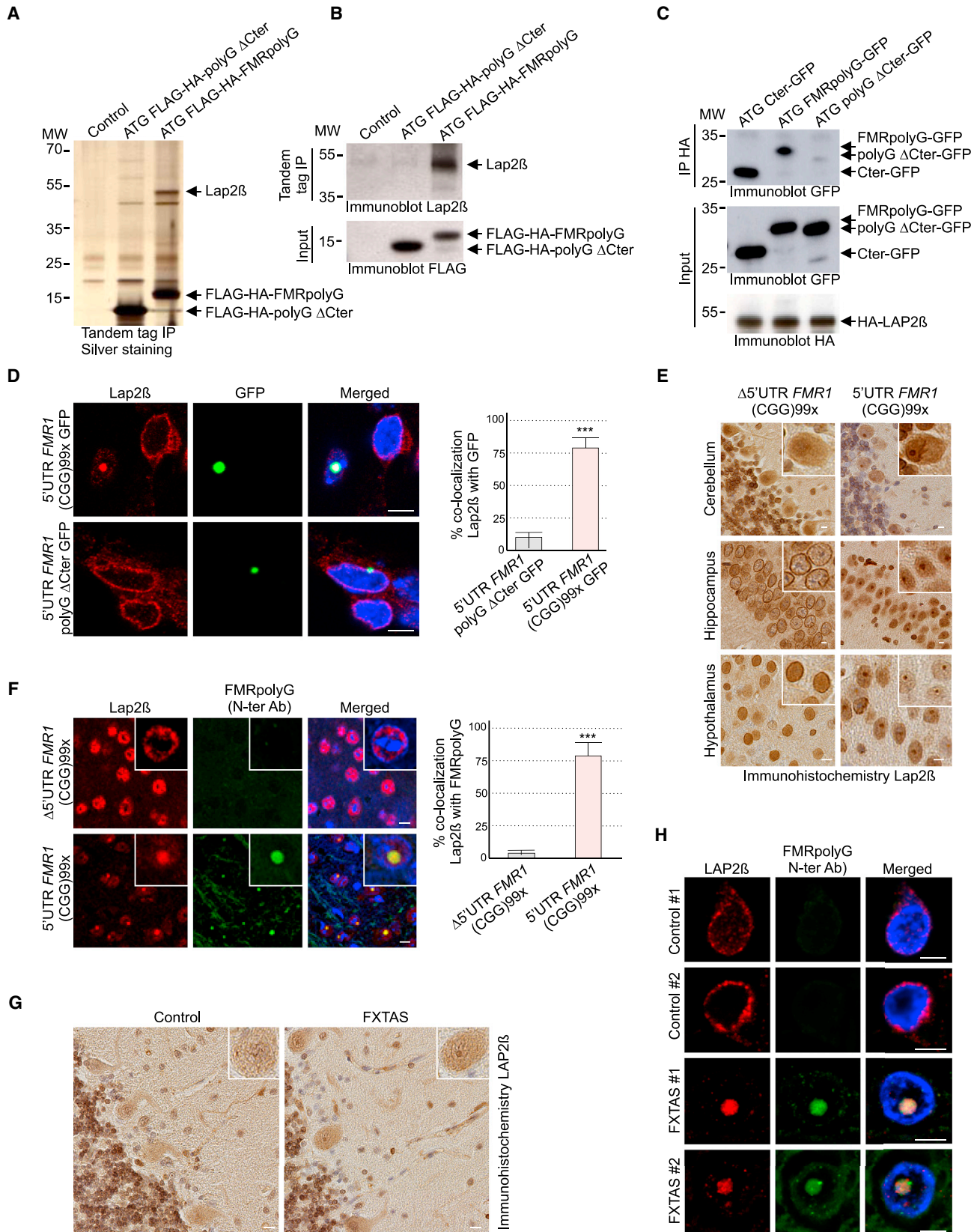
Error bars indicate SEM. Student's t test, * indicates p < 0.05, *** indicates p < 0.001.

FMRpolyG Interacts with LAP2β and Alters the Nuclear Lamina

To identify FMRpolyG binding proteins, we performed a tandem tag purification of HA-FLAG-tagged FMRpolyG transfected into Neuro2A cells followed by nano-LC-MS/MS analysis of associated proteins. This approach identified various FMRpolyG-associated proteins (Table S1), including Lap2β (Figure 6A). Lamina-associated polypeptide 2 (LAP) alpha and beta are two isoforms of the LAP2 protein that differ in their C-termini; these isoforms originate from alternative splicing of the *TMPO* pre-mRNA. LAP2α is diffusely localized in the nucleus, while LAP2β carries a transmembrane domain in its C terminus that anchors it to the inner nuclear membrane (Furukawa et al., 1995). LAP2β interacts with lamin B1 and B2 and helps to organize these proteins near the nuclear inner membrane. Consequently, alteration of LAP2β results in disorganization of the nuclear lamina architecture (Dubińska-Magiera et al., 2016; Gant et al., 1999).

Co-immunoprecipitation studies confirmed the association of LAP2β with HA-FLAG-tagged FMRpolyG, but not with the polyglycine stretch in isolation (Figure 6B). Similar results were observed when a GFP tag was used instead of the double FLAG-HA tag (Figure 6C). An interaction between LAP2β and FMRpolyG raises questions of whether translation of the expanded CGG repeats in FXTAS may alter the localization or

As a second measure of FMRpolyG toxicity, we crossed these same transgenic fly lines to a *Tubulin-Gal4* Geneswitch driver, which expression is induced upon addition of mifepristone (RU-486). As a control, adult transgenic flies reared off of RU-486 drug exhibited no differences in viability from control flies, indicating that insertion of the FMRpolyG or polyglycine transgenes had no deleterious effect. In contrast, transgenic flies fed with RU-486, activating ubiquitous transgene expression, showed a decrease in viability over time for FMRpolyG expressing flies (Figure 5D). Adult *Drosophila* expressing the polyglycine stretch alone also presented a decrease in viability (Figure 5D); however, their survival was significantly prolonged compared to flies expressing the full FMRpolyG sequence. The difference of toxicity of the polyglycine stretch in isolation between cell culture and *Drosophila* is probably due to the longer time frame and in vivo nature of the *Drosophila* studies. Overall, these results demonstrate that expression of FMRpolyG is pathogenic, with its polyglycine stretch driving aggregation and with both polyglycine and its C terminus contributing to toxicity in vivo.

(legend on next page)

the function of LAP2β. Expression of FMRpolyG-GFP in primary cultures of cortical neurons from mouse embryo indicated that FMRpolyG recruits endogenous Lap2β in nuclear aggregates (Figure 6D). Similar results were observed in HEK293 cells and when the GFP was replaced by a smaller FLAG tag (Figure S6A). In contrast, a construct containing only the polyglycine stretch of FMRpolyG did not alter Lap2β localization, suggesting that the C-terminal part of FMRpolyG is important for interaction and delocalization of LAP2β (Figure 6D). Immunohistochemistry assays confirmed that endogenous Lap2β localizes in nuclear aggregates in brain sections of the full 5′ UTR *FMR1* transgenic mice, unlike in mutant 5′ UTR mice (Figure 6E). These aggregates of Lap2β co-localized with FMRpolyG inclusions (Figure 6F). We next tried to confirm these results in brain sections of individuals with FXTAS. However, LAP2β immunohistochemistry was of poor quality on human autopsied material. We nevertheless observed some LAP2β aggregates in cerebellum sections as well as in the hippocampal area of FXTAS individuals (Figures 6G and S6B). Furthermore, immunofluorescence assays indicated that LAP2β co-localized with FMRpolyG inclusions in individuals with FXTAS, but not in age-matched control individuals (Figure 6H).

Alteration of LAP2β results in disorganization of the nuclear lamina (Gant et al., 1999; Dubińska-Magiera et al., 2016). Thus, we next investigated the consequences of FMRpolyG expression on the architecture of the nuclear lamina. Expression of 99 CGG repeats embedded within the 5′ UTR of *FMR1* fused in the glycine frame with either a GFP or a FLAG tag in HEK293 cells resulted in disorganization of the nuclear lamina, as evidenced by alterations of the lamin B1 labeling (Figure S6C). Similarly, expression of FMRpolyG-GFP in primary cultures of mouse embryonic cortical neurons altered lamin B1 nuclear organization (Figure S6D). Furthermore, and as reported previously with lamin A (Iwahashi et al., 2006), the localization of lamin B1 was altered in brain sections of FXTAS compared to controls (Figure S6E).

LAP2β and Nuclear Lamina Are Altered in Neurons Differentiated from FXTAS iPS Cells

Alterations of the nuclear lamina were observed in autopsied samples that may represent an end stage of the disease. To overcome this potential bias, we developed human induced pluripotent stem (iPS) cells derived from fibroblasts from two age-matched controls and three different FXTAS patients with expansion of 84, 90, and 99 CGG repeats. These iPS cells were differentiated into homogeneous populations of telencephalic neurons with no obvious difference of growth or differentiation between FXTAS and controls iPS cells (Figure S7A). As observed previously (Liu et al., 2012), expression of *FMR1* mRNA was increased 2- to 3-fold in differentiated neurons derived from FXTAS iPS cells compared to controls (Figure S7B). This is consistent with the increased levels of *FMR1* mRNA observed in carriers of a CGG pre-mutation (Tassone et al., 2000, 2007; Kenneson et al., 2001).

Immunofluorescence assays detected accumulation of nuclear aggregates of FMRpolyG in FXTAS neurons, but not in control neurons (Figure 7A). FMRpolyG aggregates accumulated over time post-differentiation, with 5% to 10% of neurons exhibiting small FMRpolyG aggregates at 20 days of differentiation, while 20% to 30% of neurons present FMRpolyG nuclear aggregates after 40 days of differentiation (Figure S7C). In contrast, RNA foci of expanded CGG repeats were rare or absent in FXTAS neurons at 40 days post-differentiation (Figure S7C). Importantly, immunofluorescence analysis revealed that endogenous LAP2β loses its normal localization and forms nuclear inclusions that co-localize with FMRpolyG in neurons differentiated from FXTAS iPS cells (Figure 7A). Furthermore, aggregation of LAP2β and FMRpolyG were associated with disorganization of the nuclear lamina structure, as shown by alteration of the lamin B1 labeling in FXTAS neurons (Figure 7B). In contrast, neurons differentiated from iPS cells of control individuals exhibited normal LAP2β and lamin B1 localization (Figures 7A and 7B).

Overexpression of LAP2 Rescues Neuronal Cell Death Caused by FMRpolyG

Finally, LAP2β overexpression was sufficient to rescue the cell death induced by transfection of FMRpolyG-GFP in neuronal cells (Figure 7C). Consistent with LAP2β binding to the C-terminal part of the FMRpolyG protein, expression of LAP2β also rescued cell death caused by expression of the FMRpolyG

Figure 6. FMRpolyG Interacts with LAP2β and Alters Its Nuclear Localization

(A) Silver staining of proteins captured through consecutive anti-FLAG and anti-HA affinity purification steps from N2A cells transfected for 24 hr with ATG-driven FLAG-HA-tagged FMRpolyG, either full-length or deleted of its C terminus.

(B) Immunoblotting against endogenous Lap2β protein of tandem-tag purified proteins from N2A cells transfected for 24 hr with ATG-driven FLAG-HA-tagged FMRpolyG, either full-length or deleted of its C terminus.

(C) Immunoblotting against the HA or GFP tags of HA-tagged immunoprecipitated proteins from the soluble lysate fraction of N2A cells transfected for 24 hr with HA-LAP2β and ATG-driven FMRpolyG-GFP, either full-length or deleted of its N or C terminus.

(D) Left, GFP fluorescence and immunofluorescence against endogenous Lap2β in primary cultures of E18 mouse cortical neurons transfected with expanded CGG repeats embedded within the 5′ UTR of *FMR1* either full-length or deleted of FMRpolyG C terminus and fused to the GFP in the glycine frame. Nuclei were counterstained with DAPI. Right, quantification of co-localization of Lap2β with GFP inclusions in transfected E18 mouse cortical neurons (n = 100 neurons, three independent transfections).

(E) Immunohistochemistry against Lap2β of cerebellum, hippocampal and hypothalamic areas of 9-month-old bigenic CMV-cre/full-length or mutant *FMR1* 5′ UTR transgenic mice. Sections were counter-stained with H&E staining.

(F) Left, immunofluorescence against FMRpolyG N terminus and Lap2β on hippocampal areas of 9-month-old bigenic CMV-cre/full-length or mutant *FMR1* 5′ UTR transgenic mice. Nuclei were counter-stained with DAPI. Right, quantification of co-localization of Lap2β with FMRpolyG in bigenic CMV-cre/full-length or mutant *FMR1* 5′ UTR transgenic mice (n = 50 neurons, three mice).

(G) Immunohistochemistry against LAP2β of cerebellum areas of FXTAS individual or age-matched control. Sections were counter-stained with H&E staining.

(H) Immunofluorescence against FMRpolyG N terminus and LAP2β on brain sections (hippocampal area) of FXTAS patients or age-matched controls. Nuclei were counter-stained with DAPI.

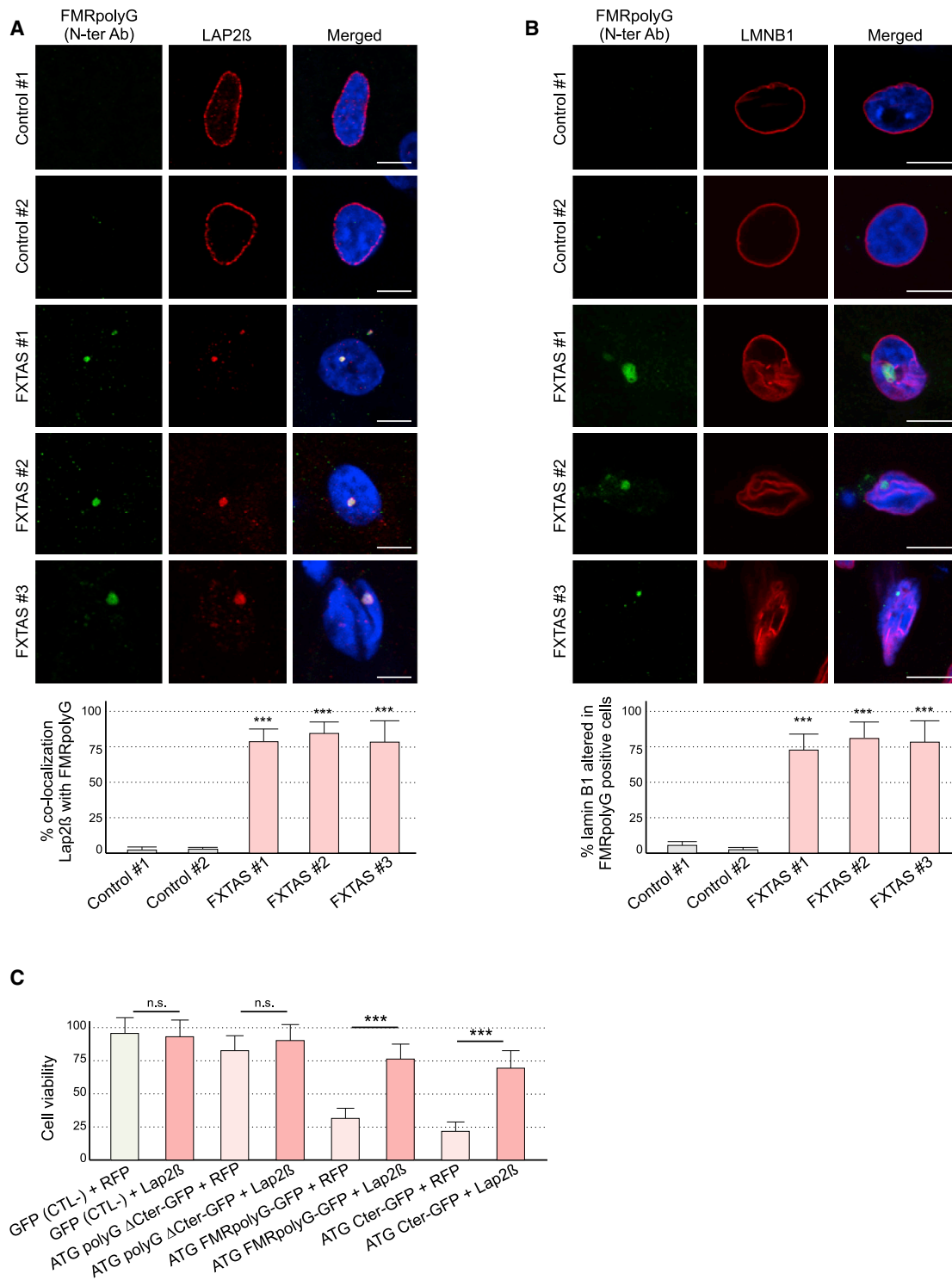Scale bars, 10 μm. Error bars indicate SEM. Student's t test, *** indicates p < 0.001.

Figure 7. LAP2β Rescues Neuronal Cell Death Induced by FMRpolyG
(A) Upper, immunofluorescence against FMRpolyG N terminus and LAP2β on neuronal cultures differentiated 40 days from iPS cells of FXTAs patients or control individuals. Lower, quantification of LAP2β co-localization with FMRpolyG in neurons from iPSC of FXTAS and control individuals (n = 100 neurons, three independent cultures).
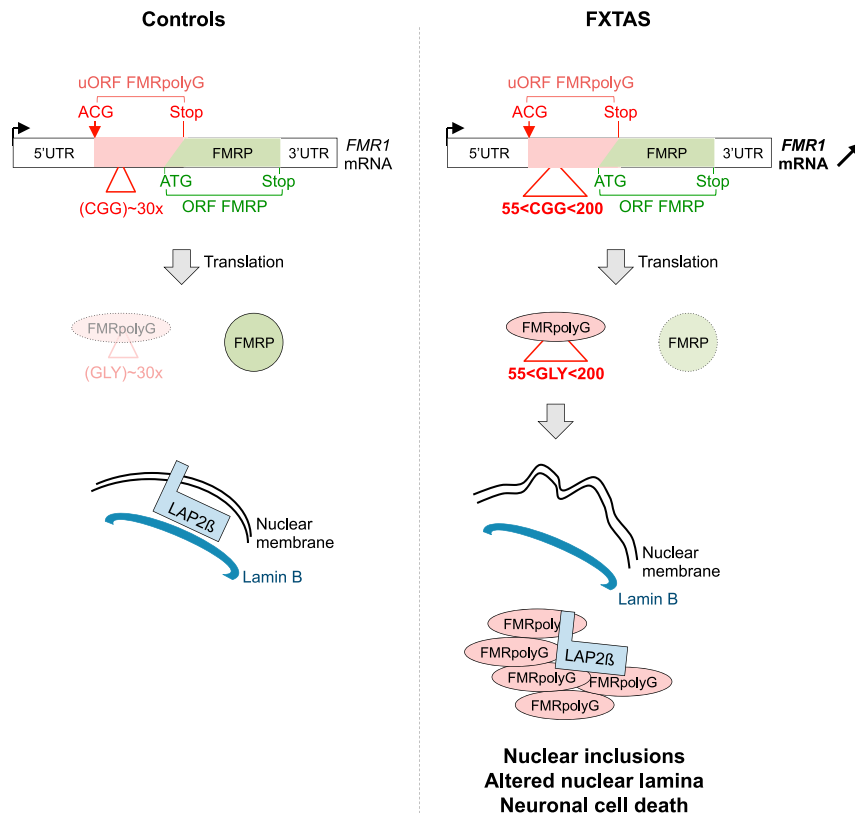
Controls

FXTAS

toxicity. This study confirms in a mammalian system previous results obtained in *Drosophila* (Todd et al., 2013). Moreover, we found that FMRpolyG disrupts nuclear lamina architecture through binding to the nuclear envelope protein LAP2β. These finding are reminiscent of Amyotrophic Lateral Sclerosis and Frontotemporal Dementia (ALS-FTD) in which expanded GGGGCC repeats in the *C9ORF72* gene (Renton et al., 2011; DeJesus-Hernandez et al., 2011) are RAN translated into pathogenic di-peptide-containing proteins (Ash et al., 2013; Mori et al., 2013; Zu et al., 2013) that disrupt nucleocytoplasmic transport (Zhang et al., 2015, 2016; Freibaum et al., 2015; Jovičić et al., 2015).

C terminus in isolation (Figure 7C). As controls, expression of LAP2β had no effect on neuronal survival in control GFP-transfected cells or in neurons expressing the polyglycine fragment in isolation. Similar results were observed when the GFP was replaced by a smaller FLAG tag (Figure S7D). Overall, these data suggest a mechanism by which FMRpolyG can elicit toxicity (Figure 8), and they provide an explanation for previously observed nuclear lamina disorganization in FXTAS patients (Arocena et al., 2005; Iwahashi et al., 2006; Hoem et al., 2011).

DISCUSSION

Previous studies have demonstrated that FXTAS is caused by expression of mutant RNAs containing expanded CGG repeats (Willemsen et al., 2003; Jin et al., 2003; Arocena et al., 2005; Hashem et al., 2009; Entezam et al., 2007; Hukema et al., 2014, 2015). However, whether expanded CGG repeats are pathogenic through an RNA gain-of-function mechanism or through translation into a toxic protein was unclear. Using novel mouse models with or without FMRpolyG expression, our study suggests a direct role for FMRpolyG in CGG-repeat-associated

In contrast to the RAN translation of CAG expanded repeats in the polyalanine frame, which was shown to initiate within the expanded repeats (Zu et al., 2011), our work confirms that expression of FMRpolyG depends largely on initiation at a near-cognate codon located upstream of the CGG repeats (Todd et al., 2013; Kearse et al., 2016). Moreover, we provide mass spectrometry data indicating that the N-terminal amino acid of FMRpolyG is a methionine in mammalian cells, suggesting a canonical mechanism of initiation with altered start-codon fidelity. Our data also suggest that expanded CGG repeats belong to a potential small ORF translated upstream of the main FMRP ORF. However, FMRpolyG is hardly detectable below 70 CGG repeats. This is consistent with the difficulties in detecting small proteins below 10 kDa, but also with the increased expression of mutant CGG RNA in carriers of pre-mutation over 70 CGG repeats (Tassone et al., 2000, 2007; Kenneson et al., 2001) and with the proposed increase in ribosomal stalling at a CGG hairpin structure, which would promote translation initiation to near-cognate codons located upstream of the CGG repeats (Todd et al., 2013; Kearse et al., 2016). Reduced expression of FMRpolyG below 70 CGG repeats may

provide some molecular basis for the threshold of severity and incomplete penetrance observed in FXTAS, as most of pre-mutation alleles contain less than 70 CGG repeats (review in Tassone et al., 2012). We also noted that translation of the FMRpolyG uORF reduces the expression of FMRP; thus, expression of FMRpolyG may contribute to the decrease of FMRP expression observed in FXTAS individuals (Tassone et al., 2000, 2007; Kenneson et al., 2001). However, other mechanisms may also contribute to the decreased expression of FMRP in FXTAS, including the increased ribosome stalling observed with expanded CGG RNA hairpin structure (Feng et al., 1995; Primerano et al., 2002). Ribosome profiling and bioinformatics analyses reveal that upstream ORFs (uORFs) are common in mammalian mRNAs and can initiate at non-canonical codons (Calvo et al., 2009; Ingolia et al., 2011; Fritsch et al., 2012; Ji et al., 2015). Mutations in uORFs are known to cause human disease, mostly by altering the expression of their downstream ORFs (reviewed in Barbosa et al., 2013). In contrast, we propose here that FXTAS is characterized by a mutation extending the length of a uORF and resulting in expression of a toxic, polyglycine-containing protein, FMRpolyG. This model is reminiscent of expansion of tri-nucleotide repeats in ORFs, resulting in expression of pathogenic polyglutamine- or polyalanine-containing proteins (reviewed in Nelson et al., 2013).

Of interest, a pathogenic mechanism in which expansion of nucleotide repeats into a uORF results in expression of a toxic protein may apply to other diseases. For example, aggregates of FMRpolyG were observed in ovarian stromal cells of a woman with Fragile-X-associated primary ovarian insufficiency (FXPOI), caused, like FXTAS, by expanded CGG repeats in the *FMR1* gene (Buijsen et al., 2016). Similarly, it is striking to note that expanded GGGGCC repeats, which are located upstream of the *C9ORF72* ORF and are the main genetic cause of ALS-FTD, are in frame with an upstream, CTG near-cognate codon in a correct Kozac sequence (gctCTGg) to encode a glycine alanine-polypeptide, which is the most common dipeptide-repeat protein detected in individuals with ALS-FTD (Mackenzie et al., 2015). While hypothetical, extending further the FXTAS and fragile X model to *C9ORF72* might predict that long expansions of GGGGCC repeats in tissues prone to somatic expansions would be transcriptionally silent due to epigenetic modification of the promoter, while shorter expansions in tissues prone to repeat contraction would be transcribed and translated.

One of the main conclusions of this work is that mice expressing FMRpolyG develop a locomotor phenotype with reduced longevity, while mice expressing only the expanded CGG repeat RNA are indistinguishable from control mice. These results suggest that accumulation of RNA with expanded CGG repeats or RAN translation initiating within these repeats is not overly pathogenic in mice, at least in the time frame of our study. While informative on the pathogenicity of FMRpolyG, these mouse models present some limitations. Notably, our behavioral and locomotor investigations were limited to the first 3 months of life, because animals expressing FMRpolyG in the brain develop obesity and grow immobile at around 6 months of age. This is in contrast to an individual with FXTAS who do not develop morbid obesity. Furthermore, no FMRpolyG-expressing mice survived past 1 year of age, impairing our ability to observe potential neurode-generation at later ages. As transgene expression is driven by an artificial promoter in these mice, it is likely that their premature death is due to a non-physiological level and/or distribution of transgene expression that may be unrelated to what occurs in FXTAS. Thus, analysis of potential Purkinje cell dropout and ataxia at old age will require specific expression of FMRpolyG in mouse cerebellum.

Direct comparison of these novel mice with previous models of FXTAS expressing either mRNA with 90 CGG repeats and FMRpolyG protein (Willemsen et al., 2003; Hashem et al., 2009; Hukema et al., 2015) or only mRNA with ~118 CGG repeats (Entezam et al., 2007; Qin et al., 2011) is difficult, as these mice either express endogenous levels of *FMR1* mRNA with knockin of the expanded CGG repeats into the endogenous *Fmr1* gene (Willemsen et al., 2003; Entezam et al., 2007) or express expanded repeats in more restricted brain regions than the mouse models reported here (Hashem et al., 2009; Hukema et al., 2015). It is noteworthy that knockin mice with expanded CGG repeats, which do not express FMRpolyG due to presence of a stop codon located upstream of the CGG repeats (Entezam et al., 2007; Todd et al., 2013), exhibit some Purkinje cell dropout at 2 years of age but exhibit no ataxia or rotarod deficiency and only a mild behavioral phenotype with hyperactivity, reduced anxiety, and some subtle deficits in social interaction (Entezam et al., 2007; Qin et al., 2011). These behavioral alterations are similar to those observed in *Fmr1* knockout mouse models; this suggests that they might be caused, in part, by the reduced expression of FMRP observed in this CGG knockin mouse model (Qin et al., 2011). In contrast, the mild Purkinje cell dropout observed at 2 years of age in that mouse model (Entezam et al., 2007) could be caused by expression of the mutant RNA containing expanded CGG repeats and/or RAN translation from these repeats. We failed to detect RNA foci of expanded CGG repeats or nuclear aggregates of FMRpolyA in our mouse models or in neurons from human iPS cells derived from individuals with FXTAS. However, this could originate from technical issues, notably a lack of sufficient sensitivity of our FISH probes of our FMRpolyA antibody. It is also possible that CGG RNA foci and FMRpolyA appear at a later stage of the disease. Thus, whether accumulation of CGG RNA and/or RAN translation of CGG repeats into FMRpolyA might contribute to pathogenicity in FXTAS remains an open question.

Finally, we found that toxicity of FMRpolyG is mediated at least in part through sequestration of the LAP2β protein into nuclear aggregates, leading to disruption of the nuclear lamina architecture and neuronal cell death. These data provide a molecular mechanism for the previously reported nuclear lamina disorganization in FXTAS (Arocena et al., 2005; Iwahashi et al., 2006; Hoem et al., 2011). However, the cell toxicity rescue observed upon overexpression of LAP2β may also originate from high levels of LAP2β blocking the interaction of FMRpolyG with various other proteins. Thus, it is possible that FMRpolyG mediates its toxic effects through more than one mechanism and/or protein partner. Notably, LAP2β regulates gene expression through association with various transcription factors as well as with DNA and with the DNA-binding protein BAF1 (barrier to auto-integration factor 1) protein (Shumaker et al., 2001). It remains to be determined whether FMRpolyG alters this LAP2 transcriptional regulatory activity. Also, proteomic analysis indicates that FMRpolyG pulls

down various other proteins, including cytoskeleton, mitochondrial, and proteasome proteins. Some of these proteins are good candidates to contribute to the mitochondrial alterations that have been observed in FXTAS (Ross-Inta et al., 2010; Kaplan et al., 2012; Hukema et al., 2014). Similarly, FMRpolyG appears critical for elicitation of ubiquitin proteasome impairment when expressed in cells or *Drosophila* (Oh et al., 2015). Further investigations will be needed to test the potential pathological effect of FMRpolyG on mitochondria, protein degradation mechanisms, and cell cytoskeleton in FXTAS.

In conclusion, the *FMR1* gene encodes for at least two different proteins, FMRpolyG and FMRP, whose levels are inversely modulated by the size of the CGG repeat expansion and which are involved in two different genetic diseases, FXTAS and fragile X syndrome.

EXPERIMENTAL PROCEDURES

Human Samples

All brain samples were obtained from the FXTAS brain repository at the UC Davis School of Medicine with the informed consent of individuals and were approved by the Institutional Review Board of the University of California, Davis. Patients have been described previously (Cases LR, 58-02-WD and 1007-05-HP from Greco et al., 2006 and Pretto et al., 2014). Fibroblasts of three FXTAS male individuals with confirmed pre-mutation of 84, 90, and 99 CGG repeats were obtained with the informed consent of individuals and approved by the Institutional Review Board of the Hospital La Pitié Salpêtrière.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, one table, and six videos and can be found with this article online at http://dx.doi.org/10.1016/j.neuron.2016.12.016.

AUTHOR CONTRIBUTIONS

Experiments were performed by C.S., R.A.M.B., F.H., S.N., L.J., P.T., A.G., H.J., H.M., A.V., M.-F.C., M.W.-D., M.-C.B., M.O.-A., P.E., F.R., and M.J. Clinical samples and patient data were obtained by M.A., V.M.-C., F.T., and R.W. Data were collected and analyzed by C.S., F.H., T.S., G.P., M.J., R.W., R.K.H., S.V., C.M., P.K.T., and N.C.-B. The study was designed, coordinated, and written by C.S. and N.C.B. with editorial input from all authors.

ACKNOWLEDGMENTS

REFERENCES

Arocena, D.G., Iwahashi, C.K., Won, N., Beilina, A., Ludwig, A.L., Tassone, F., Schwartz, P.H., and Hagerman, P.J. (2005). Induction of inclusion formation and disruption of lamin A/C structure by premutation CGG-repeat RNA in human cultured neural cells. Hum. Mol. Genet. *14*, 3661–3671.

Ash, P.E., Bieniek, K.F., Gendron, T.F., Caulfield, T., Lin, W.L., Dejesus-Hernandez, M., van Blitterswijk, M.M., Jansen-West, K., Paul, J.W., 3rd, Rademakers, R., et al. (2013). Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. Neuron *77*, 639–646.

Aspden, J.L., Eyre-Walker, Y.C., Phillips, R.J., Amin, U., Mumtaz, M.A., Brocard, M., and Couso, J.P. (2014). Extensive translation of small Open Reading Frames revealed by Poly-Ribo-Seq. Elife *3*, e03528.

Barbosa, C., Peixeiro, I., and Romão, L. (2013). Gene expression regulation by upstream open reading frames and human disease. PLoS Genet. *9*, e1003529.

Buijsen, R.A., Sellier, C., Severijnen, L.A., Oulad-Abdelghani, M., Verhagen, R.F., Berman, R.F., Charlet-Berguerand, N., Willemsen, R., and Hukema, R.K. (2014). FMRpolyG-positive inclusions in CNS and non-CNS organs of a fragile X premutation carrier with fragile X-associated tremor/ataxia syndrome. Acta Neuropathol. Commun. *2*, 162.

Buijsen, R.A., Visser, J.A., Kramer, P., Severijnen, E.A., Gearing, M., Charlet-Berguerand, N., Sherman, S.L., Berman, R.F., Willemsen, R., and Hukema, R.K. (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.

Calvo, S.E., Pagliarini, D.J., and Mootha, V.K. (2009). Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. Proc. Natl. Acad. Sci. USA *106*, 7507–7512.

Cleary, J.D., and Ranum, L.P. (2014). Repeat associated non-ATG (RAN) translation: new starts in microsatellite expansion disorders. Curr. Opin. Genet. Dev. *26*, 6–15.

DeJesus-Hernandez, M., Mackenzie, I.R., Boeve, B.F., Boxer, A.L., Baker, M., Rutherford, N.J., Nicholson, A.M., Finch, N.A., Flynn, H., Adamson, J., et al. (2011). Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron *72*, 245–256.

Dubińska-Magiera, M., Chmielewska, M., Kozioł, K., Machowska, M., Hutchison, C.J., Goldberg, M.W., and Rzepecki, R. (2016). Xenopus LAP2β protein knockdown affects location of lamin B and nucleoporins and has effect on assembly of cell nucleus and cell viability. Protoplasma *253*, 943–956.

Entezam, A., Biacsi, R., Orrison, B., Saha, T., Hoffman, G.E., Grabczyk, E., Nussbaum, R.L., and Usdin, K. (2007). Regional FMRP deficits and large repeat expansions into the full mutation range in a new Fragile X premutation mouse model. Gene *395*, 125–134.

Feng, Y., Zhang, F., Lokey, L.K., Chastain, J.L., Lakkis, L., Eberhart, D., and Warren, S.T. (1995). Translational suppression by trinucleotide repeat expansion at FMR1. Science *268*, 731–734.

Freibaum, B.D., Lu, Y., Lopez-Gonzalez, R., Kim, N.C., Almeida, S., Lee, K.H., Badders, N., Valentine, M., Miller, B.L., Wong, P.C., et al. (2015). GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. Nature *525*, 129–133.

Fritsch, C., Herrmann, A., Nothnagel, M., Szafranski, K., Huse, K., Schumann, F., Schreiber, S., Platzer, M., Krawczak, M., Hampe, J., and Brosch, M. (2012). Genome-wide search for novel human uORFs and N-terminal protein extensions using ribosomal footprinting. Genome Res. *22*, 2208–2218.

Furukawa, K., Panté, N., Aebi, U., and Gerace, L. (1995). Cloning of a cDNA for lamina-associated polypeptide 2 (LAP2) and identification of regions that specify targeting to the nuclear envelope. EMBO J. *14*, 1626–1636.

Gant, T.M., Harris, C.A., and Wilson, K.L. (1999). Roles of LAP2 proteins in nuclear assembly and DNA replication: truncated LAP2beta proteins alter lamina assembly, envelope formation, nuclear size, and DNA replication efficiency in Xenopus laevis extracts. J. Cell Biol. *144*, 1083–1096.

Greco, C.M., Berman, R.F., Martin, R.M., Tassone, F., Schwartz, P.H., Chang, A., Trapp, B.D., Iwahashi, C., Brunberg, J., Grigsby, J., et al. (2006). Neuropathology of fragile X-associated tremor/ataxia syndrome (FXTAS). Brain *129*, 243–255.

Hagerman, P.J., and Hagerman, R.J. (2004). The fragile-X premutation: a maturing perspective. Am. J. Hum. Genet. *74*, 805–816.

Hagerman, R.J., Leehey, M., Heinrichs, W., Tassone, F., Wilson, R., Hills, J., Grigsby, J., Gage, B., and Hagerman, P.J. (2001). Intention tremor, parkinsonism, and generalized brain atrophy in male carriers of fragile X. Neurology *57*, 127–130.

Hashem, V., Galloway, J.N., Mori, M., Willemsen, R., Oostra, B.A., Paylor, R., and Nelson, D.L. (2009). Ectopic expression of CGG containing mRNA is neurotoxic in mammals. Hum. Mol. Genet. *18*, 2443–2451.

Hoem, G., Raske, C.R., Garcia-Arocena, D., Tassone, F., Sanchez, E., Ludwig, A.L., Iwahashi, C.K., Kumar, M., Yang, J.E., and Hagerman, P.J. (2011). CGG-repeat length threshold for FMR1 RNA pathogenesis in a cellular model for FXTAS. Hum. Mol. Genet. *20*, 2161–2170.

Hukema, R.K., Buijsen, R.A., Raske, C., Severijnen, L.A., Nieuwenhuizen-Bakker, I., Minneboo, M., Maas, A., de Crom, R., Kros, J.M., Hagerman, P.J., et al. (2014). Induced expression of expanded CGG RNA causes mitochondrial dysfunction in vivo. Cell Cycle *13*, 2600–2608.

Hukema, R.K., Buijsen, R.A., Schonewille, M., Raske, C., Severijnen, L.A., Nieuwenhuizen-Bakker, I., Verhagen, R.F., van Dessel, L., Maas, A., Charlet-Berguerand, N., et al. (2015). Reversibility of neuropathology and motor deficits in an inducible mouse model for FXTAS. Hum. Mol. Genet. *24*, 4948–4957.

Ingolia, N.T., Lareau, L.F., and Weissman, J.S. (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. Cell *147*, 789–802.

Iwahashi, C.K., Yasui, D.H., An, H.J., Greco, C.M., Tassone, F., Nannen, K., Babineau, B., Lebrilla, C.B., Hagerman, R.J., and Hagerman, P.J. (2006). Protein composition of the intranuclear inclusions of FXTAS. Brain *129*, 256–271.

Jacquemont, S., Hagerman, R.J., Leehey, M., Grigsby, J., Zhang, L., Brunberg, J.A., Greco, C., Des Portes, V., Jardini, T., Levine, R., et al. (2003). Fragile X premutation tremor/ataxia syndrome: molecular, clinical, and neuroimaging correlates. Am. J. Hum. Genet. *72*, 869–878.

Jacquemont, S., Hagerman, R.J., Leehey, M.A., Hall, D.A., Levine, R.A., Brunberg, J.A., Zhang, L., Jardini, T., Gane, L.W., Harris, S.W., et al. (2004). Penetrance of the fragile X-associated tremor/ataxia syndrome in a premutation carrier population. JAMA *291*, 460–469.

Ji, Z., Song, R., Regev, A., and Struhl, K. (2015). Many lncRNAs, 5'UTRs, and pseudogenes are translated and some are likely to express functional proteins. eLife *4*, e08890.

Jin, P., Zarnescu, D.C., Zhang, F., Pearson, C.E., Lucchesi, J.C., Moses, K., and Warren, S.T. (2003). RNA-mediated neurodegeneration caused by the fragile X premutation rCGG repeats in Drosophila. Neuron *39*, 739–747.

Jin, P., Duan, R., Qurashi, A., Qin, Y., Tian, D., Rosser, T.C., Liu, H., Feng, Y., and Warren, S.T. (2007). Pur alpha binds to rCGG repeats and modulates repeat-mediated neurodegeneration in a Drosophila model of fragile X tremor/ataxia syndrome. Neuron *55*, 556–564.

Jovičić, A., Mertens, J., Boeynaems, S., Bogaert, E., Chai, N., Yamada, S.B., Paul, J.W., 3rd, Sun, S., Herdy, J.R., Bieri, G., et al. (2015). Modifiers of C9orf72 dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS. Nat. Neurosci. *18*, 1226–1229.

Kaplan, E.S., Cao, Z., Hulsizer, S., Tassone, F., Berman, R.F., Hagerman, P.J., and Pessah, I.N. (2012). Early mitochondrial abnormalities in hippocampal neurons cultured from Fmr1 pre-mutation mouse model. J. Neurochem. *123*, 613–621.

Kearse, M.G., Green, K.M., Krans, A., Rodriguez, C.M., Linsalata, A.E., Goldstrohm, A.C., and Todd, P.K. (2016). CGG repeat-associated non-AUG translation utilizes a cap-dependent scanning mechanism of initiation to produce toxic proteins. Mol. Cell *62*, 314–322.

Kenneson, A., Zhang, F., Hagedorn, C.H., and Warren, S.T. (2001). Reduced FMRP and increased FMR1 transcription is proportionally associated with CGG repeat number in intermediate-length and premutation carriers. Hum. Mol. Genet. *10*, 1449–1454.

Liu, J., Koscielska, K.A., Cao, Z., Hulsizer, S., Grace, N., Mitchell, G., Nacey, C., Githinji, J., McGee, J., Garcia-Arocena, D., et al. (2012). Signaling defects in iPSC-derived fragile X premutation neurons. Hum. Mol. Genet. *21*, 3795–3805.

Mackenzie, I.R., Frick, P., Grässer, F.A., Gendron, T.F., Petrucelli, L., Cashman, N.R., Edbauer, D., Kremmer, E., Prudlo, J., Troost, D., and Neumann, M. (2015). Quantitative analysis and clinico-pathological correlations of different dipeptide repeat protein pathologies in C9ORF72 mutation carriers. Acta Neuropathol. *130*, 845–861.

Mori, K., Weng, S.M., Arzberger, T., May, S., Rentzsch, K., Kremmer, E., Schmid, B., Kretzschmar, H.A., Cruts, M., Van Broeckhoven, C., et al. (2013). The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. Science *339*, 1335–1338.

Nelson, D.L., Orr, H.T., and Warren, S.T. (2013). The unstable repeats—three evolving faces of neurological disease. Neuron *77*, 825–843.

Oh, S.Y., He, F., Krans, A., Frazer, M., Taylor, J.P., Paulson, H.L., and Todd, P.K. (2015). RAN translation at CGG repeats induces ubiquitin proteasome system impairment in models of fragile X-associated tremor ataxia syndrome. Hum. Mol. Genet. *24*, 4317–4326.

Pretto, D.I., Kumar, M., Cao, Z., Cunningham, C.L., Durbin-Johnson, B., Qi, L., Berman, R., Noctor, S.C., Hagerman, R.J., Pessah, I.N., and Tassone, F. (2014). Reduced excitatory amino acid transporter 1 and metabotropic glutamate receptor 5 expression in the cerebellum of fragile X mental retardation gene 1 premutation carriers with fragile X-associated tremor/ataxia syndrome. Neurobiol. Aging *35*, 1189–1197.

Primerano, B., Tassone, F., Hagerman, R.J., Hagerman, P., Amaldi, F., and Bagni, C. (2002). Reduced FMR1 mRNA translation efficiency in fragile X patients with premutations. RNA *8*, 1482–1488.

Qin, M., Entezam, A., Usdin, K., Huang, T., Liu, Z.H., Hoffman, G.E., and Smith, C.B. (2011). A mouse model of the fragile X premutation: effects on behavior, dendrite morphology, and regional rates of cerebral protein synthesis. Neurobiol. Dis. *42*, 85–98.

Renton, A.E., Majounie, E., Waite, A., Simón-Sánchez, J., Rollinson, S., Gibbs, J.R., Schymick, J.C., Laaksovirta, H., van Swieten, J.C., Myllykangas, L., et al.; ITALSGEN Consortium (2011). A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron *72*, 257–268.

Ross-Inta, C., Omanska-Klusek, A., Wong, S., Barrow, C., Garcia-Arocena, D., Iwahashi, C., Berry-Kravis, E., Hagerman, R.J., Hagerman, P.J., and Giulivi, C. (2010). Evidence of mitochondrial dysfunction in fragile X-associated tremor/ataxia syndrome. Biochem. J. *429*, 545–552.

Sellier, C., Rau, F., Liu, Y., Tassone, F., Hukema, R.K., Gattoni, R., Schneider, A., Richard, S., Willemsen, R., Elliott, D.J., et al. (2010). Sam68 sequestration and partial loss of function are associated with splicing alterations in FXTAS patients. EMBO J. *29*, 1248–1261.

Sellier, C., Freyermuth, F., Tabet, R., Tran, T., He, F., Ruffenach, F., Alunni, V., Moine, H., Thibault, C., Page, A., et al. (2013). Sequestration of DROSHA and DGCR8 by expanded CGG RNA repeats alters microRNA processing in fragile X-associated tremor/ataxia syndrome. Cell Rep. *3*, 869–880.

Shumaker, D.K., Lee, K.K., Tanhehco, Y.C., Craigie, R., and Wilson, K.L. (2001). LAP2 binds to BAF.DNA complexes: requirement for the LEM domain and modulation by variable regions. EMBO J. *20*, 1754–1764.

Sofola, O.A., Jin, P., Qin, Y., Duan, R., Liu, H., de Haro, M., Nelson, D.L., and Botas, J. (2007). RNA-binding proteins hnRNP A2/B1 and CUGBP1 suppress fragile X CGG premutation repeat-induced neurodegeneration in a Drosophila model of FXTAS. Neuron *55*, 565–571.

Sonenberg, N., and Hinnebusch, A.G. (2009). Regulation of translation initiation in eukaryotes: mechanisms and biological targets. Cell *136*, 731–745.

Tassone, F., Hagerman, R.J., Loesch, D.Z., Lachiewicz, A., Taylor, A.K., and Hagerman, P.J. (2000). Fragile X males with unmethylated, full mutation trinucleotide repeat expansions have elevated levels of FMR1 messenger RNA. Am. J. Med. Genet. *94*, 232–236.

Tassone, F., Beilina, A., Carosi, C., Albertosi, S., Bagni, C., Li, L., Glover, K., Bentley, D., and Hagerman, P.J. (2007). Elevated FMR1 mRNA in premutation carriers is due to increased transcription. RNA *13*, 555–562.

Tassone, F., Iong, K.P., Tong, T.H., Lo, J., Gane, L.W., Berry-Kravis, E., Nguyen, D., Mu, L.Y., Laffin, J., Bailey, D.B., and Hagerman, R.J. (2012). FMR1 CGG allele size and prevalence ascertained through newborn screening in the United States. Genome Med. *4*, 100.

Todd, P.K., Oh, S.Y., Krans, A., He, F., Sellier, C., Frazer, M., Renoux, A.J., Chen, K.C., Scaglione, K.M., Basrur, V., et al. (2013). CGG repeat-associated translation mediates neurodegeneration in fragile X tremor ataxia syndrome. Neuron *78*, 440–455.

Willemsen, R., Hoogeveen-Westerveld, M., Reis, S., Holstege, J., Severijnen, L.A., Nieuwenhuizen, I.M., Schrier, M., van Unen, L., Tassone, F., Hoogeveen, A.T., et al. (2003). The FMR1 CGG repeat mouse displays ubiquitin-positive intranuclear neuronal inclusions; implications for the cerebellar tremor/ataxia syndrome. Hum. Mol. Genet. *12*, 949–959.

Zhang, K., Donnelly, C.J., Haeusler, A.R., Grima, J.C., Machamer, J.B., Steinwald, P., Daley, E.L., Miller, S.J., Cunningham, K.M., Vidensky, S., et al. (2015). The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. Nature *525*, 56–61.

Zhang, Y.J., Gendron, T.F., Grima, J.C., Sasaguri, H., Jansen-West, K., Xu, Y.F., Katzman, R.B., Gass, J., Murray, M.E., Shinohara, M., et al. (2016). C9ORF72 poly(GA) aggregates sequester and impair HR23 and nucleocytoplasmic transport proteins. Nat. Neurosci. http://dx.doi.org/10.1038/nn.4272.

Zu, T., Gibbens, B., Doty, N.S., Gomes-Pereira, M., Huguet, A., Stone, M.D., Margolis, J., Peterson, M., Markowski, T.W., Ingram, M.A., et al. (2011). Non-ATG-initiated translation directed by microsatellite expansions. Proc. Natl. Acad. Sci. USA *108*, 260–265.

Zu, T., Liu, Y., Bañez-Coronel, M., Reid, T., Pletnikova, O., Lewis, J., Miller, T.M., Harms, M.B., Falchook, A.E., Subramony, S.H., et al. (2013). RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. Proc. Natl. Acad. Sci. USA *110*, E4968–E4977.