Article

Neuron

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.

Sellier et al., 2017, Neuron *93*, 331–347
January 18, 2017
http://dx.doi.org/10.1016/j.neuron.2016.12.016

CellPress

Neuron
Article

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

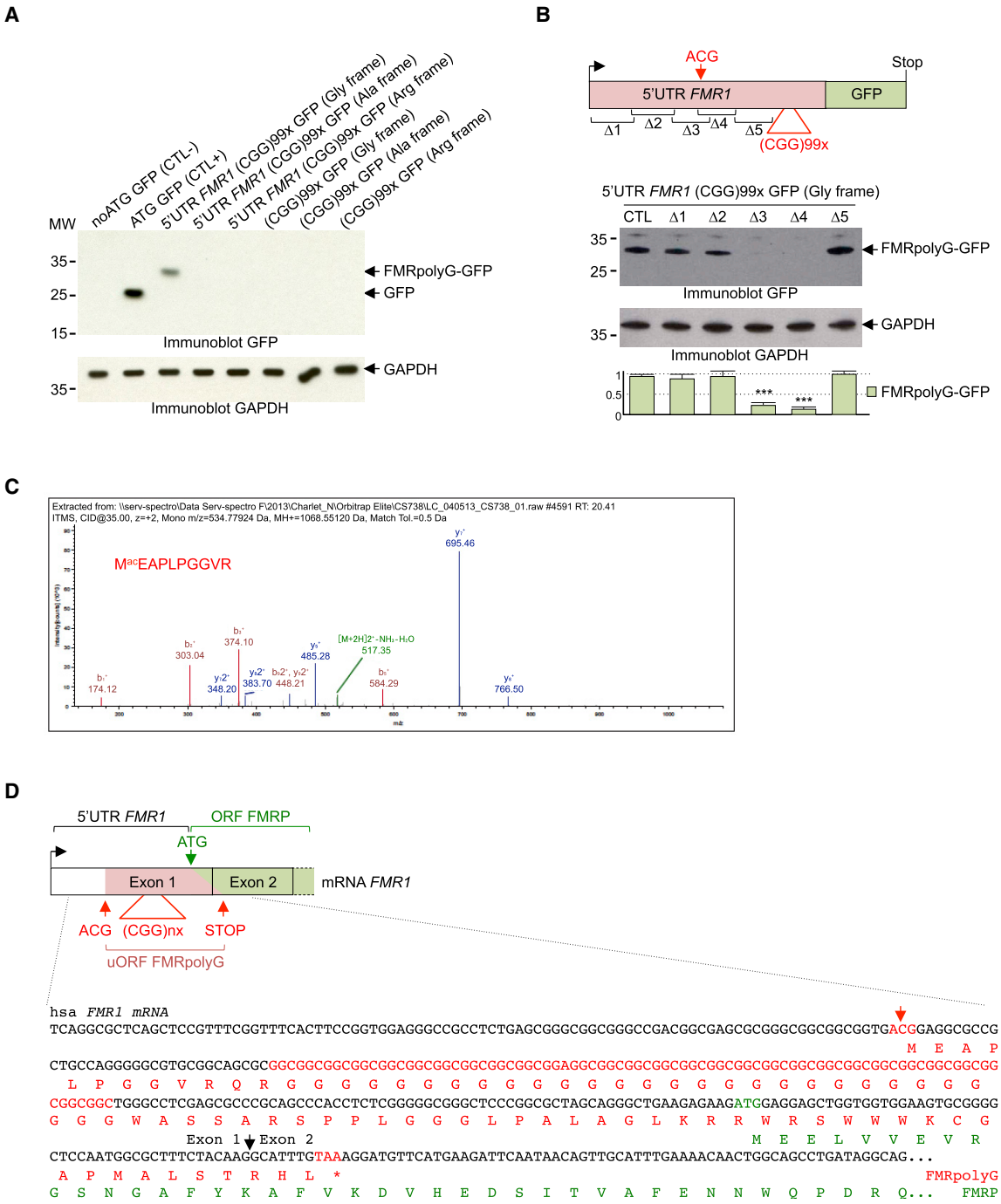


Figure 1. Translation of CGG Repeats Initiates at an Upstream Near-Cognate Codon

(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 $p < 0.001$.

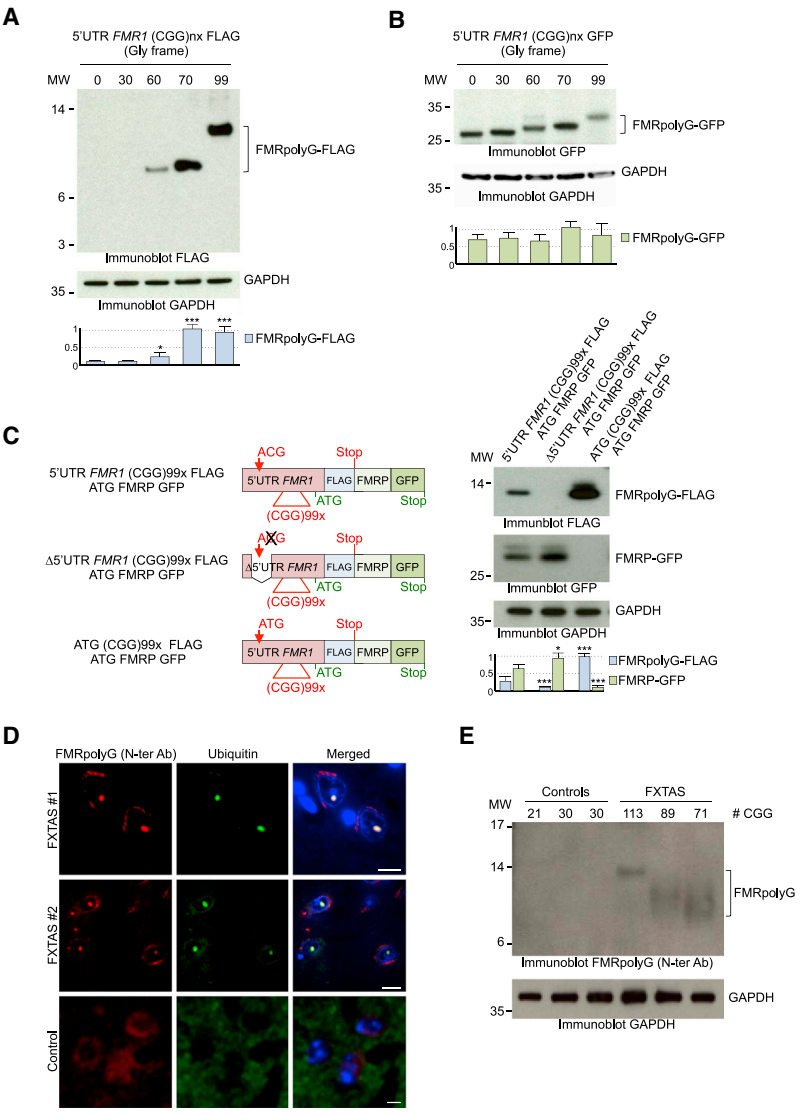


Figure 2. A Minimum of 60 Expanded CGG Repeats Is Required to Detect FMRpolyG

(A) Upper, immunoblotting against the FLAG tag on the soluble lysate fraction of HeLa cells transfected for 24 hr with various lengths of expanded CGG repeats embedded in the 5′ UTR of *FMR1* and fused in the glycine frame with the FLAG tag. Middle, control immunoblotting against GAPDH. Lower, quantification of FMRpolyG-FLAG levels reported to GAPDH.

(B) Identical to (A), but with CGG repeats embedded in the 5′ UTR of *FMR1* fused in the glycine frame with a GFP tag instead of a FLAG tag.

(C) Left, schemes of the 5′ UTR of *FMR1* constructs with ACG mutations. FMRpolyG uORF is FLAG-tagged in the glycine frame, while FMRP ORF is GFP-tagged. Right, upper, and middle, immunoblotting against FLAG, GFP, and GAPDH on the soluble lysate fraction of HeLa cells transfected for 24 hr with mutants of the 5′ UTR of *FMR1*; expanded CGG repeats are fused to the FLAG tag in the glycine frame, while the downstream FMRP ORF is fused to the GFP. Right lower, quantification of FMRpolyG-FLAG and FMRP-GFP levels reported to GAPDH.

(D) Immunofluorescence against the FMRpolyG N terminus and ubiquitin on brain sections (hippocampal area) of FXTAS or control individuals. Scale bars, 10 μm. Nuclei were counter-stained with DAPI.

(E) Immunoblotting against the FMRpolyG N terminus of insoluble fraction of brain lysate (cerebellum area) of FXTAS and age-matched individuals.

The numbers of expanded CGG in *FMR1* are indicated as # CGG. Error bars indicate SEM of three independent transfections. Student's t test, * indicates $p < 0.05$, *** indicates $p < 0.001$.

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

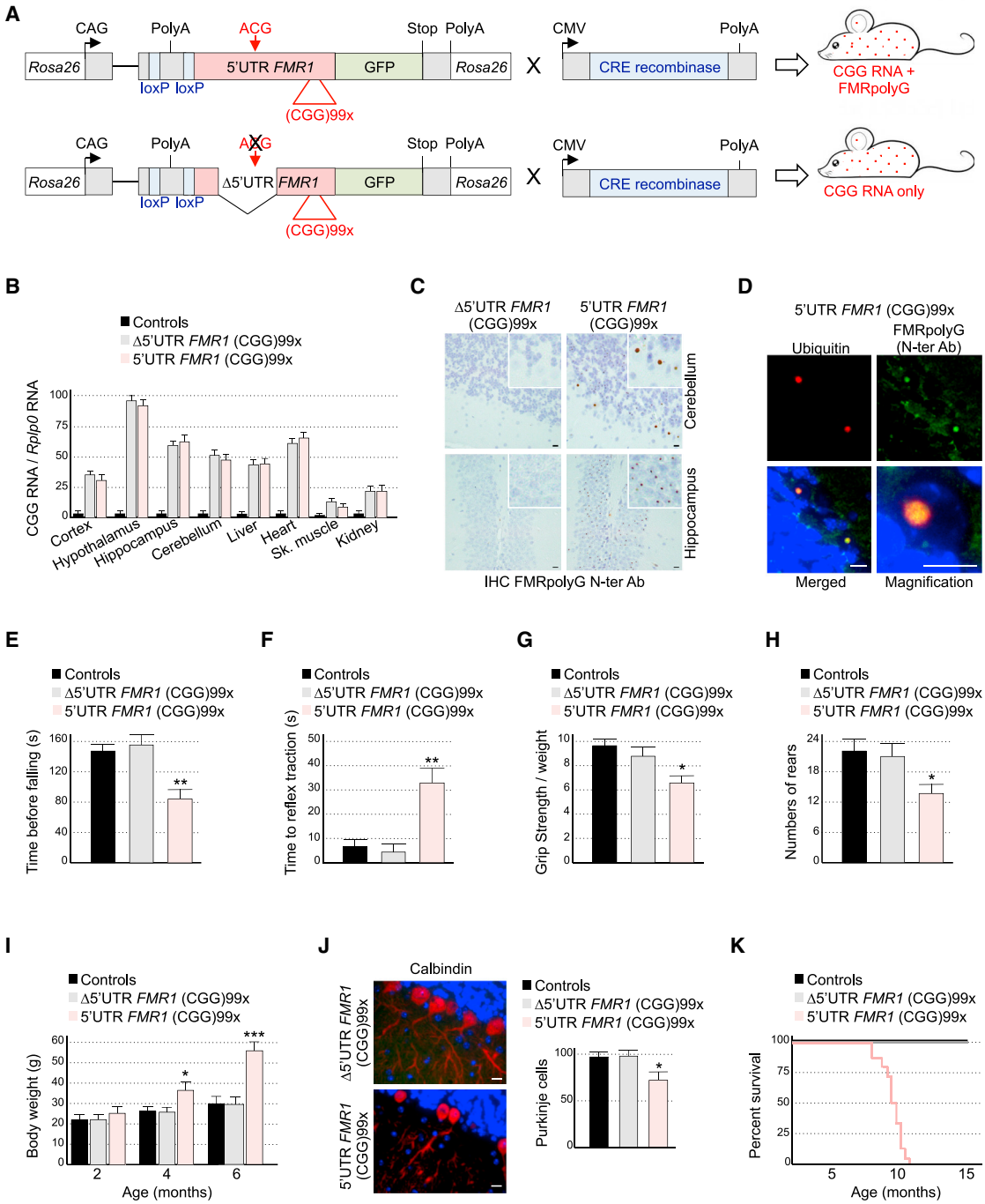


Figure 3. Expression of FMRpolyG Is Pathogenic in Mice

(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 $p < 0.001$.

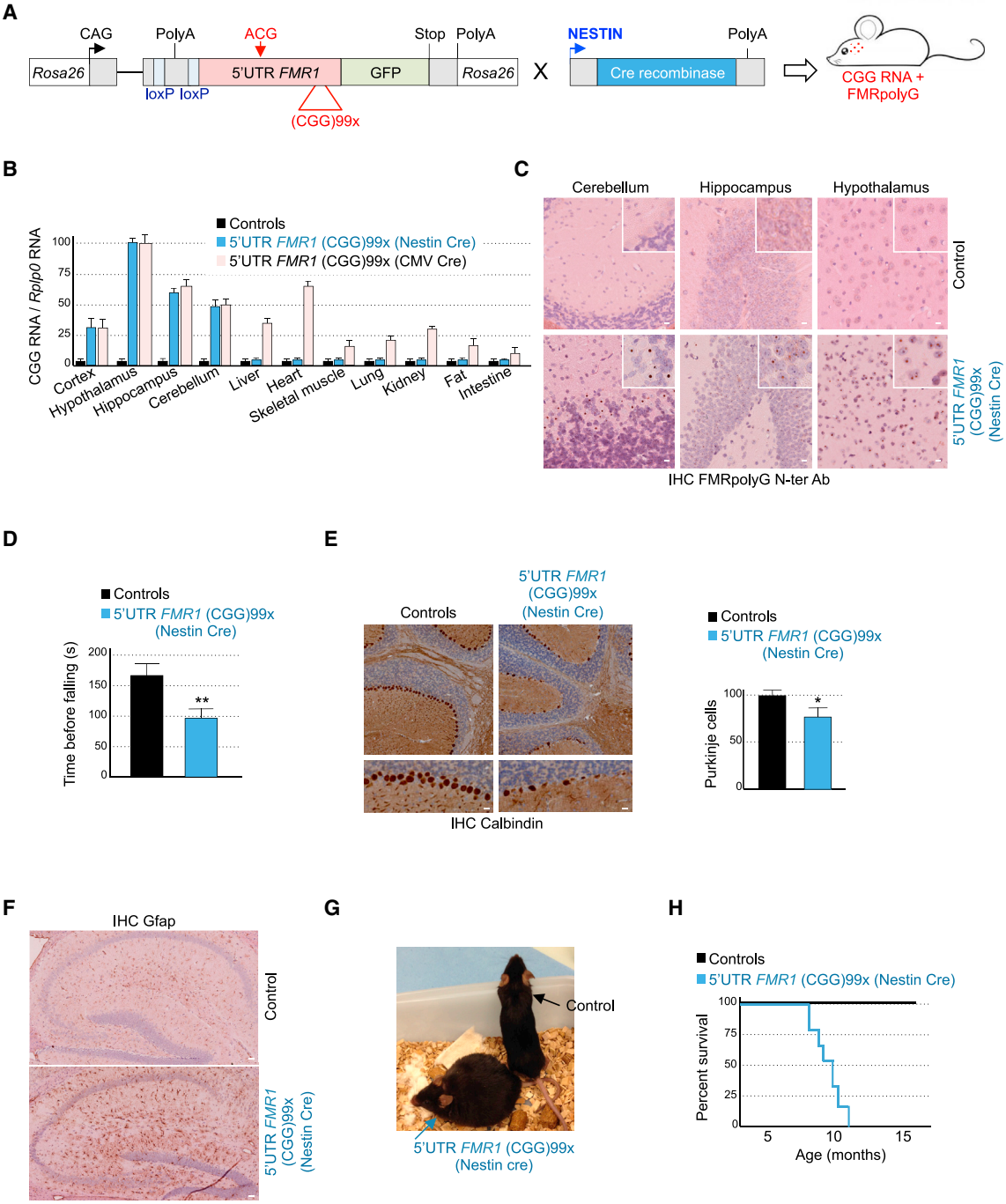


Figure 4. Neuronal Expression of FMRpolyG Is Pathogenic in Mice

(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 $p < 0.01$.

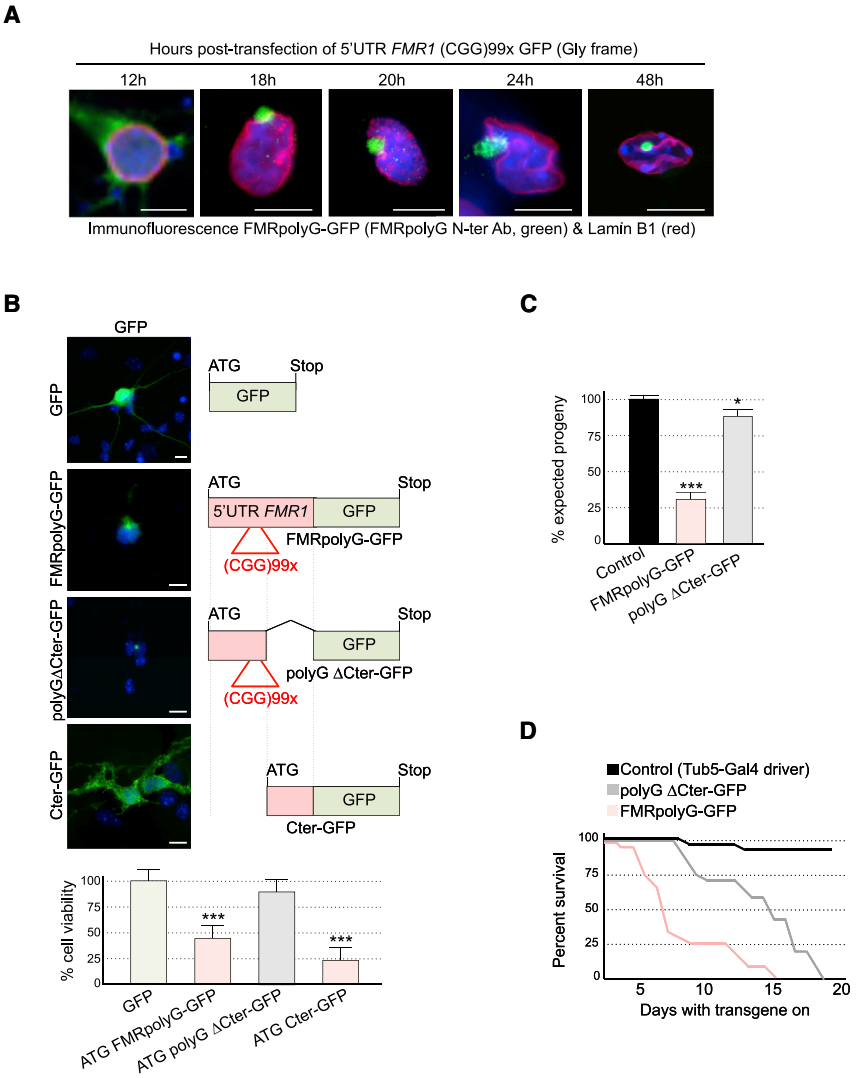


Figure 5. FMRpolyG Toxicity Is Influenced by Its Carboxyl Terminus

(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*/+).

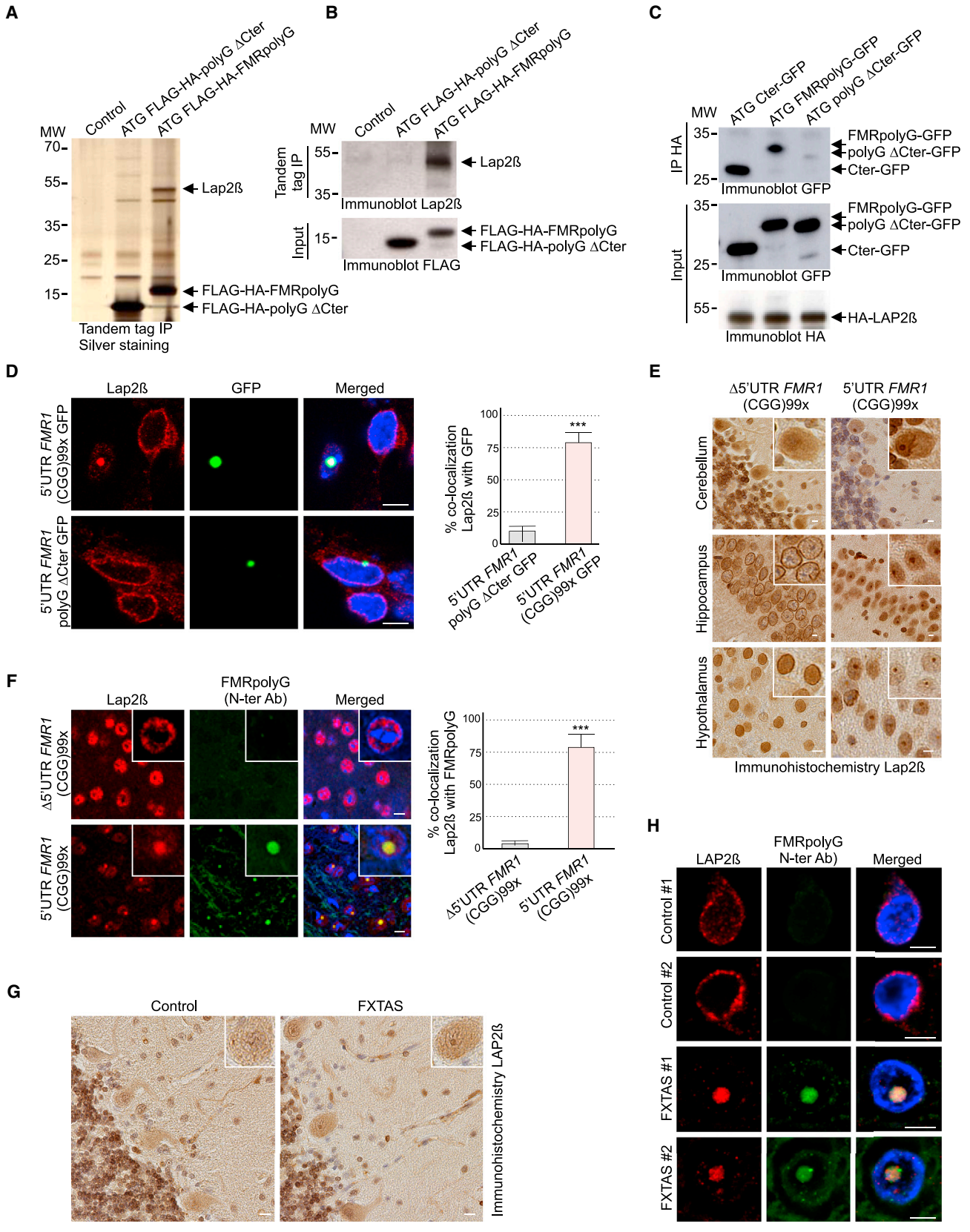
Error bars indicate SEM. Student's t test, * indicates $p < 0.05$, *** indicates $p < 0.001$.

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.

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



(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 de-localization 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 $p < 0.001$.

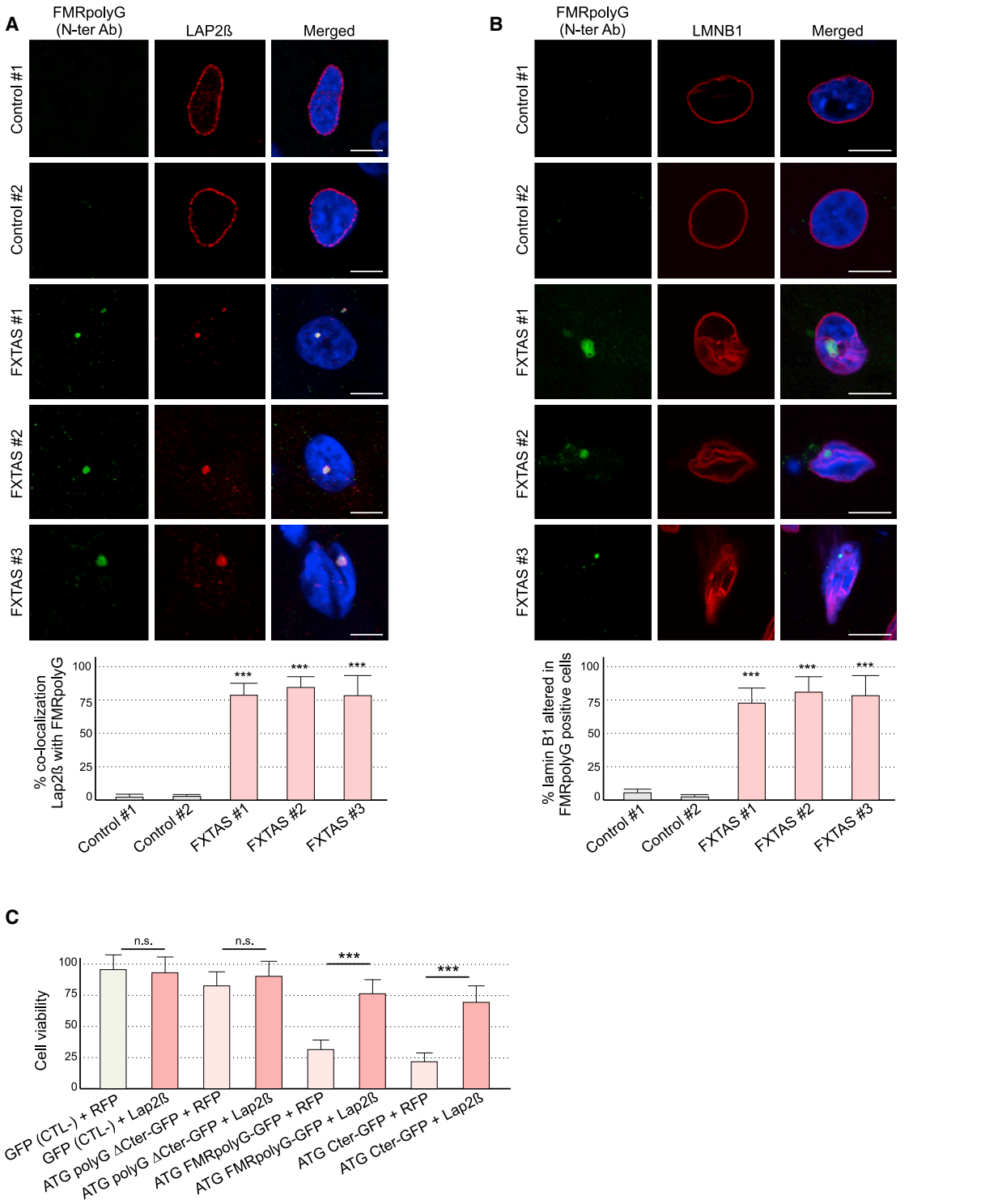


Figure 7. LAP2β Rescues Neuronal Cell Death Induced by FMRpolyG

(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).

(legend continued on next page)

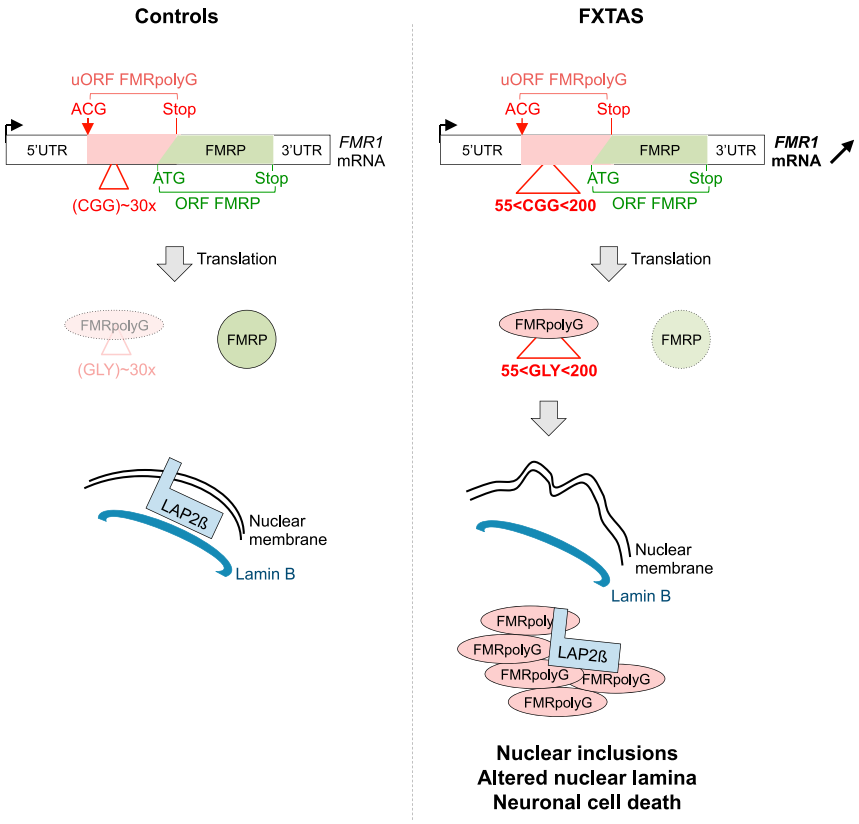


Figure 8. A Working Model for Pathogenicity in FXTAS

Expanded CGG repeats are translated into a polyglycine-containing protein, FMRpolyG, through initiation to a non-canonical ACG codon located upstream of the CGG repeats. In FXTAS, higher expression of *FMR1* mRNA and increased translation and stability of the expanded CGG repeats result in accumulation of FMRpolyG in nuclear aggregates that sequester the LAP2β protein and alter the nuclear lamina architecture.

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 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; Jovićić et al., 2015).

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

(B) Upper, immunofluorescence against FMRpolyG N terminus and LMNB1 on neuronal cultures differentiated 40 days from iPS cells of FXTAs patients or control individuals. Lower, quantification of lamin B1 alteration in FMRpolyG-positive cells in neurons from iPSC of FXTAS and control individuals (n = 100 neurons, three independent cultures).

(C) Cell viability of neuronal N2A cells transfected (n = 3 transfections) with ATG-driven FMRpolyG-GFP either full-length or deleted of its N or C terminus and with a plasmid expressing RFP as control or Ha-tagged LAP2β.

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

provide some molecular basis for the threshold of severity and incomplete penetrance observed in FXTAS, as most of premutation 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 neurodegeneration 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

We thank Prof. P.J. Hagerman (Department of Biochemistry and Molecular Medicine, School of Medicine, University of California, Davis) for the kind gift of human samples. We thank all IGBMC and ICS common core facilities for assistance. This work was supported by ANR-12-RARE-0001 E-RARE "CURE FXTAS" (N.C.B. and R.K.H.), ANR-14-RARE-0003 E-RARE "Drug_FXSPreMut" (N.C.B., R.W., and C.M.), ERC-2012-StG 310659 "RNA DISEASES", ANR-14-CE10-0016-01 "MITO-FXTAS" (N.C.B.), ANR-10-INBS-07 PHENOMIN (I.C.S.), ANR-10-LABX-0030-INRT (IGBMC), ANR-10-IDEX-0002-02 (IGBMC), Hersenstichting Nederland F2012(1)-101 (R.W.), Hersenstichting Nederland F2015(1)-02 (R.K.H.), National Institutes of Health NINDS RO1 NS079775 (R.W.), NIH R01NS08681001 (P.K.T.), and the Department of Veterans Affairs BLRD 1I21BX001841 (P.K.T). P.K.T. serves as a consultant for Denali Therapeutics; this consultation role does not relate to the work described here.

Received: March 24, 2016
Revised: October 6, 2016
Accepted: December 2, 2016
Published: January 5, 2017

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.

Neuron, Volume *93*

Supplemental Information

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

Chantal Sellier, Ronald A.M. Buijsen, Fang He, Sam Natla, Laura Jung, Philippe Tropel, Angeline Gaucherot, Hugues Jacobs, Hamid Meziane, Alexandre Vincent, Marie-France Champy, Tania Sorg, Guillaume Pavlovic, Marie Wattenhofer-Donze, Marie-Christine Birling, Mustapha Oulad-Abdelghani, Pascal Eberling, Frank Ruffenach, Mathilde Joint, Mathieu Anheim, Veronica Martinez-Cerdeno, Flora Tassone, Rob Willemsen, Renate K. Hukema, Stéphane Viville, Cecile Martinat, Peter K. Todd, and Nicolas Charlet-Berguerand

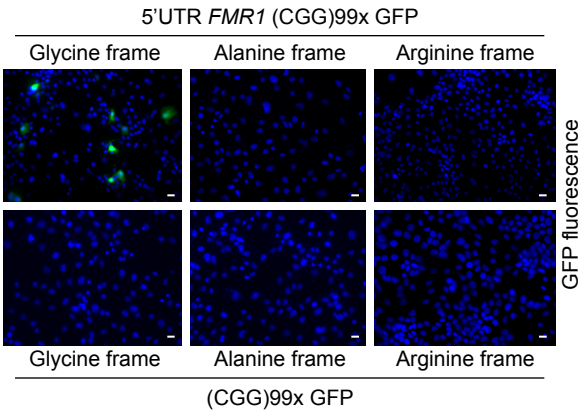
Figure S1

A

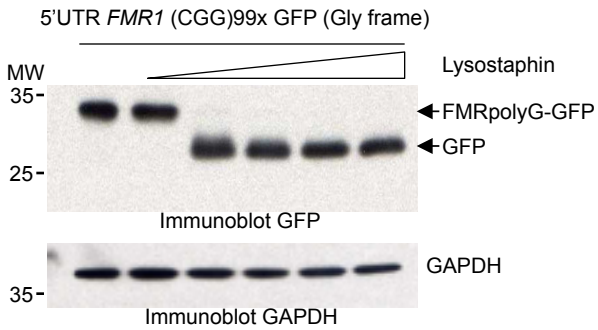
5'UTR *FMR1* (CGG)99x GFP (Gly frame)

```
...ACGGTGGGAGGTCTATATAAGCAGAGCTCtctggctaactagagaacccactgcttactggcttatcgaaattaatacgactcactataggg
agacccaagctggctagcgtttaaacttaagcttcgtcgacagggcccaccgcggaggccggccTCAGGCGCTCAGCTCCGTTTCGGTTTCACTT
CCGGTGGAGGGCCGCCTCTGAGCGGGCGGCGGGCCGACGGCGAGCGCGGGCGGCGGCGGTGACGGAGGCGCCGCTGCCAGGGGGCGTGCGGCAGC
GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGC
GGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG
GCGGCGGCGGCGGCTGGGCCTCGAGCGCCCGCAGCCCACCTCTCGGGGGCGGGCTCCCGGCGCTAGCAGGGCTGAAGAGAAGATGGAGGAGCTGG
TGGTGGAAGTGCGGGGCTCCAATGGCGCTTTCTACAAGGCACCTAggaGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGT
CGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCT
GCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAG
CAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGA
GGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGT
ACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGC
AGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGC
CCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACA
AGTAActctagagcggccgcttcgagCAGACATGATAAGATA...
```

B



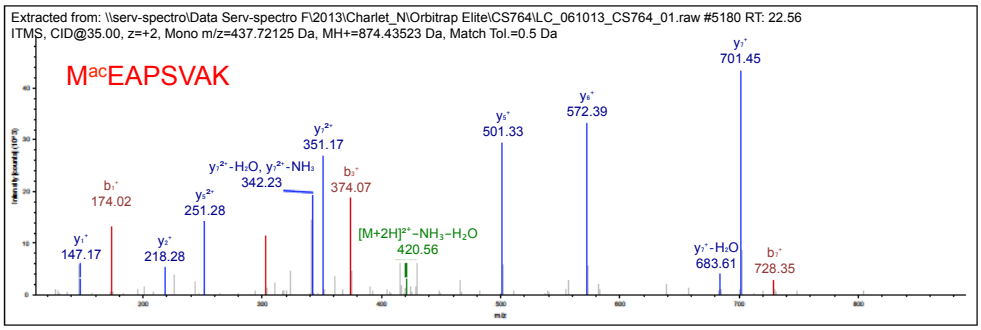
C



D

5'UTR *FMR1* LysC (CGG)99x GFP

```
...AGCGCGGGCGGCGGCGGTGACGGAGGCGCCgtccgtcgcaaaactagttccccgcggctcccgGGCGCCGCTGCCAGGGGGCGTGCGGCAGCGCGG...
                              M  E  A  P  S  V  A  K  L  V  P  R  G  S  R  A  P  L  P  G  G  V  R  Q  R  G
```



E

```
Homo Sapiens          ...CGGGCCGACGGCGAGCGCGGGCGGCGGCGGTGACGGAGGCGCCGCTGCCAGGGGGCGTGCGGCAGCGCGGCGGCGG...
Pan troglodytes       ...CGGGCCGACGGCGAGCGCGGGCGGCGGCGGTGACGGAGGCGCCGCTGCCAGGGGGCGTGCGGCAGCGCGGCGGCGG...
Macaca arctoides      ...CGGGCCGACGGCGAGCGCGGGCGGCGGCGGTGACGGAGGCGCCGCTGCCAGGGGGCGTGCGGCAGCGCGGCGGCGG...
Nomascus leucogenys   ...CGGGCCGACGGCGAGCGCGGGCGGCGGCGGTGACGGAGGCGCCGCTGCCAGGGGGCGTGCGGCAGCGCGGCGGCGG...
Otolemur garnettii    ...CGGGCCGACGGCtAGCGCGGGCGGCGaCGGTGACGGtGGCGCCGCTGCCgGGGGGCGTGCGGaAGCGCGGCGGCGG...
Sus scrofa            ...tGaGCCGACGGCGAGCGCaGGCGGCGGCGGTGACGGcGGCGCCGCTGCCgGGGGGCGTGCGGCAGCGCGGCGGCGG...
Canis familiaris      ...tGaGCCGACGGCGAGCGCGGGtGGCGGCGGTGACGGcGGCGCCGCTGCCgGGGGGCGTGCGaGAGCGCGGCGGCGG...
Rattus norvegicus     ...CCaaCgGACGGCGAGCGCGGGCGGtGtCaGTGACGGcGGCGCCGCTGCCgGGGGGCGTGCGGtAaCGCGGCGGCGG...
Mus musculus          ...CaaaCgGACGGCGAGCGCGGGCGGtGGCaGTGACGGcGGCGCCGCTGCCgGGGGGCGTGCGGtAaCGCGGCGGCGG...
Kozac consensus sequence                                RCCRCCATGG
```

Figure S1 (related to figure 1).
Translation of CGG repeats requires a near-cognate codon.

(**A**) Partial sequence of the 5'UTR FMR1 CGG99x GFP (glycine frame) plasmid (Addgene #63091). CMV promoter and sv40 polyadenylation sequences are indicated in black uppercase at the start and the end of the sequence, respectively. The plasmid multiple cloning site sequence is in black lowercase with Fse1 and Xba1 cloning site underlined. The 5'UTR sequence of human *FMR1* is in bold black uppercase with expanded CGG repeats in red. The ACG near cognate codon is also indicated in red. The stop codon of FMRpolyG and the ATG codon of GFP were deleted and replaced by a glycine codon (gga). The GFP sequence is in green uppercase.

(**B**) GFP fluorescence of HeLa cells transfected for 24 hours 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.

(**C**) Immunoblotting against the GFP on the soluble lysate fraction of HeLa cells transfected for 24 hours with expanded CGG repeats embedded within the 5'UTR of *FMR1* and fused to the GFP in the glycine frame and treated with increasing quantity of lysostaphin, a glycine endopeptidase.

(**D**) Upper panel, sequence of the mutant *FMR1* 5'UTR including a LysC restriction site. Lower panel, LC-MS/MS spectra of the N-terminal part of the immunoprecipitated and LysC digested protein translated from expanded CGG embedded in the LysC mutant 5'UTR of *FMR1*.

(**E**) Sequence alignment of the 5'UTR of *FMR1* from human, chimpanzee, macaque, gibbon, galago, wild boar, dog, rat and mouse with Kozac consensus sequence. ACG translation start of FMRpolyG is indicated in red, the conserved nucleotides in the Kozac sequence are indicated in bold black.

Figure S2

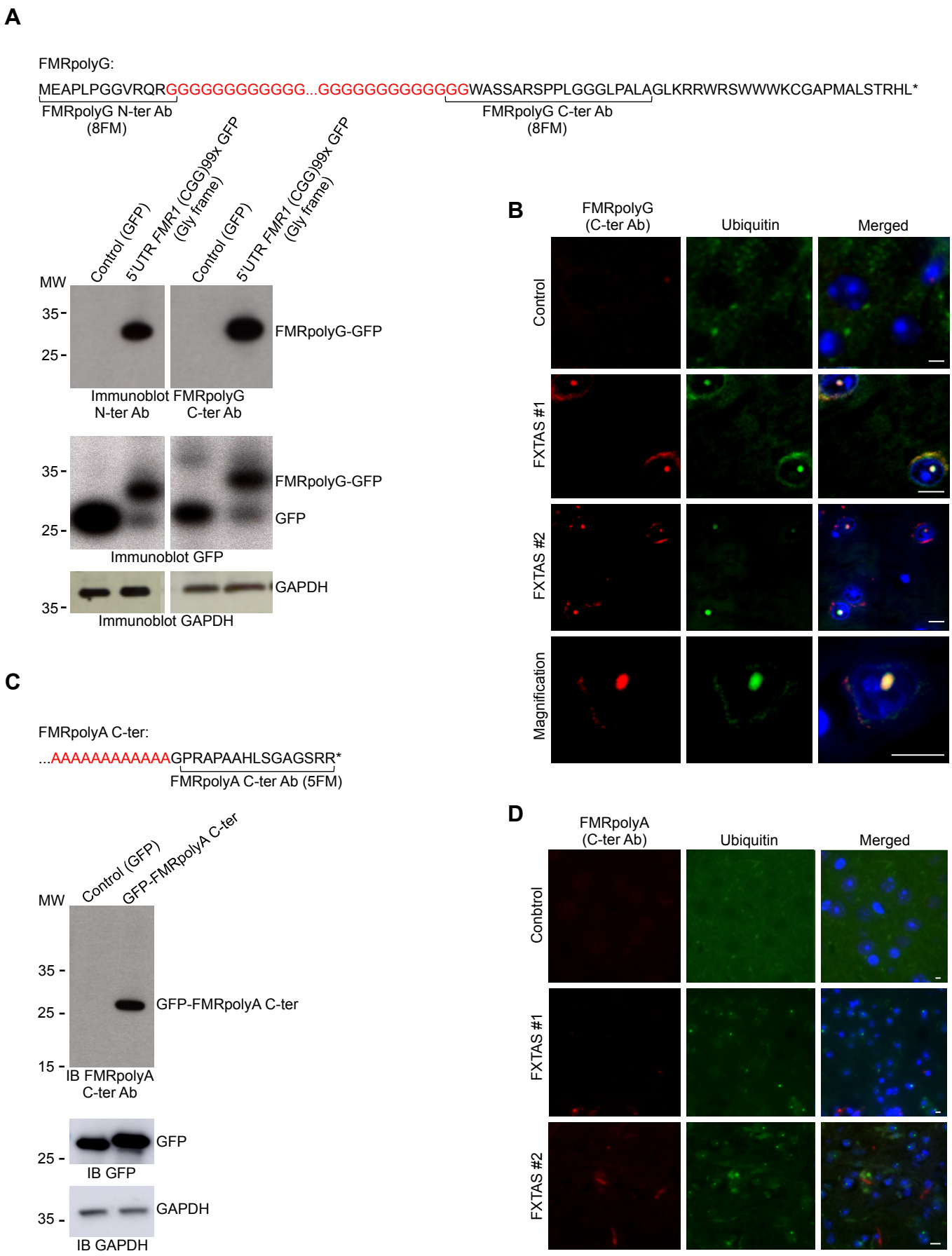

A
FMRpolyG:
MEAPLPGGVRQRGGGGGGGGGGGG...GGGGGGGGGGGGGGWASSARSPPLGGGLPALAGLKRRWRSWWWKCGAPMALSTRHL*
FMRpolyG N-ter Ab (8FM)
FMRpolyG C-ter Ab (8FM)
Control (GFP)
5'UTR FMR1 (CGG)99x GFP (Gly frame)
Control (GFP)
5'UTR FMR1 (CGG)99x GFP (Gly frame)
MW
35
25
FMRpolyG-GFP
Immunoblot FMRpolyG N-ter Ab
Immunoblot FMRpolyG C-ter Ab
35
25
FMRpolyG-GFP
GFP
Immunoblot GFP
35
GAPDH
Immunoblot GAPDH
B
FMRpolyG (C-ter Ab)
Ubiquitin
Merged
Control
FXTAS #1
FXTAS #2
Magnification
C
FMRpolyA C-ter:
...AAAAAAAAAAAAGPRAPAAHLSGAGSRR*
FMRpolyA C-ter Ab (5FM)
Control (GFP)
GFP-FMRpolyA C-ter
MW
35
25
15
GFP-FMRpolyA C-ter
IB FMRpolyA C-ter Ab
25
GFP
IB GFP
35
GAPDH
IB GAPDH
D
FMRpolyA (C-ter Ab)
Ubiquitin
Merged
Conbtrol
FXTAS #1
FXTAS #2

Figure S2 (related to figure 2).
Generation of antibodies specific to FMRpolyG.

(**A**) Upper panel, amino acid sequence of FMRpolyG. The N- and C-terminal peptide sequences used to generate 8FM and 9FM mouse monoclonal antibodies are underlined. Lower panel, immunoblotting validation of 8FM and 9FM antibodies directed respectively against the N- and C-terminal parts of FMRpolyG on the soluble lysate fraction of HeLa cells transfected for 24 hours with expanded CGG repeats embedded in the 5'UTR of *FMR1* and fused to the GFP in the glycine frame.

(**B**) Immunofluorescence against the C-terminal part of FMRpolyG (red, 9FM antibody) and ubiquitin (green) on brain sections (hippocampal area) of FXTAS patients compared to a age-matched control individual.

(**C**) Upper panel, putative amino acid sequence of the C-terminal part of FMRpolyA. The peptide sequence used to generate the 5FM mouse monoclonal antibody is underlined. Lower panel, immunoblotting validation of the 5FM antibody on lysate of HeLa cells transfected for 24 hours with a construct expressing the GFP fused to the putative C-terminal part of FMRpolyA.

(**D**) Immunofluorescence against the C-terminal part of FMRpolyA (red, 5FM antibody) and ubiquitin (green) on brain sections (hippocampal area) of FXTAS patients. Scale bars, 10 μm. Nuclei were counterstained with DAPI.

Figure S3

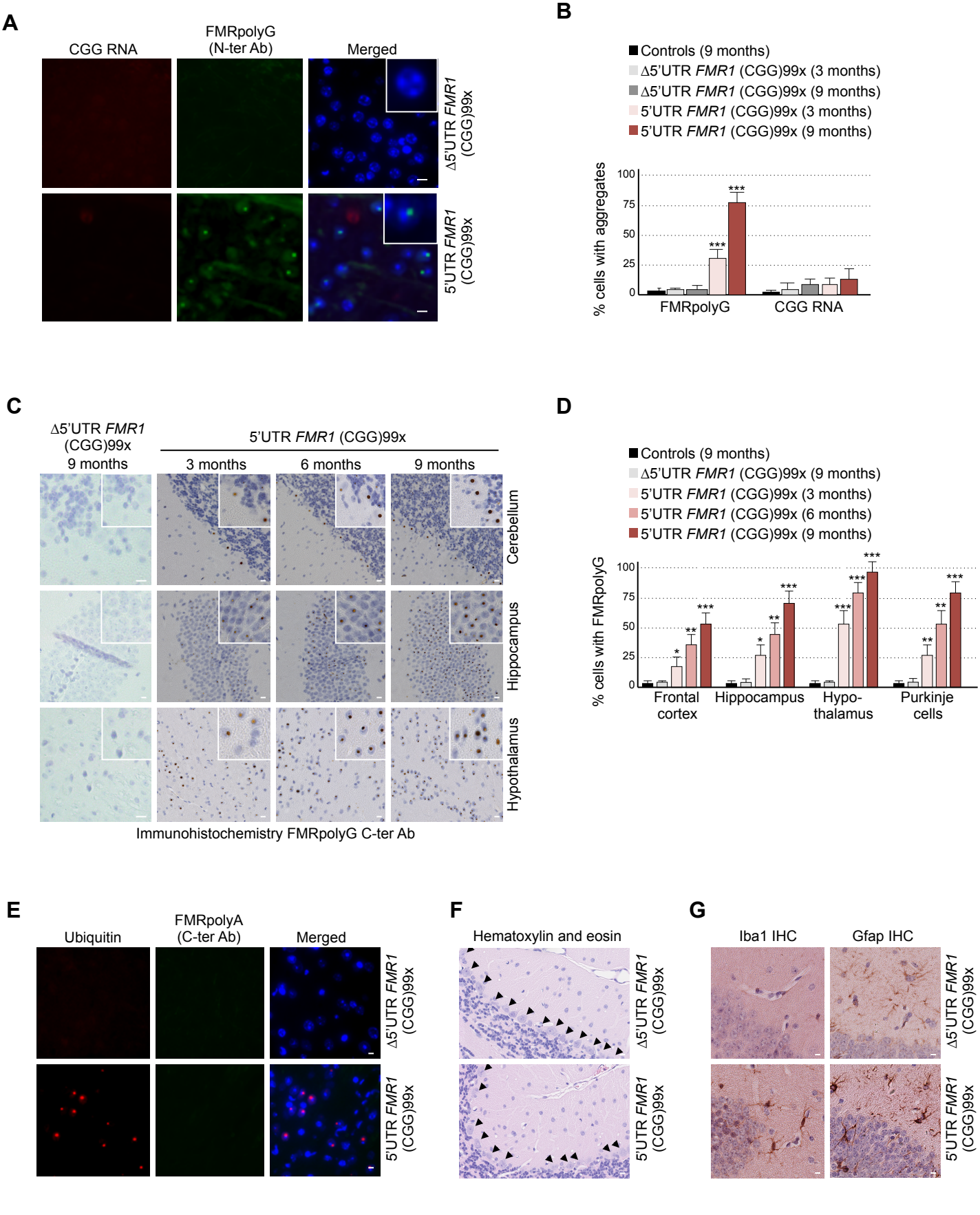

A
CGG RNA
FMRpolyG (N-ter Ab)
Merged
Δ5'UTR FMR1 (CGG)99x
5'UTR FMR1 (CGG)99x
B
Controls (9 months)
Δ5'UTR FMR1 (CGG)99x (3 months)
Δ5'UTR FMR1 (CGG)99x (9 months)
5'UTR FMR1 (CGG)99x (3 months)
5'UTR FMR1 (CGG)99x (9 months)
% cells with aggregates
100
75
50
25
0
***
FMRpolyG
CGG RNA
C
Δ5'UTR FMR1 (CGG)99x 9 months
5'UTR FMR1 (CGG)99x
3 months
6 months
9 months
Cerebellum
Hippocampus
Hypothalamus
Immunohistochemistry FMRpolyG C-ter Ab
D
Controls (9 months)
Δ5'UTR FMR1 (CGG)99x (9 months)
5'UTR FMR1 (CGG)99x (3 months)
5'UTR FMR1 (CGG)99x (6 months)
5'UTR FMR1 (CGG)99x (9 months)
% cells with FMRpolyG
100
75
50
25
0
*
**
***
Frontal cortex
Hippocampus
Hypo-thalamus
Purkinje cells
E
Ubiquitin
FMRpolyA (C-ter Ab)
Merged
Δ5'UTR FMR1 (CGG)99x
5'UTR FMR1 (CGG)99x
F
Hematoxylin and eosin
Δ5'UTR FMR1 (CGG)99x
5'UTR FMR1 (CGG)99x
G
Iba1 IHC
Gfap IHC
Δ5'UTR FMR1 (CGG)99x
5'UTR FMR1 (CGG)99x

Figure S3 (related to figure 3).
Expression of FMRpolyG is pathogenic in mice.

(**A**) RNA FISH against CGG RNA foci (red) coupled to immunofluorescence against FMRpolyG N-terminal part (8FM antibody, green) in brain of 6 months old bigenic CMV-cre/full-length or mutant FMR1 5'UTR transgenic mice. Nuclei were counterstained with DAPI.

(**B**) Quantification of FMRpolyG protein aggregates and CGG RNA foci in 3 or 9 months old control (n=3), bigenic CMV-cre/full-length (n=3) or mutant (n=3) FMR1 5'UTR transgenic mice (n=100 cells).

(**C**) Immunohistochemistry against FMRpolyG C-terminal part (9FM) of cerebellum, hippocampal and hypothalamic areas of 3, 6 or 9 months old bigenic CMV-cre/full-length or mutant FMR1 5'UTR transgenic mice. Sections were counterstained with Nissl staining.

(**D**) Quantification of FMRpolyG protein aggregates in control (n=3), bigenic CMV-cre/full-length (n=3) or mutant (n=3) FMR1 5'UTR transgenic mice.

(**E**) Immunofluorescence against ubiquitin (red) and the C-terminal part of FMRpolyA (5FM, green) on hippocampal areas of 6 months old bigenic CMV-cre/full-length or mutant FMR1 5'UTR transgenic mice. Nuclei were counterstained with DAPI.

(**F**) Hematoxylin and eosin staining of cerebellum of 9 months old bigenic CMV-cre/full-length or mutant FMR1 5'UTR transgenic mice. Arrowheads indicate Purkinje cells.

(**G**) Immunohistochemistry against Iba1 and Gfap of the hippocampal areas of 9 months old bigenic CMV-cre/full-length or mutant FMR1 5'UTR transgenic mice. Sections were counterstained with H&E staining. Scale bars, 10 µm. Error bars indicate s.e.m. Student t-test, * indicates $p<0.05$, ** indicates $p<0.01$, *** indicates $p<0.001$.

Figure S4

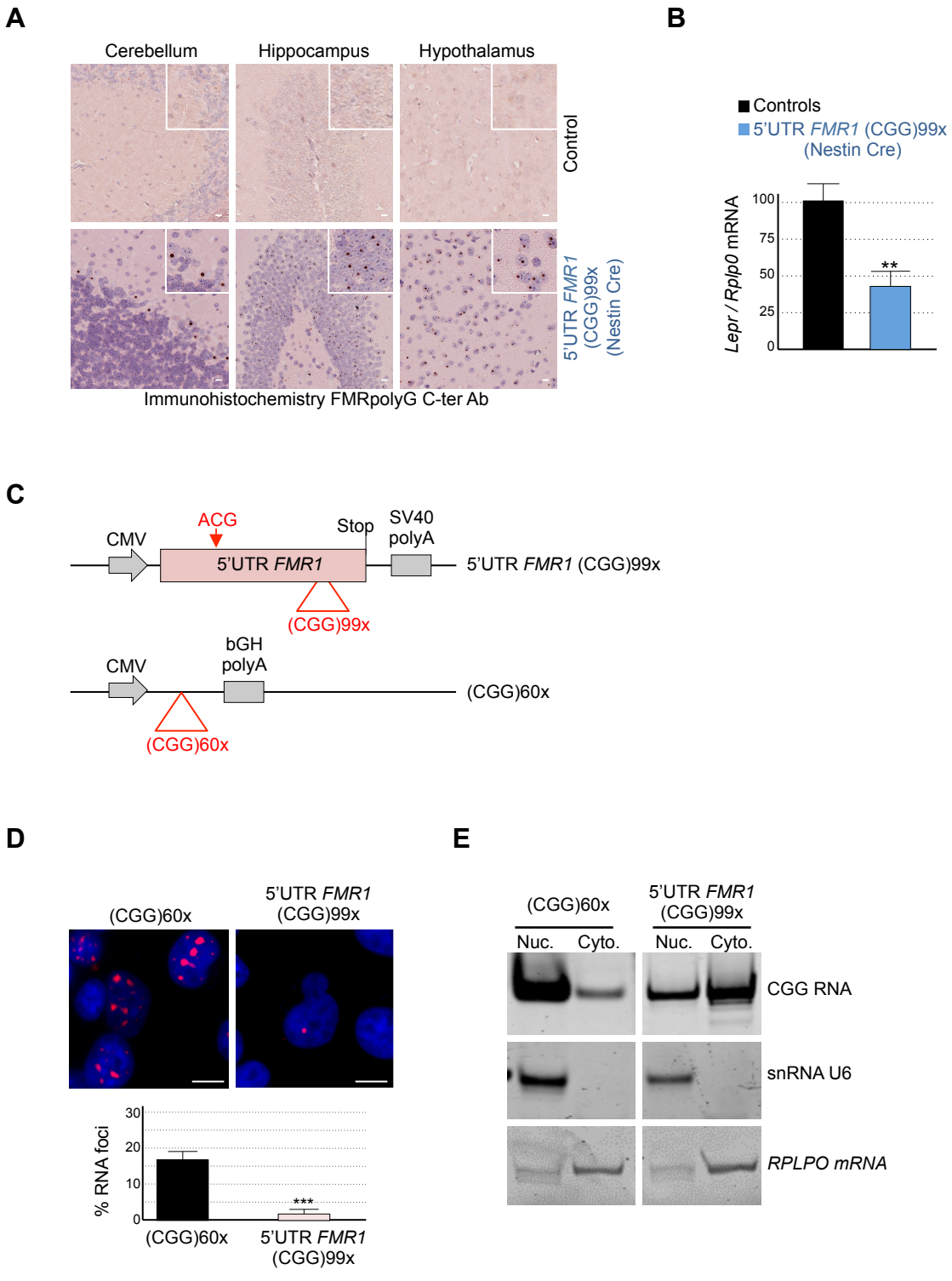

A
Cerebellum
Hippocampus
Hypothalamus
Control
5'UTR FMR1 (CGG)99x (Nestin Cre)
Immunohistochemistry FMRpolyG C-ter Ab
B
Controls
5'UTR FMR1 (CGG)99x (Nestin Cre)
Lepr / Rplp0 mRNA
100
75
50
25
0
**
C
CMV
ACG
5'UTR FMR1
Stop
SV40 polyA
5'UTR FMR1 (CGG)99x
(CGG)99x
CMV
bGH polyA
(CGG)60x
(CGG)60x
D
(CGG)60x
5'UTR FMR1 (CGG)99x
% RNA foci
30
20
10
0
***
(CGG)60x
5'UTR FMR1 (CGG)99x
E
(CGG)60x
5'UTR FMR1 (CGG)99x
Nuc.
Cyto.
Nuc.
Cyto.
CGG RNA
snRNA U6
RPLPO mRNA

Figure S4 (related to figure 4).
The 5'UTR sequence of FMR1 impairs formation of RNA foci.

(**A**) Immunohistochemistry against FMRpolyG C-terminal part (9FM antibody) in the cerebellum, hippocampus and hypothalamus of 6 months old bigenic Nestin-cre/full-length FMR1 5'UTR mice. Scale bars, 10 µm. Sections were counterstained with H&E staining.

(**B**) Quantitative RT-PCR analysis of Leptin receptor mRNA expression relative to the *RplpO* mRNA in the hypothalamic brain areas of 9 months old control (n=3) or Nestin-cre/full-length FMR1 5'UTR bigenic (n=3) mice.

(**C**) Schematic description of the expanded CGG repeats constructs embedded in the human 5'UTR sequence of FMR1 or cloned without any FMR1 sequence. The 99 CGG repeats embedded in their natural 5'UTR sequence is deposited at Addgene, #63089. The construct with 60 CGG repeats inserted into the multiple cloning site of pcDNA3.1 is described in Sellier et al, 2010.

(**D**) Upper panel, RNA FISH against CGG RNA foci in cells transfected for 24 hours with expanded CGG repeats embedded or not in the 5'UTR of FMR1 (untagged). Scale bars, 10 µm. Nuclei were counterstained with DAPI. Lower panel, quantification of cells (n=50) with CGG RNA foci (n=3 transfection). Error bars indicate s.e.m. Student t-test, *** indicates $p<0.001$.

(**E**) RT-PCR analysis of nuclear and cytoplasmic fractions of neuronal cells transfected for 24 hours with expanded CGG repeats embedded or not in the 5'UTR of FMR1 (untagged). Correct nuclear and cytoplasmic fractionation was controlled by RT-PCR for nuclear U6 snRNA and for cytoplasmic RPLP0 mRNA.

Figure S5

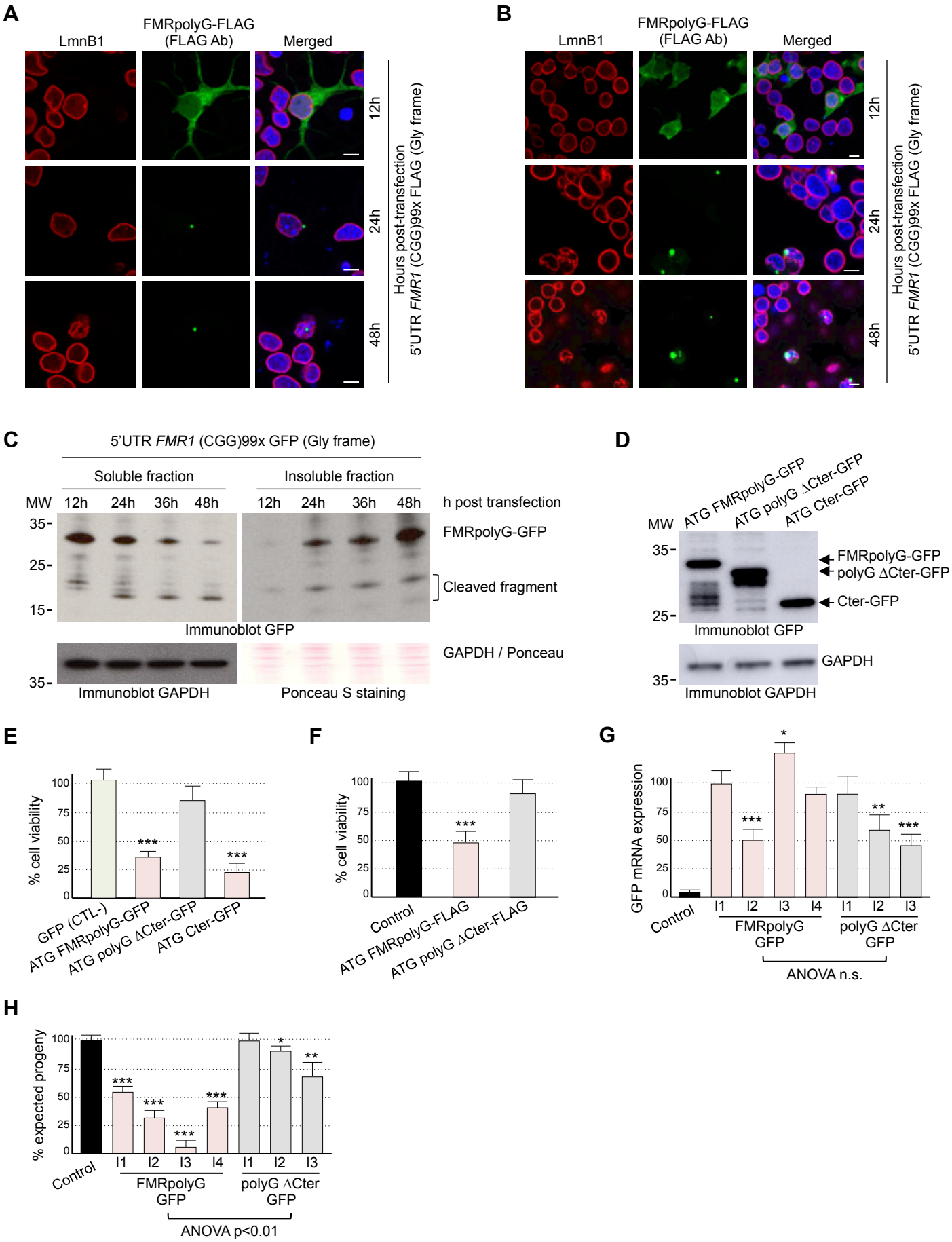

A
LmnB1
FMRpolyG-FLAG (FLAG Ab)
Merged
12h
24h
48h
Hours post-transfection
5'UTR FMR1 (CGG)99x FLAG (Gly frame)
B
LmnB1
FMRpolyG-FLAG (FLAG Ab)
Merged
12h
24h
48h
Hours post-transfection
5'UTR FMR1 (CGG)99x FLAG (Gly frame)
C
5'UTR FMR1 (CGG)99x GFP (Gly frame)
Soluble fraction
Insoluble fraction
MW
12h
24h
36h
48h
12h
24h
36h
48h
h post transfection
35-
25-
15-
FMRpolyG-GFP
Cleaved fragment
Immunoblot GFP
GAPDH / Ponceau
35-
Immunoblot GAPDH
Ponceau S staining
D
MW
ATG FMRpolyG-GFP
ATG polyG ΔCter-GFP
ATG Cter-GFP
35-
25-
FMRpolyG-GFP
polyG ΔCter-GFP
Cter-GFP
Immunoblot GFP
GAPDH
35-
Immunoblot GAPDH
E
% cell viability
100
75
50
25
0
***
***
GFP (CTL-)
ATG FMRpolyG-GFP
ATG polyG ΔCter-GFP
ATG Cter-GFP
F
% cell viability
100
75
50
25
0
***
Control
ATG FMRpolyG-FLAG
ATG polyG ΔCter-FLAG
G
GFP mRNA expression
100
75
50
25
0
*
***
**
***
Control
I1
I2
I3
I4
I1
I2
I3
FMRpolyG GFP
polyG ΔCter GFP
ANOVA n.s.
H
% expected progeny
100
75
50
25
0
***
***
***
***
*
**
Control
I1
I2
I3
I4
I1
I2
I3
FMRpolyG GFP
polyG ΔCter GFP
ANOVA p<0.01

Figure S5 (related to figure 5).
The C-terminal of FMRpolyG contributes to its toxicity.

(**A**) Immunofluorescence against the FLAG tag (green) and lamin B1 (Lmnb1, red) 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 a FLAG tag in the glycine frame. Scale bars, 10 μm. Nuclei were counterstained with DAPI.

(**B**) Immunofluorescence against the FLAG tag (green) and lamin B1 (Lmnb1, red) in HEK293 cells transfected for the indicated time period with expanded CGG repeats embedded within the 5'UTR of FMR1 and fused to a FLAG tag in the glycine frame. Scale bars, 10 μm. Nuclei were counterstained with DAPI.

(**C**) Immunoblotting against the GFP of the soluble and insoluble fractions of neuronal cells 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.

(**D**) Immunoblotting against the GFP of the soluble fractions of N2A cells transfected for 24 hours with ATG-driven FMRpolyG-GFP full-length or deleted of its N- or C-terminal part.

(**E**) Cell viability of N2A cells transfected for 24 hours with ATG-driven FMRpolyG-GFP full-length or deleted of its N- or C-terminal part (5 independent transfections).

(**F**) Cell viability of N2A cells transfected for 24 hours with ATG-driven FMRpolyG-FLAG full-length or deleted of its C-terminal part (3 independent transfections).

(**G**) Transgene mRNA expression in independent Drosophila lines expressing either FMRpolyG-GFP full-length or deleted of its C-terminus. RNA levels are significantly different from one line to another by student t-test. However, this is independent of the construct as there is no significant difference of transgene expression between FMRpolyG-GFP and polyG ΔCter-GFP lines by ANOVA. FMRpolyG-GFP line #2 and polyG ΔCter-GFP line #2 correspond to the Drosophila lines analyzed in figure 5C and present similar transgene expression.

(**H**) Progeny eclosion ratio (n=100, 3 independent crosses) of independent Drosophila lines expressing FMRpolyG full-length or deleted of its C-terminus compared to control driver line (Actin5C-Gal4/+). Fly lines expressing FMRpolyG-GFP present a decrease eclosion rate compared to Drosophila lines expressing polyG ΔCter-GFP (one way ANOVA, $P<0.01$). Error bars indicate s.e.m. Student t-test, * indicates $p<0.05$, ** indicates $p<0.01$, *** indicates $p<0.001$.

Figure S6

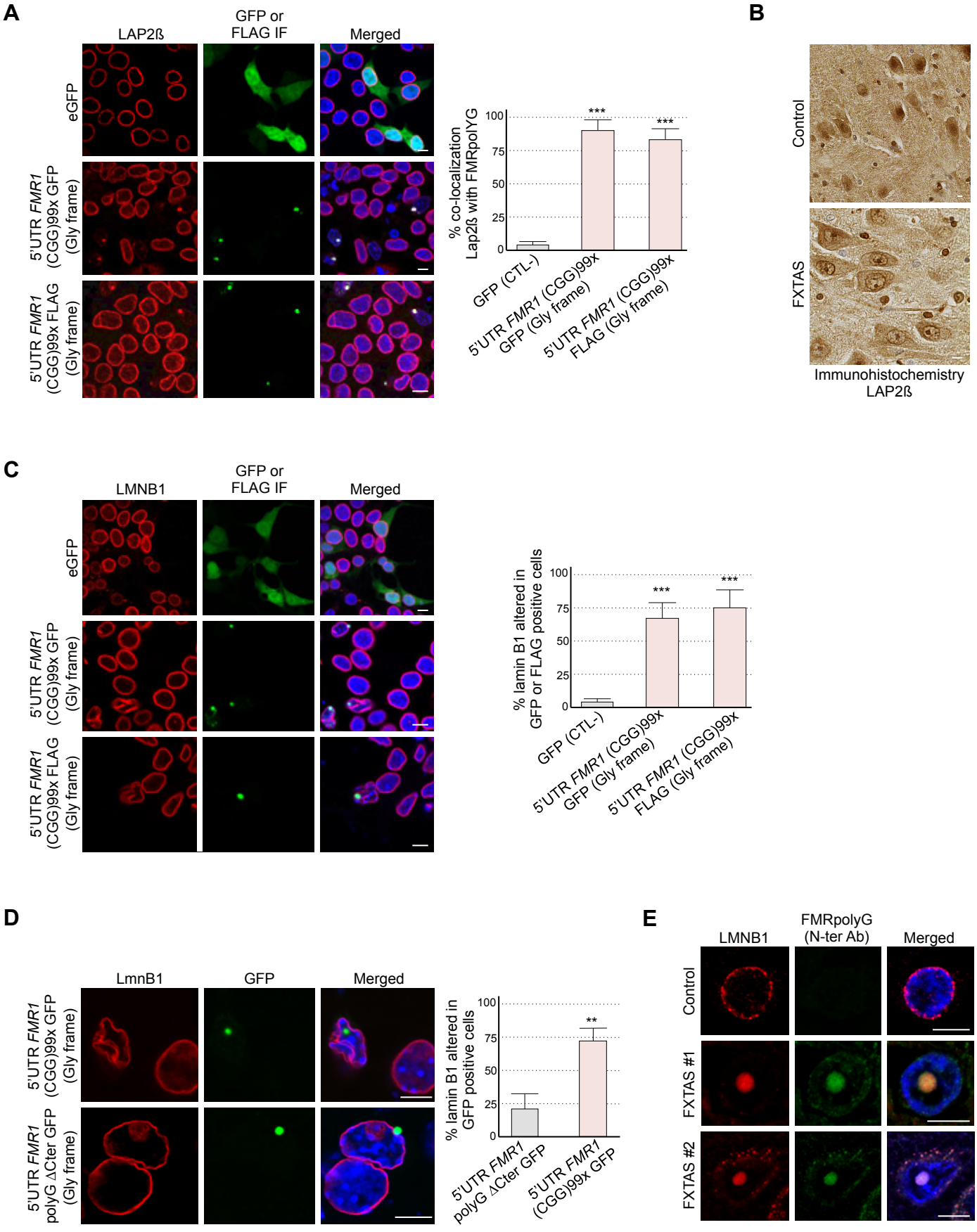

A
LAP2ß
GFP or FLAG IF
Merged
eGFP
5'UTR FMR1 (CGG)99x GFP (Gly frame)
5'UTR FMR1 (CGG)99x FLAG (Gly frame)
% co-localization Lap2ß with FMRpolyG
100
75
50
25
0
***
***
GFP (CTL-)
5'UTR FMR1 (CGG)99x GFP (Gly frame)
5'UTR FMR1 (CGG)99x FLAG (Gly frame)
B
Control
FXTAS
Immunohistochemistry LAP2ß
C
LMNB1
GFP or FLAG IF
Merged
eGFP
5'UTR FMR1 (CGG)99x GFP (Gly frame)
5'UTR FMR1 (CGG)99x FLAG (Gly frame)
% lamin B1 altered in GFP or FLAG positive cells
100
75
50
25
0
***
***
GFP (CTL-)
5'UTR FMR1 (CGG)99x GFP (Gly frame)
5'UTR FMR1 (CGG)99x FLAG (Gly frame)
D
LmnB1
GFP
Merged
5'UTR FMR1 (CGG)99x GFP (Gly frame)
5'UTR FMR1 polyG ΔCter GFP (Gly frame)
% lamin B1 altered in GFP positive cells
100
75
50
25
0
**
5'UTR FMR1 polyG ΔCter GFP
5'UTR FMR1 (CGG)99x GFP
E
LMNB1
FMRpolyG (N-ter Ab)
Merged
Control
FXTAS #1
FXTAS #2

Figure S6 (related to figure 6).
FMRpolyG alters LAP2ß and lamin B1 nuclear organization.

(**A**) Left panel, GFP fluorescence and immunofluorescence against the FLAG tag (green) and LAP2ß (red) in HEK293 cells transfected for 24 hours with GFP as control or with expanded CGG repeats embedded within the 5'UTR of FMR1 and fused either to the GFP or to a FLAG tag in the glycine frame. Right panel, quantification of LAP2ß co-localization with FMRpolyG (n=50 cells, 3 independent transfections).

(**B**) Immunohistochemistry against LAP2ß of hippocampal areas of FXTAS individual or age-matched control. Sections were counterstained with H&E staining.

(**C**) Left panel, GFP fluorescence and immunofluorescence against the FLAG tag (green) and lamin B1 (LMNB1, red) in HEK293 cells transfected for 24 hours with GFP as control or with expanded CGG repeats embedded within the 5'UTR of FMR1 and fused either to the GFP or to a FLAG tag in the glycine frame. Right panel, quantification of lamin B1 alterations in FMRpolyG positive cells (n=50 cells, 3 independent transfections).

(**D**) Left panel, GFP fluorescence and immunofluorescence against lamin B1 (LmnB1, red) in primary cultures of E18 mouse cortical neurons transfected with expanded CGG repeats embedded within the 5'UTR of FMR1 full-length or deleted of FMRpolyG C-terminus and fused to the GFP in the glycine frame. Right panel, quantification of lamin B1 alteration in GFP-positive transfected neurons (n=100 neurons, 3 independent transfections).

(**E**) Immunofluorescence against the N-terminal part of FMRpolyG (green, 8FM antibody) and lamin B1 (LMNB1, red) on brain sections (hippocampal area) of FXTAS patients or age-matched control individual. Scale bars, 10 µm. Nuclei were counterstained with DAPI. Error bars indicate s.e.m. Student t-test, ** indicates $p<0.01$, *** indicates $p<0.001$.

Figure S7

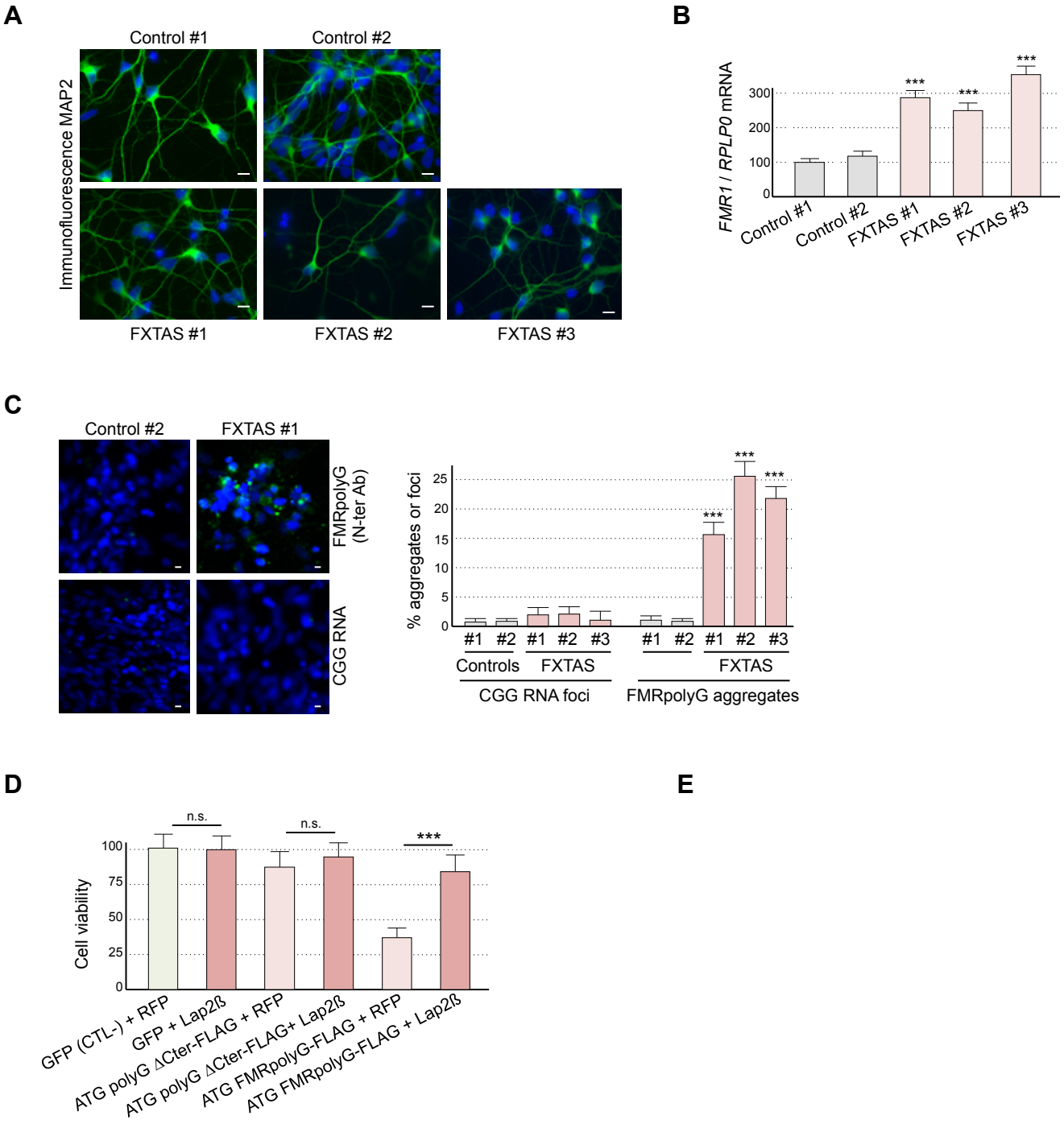

A
Control #1
Control #2
Immunofluorescence MAP2
FXTAS #1
FXTAS #2
FXTAS #3
B
FMR1 / RPLP0 mRNA
300
200
100
0
***
***
***
Control #1
Control #2
FXTAS #1
FXTAS #2
FXTAS #3
C
Control #2
FXTAS #1
FMRpolyG (N-ter Ab)
CGG RNA
% aggregates or foci
25
20
15
10
5
0
***
***
***
#1 #2 #1 #2 #3 #1 #2 #1 #2 #3
Controls
FXTAS
FXTAS
CGG RNA foci
FMRpolyG aggregates
D
Cell viability
100
75
50
25
0
n.s.
n.s.
***
GFP (CTL-) + RFP
GFP + Lap2ß
ATG polyG ΔCter-FLAG + RFP
ATG polyG ΔCter-FLAG+ Lap2ß
ATG FMRpolyG-FLAG + RFP
ATG FMRpolyG-FLAG + Lap2ß
E

Figure S7 (related to figure 7).
Neurons from iPS cell of FXTAS express FMRpolyG.

(**A**) Immunofluorescence against MAP2 on neuronal cultures differentiated 40 days from iPS cells of FXTAS or control individuals.

(**B**) *FMR1* mRNA expression relative to RPLP0 in neuronal cultures (n=3 independent cultures) differentiated 40 days from iPS cells of FXTAS or control individuals.

(**C**) Left panel, RNA FISH against expanded CGG repeats (red) coupled to immunofluorescence against FMRpolyG N-terminus (8FM antibody, green) in neurons differentiated 40 days from iPS cells originating from FXTAS or control individuals. Right panel, quantification of CGG nuclear RNA foci and FMRpolyG nuclear aggregates in neurons differentiated 40 days from iPS cells originating from FXTAS or control individuals (n=100 neurons, 3 independent cell cultures).

(**D**) Cell viability of neuronal N2A cells transfected with ATG-driven FMRpolyG-FLAG full-length or deleted of its C-terminus and with a plasmid expressing RFP as control or Ha-tagged LAP2ß (n=3 transfections). Scale bars, 10 µm. Nuclei were counterstained with DAPI. Error bars indicate s.e.m. Student t-test, *** indicates $p<0.001$.

Table S1

Accession	Description	Score	Coverage	# Peptides	# PSM	# AAs	MW [kDa]
Q61029-3	Lamina-associated polypeptide 2, isoforms beta/delta/epsilon/gamma OS=Mus musculus GN=Tmpo - [LAP2B_MOUSE]	48	16	13	15	412	46
P57780	Alpha-actinin-4 OS=Mus musculus GN=Actn4 PE=1 SV=1 - [ACTN4_MOUSE]	23	7	5	7	912	105
Q9CPW4	Actin-related protein 2/3 complex subunit 5 OS=Mus musculus GN=Arpc5 PE=2 SV=3 - [ARPC5_MOUSE]	20	31	3	4	151	16
Q5SW19-2	Isoform 2 of Clustered mitochondria protein homolog OS=Mus musculus GN=Cluh - [CLU_MOUSE]	17	8	5	5	818	92
Q61033	Lamina-associated polypeptide 2, isoforms alpha/zeta OS=Mus musculus GN=Tmpo PE=1 SV=4 - [LAP2A_MOUSE]	15	7	4	5	693	75
Q9WUI1	Mitogen-activated protein kinase 11 OS=Mus musculus GN=Mapk11 PE=1 SV=2 - [MK11_MOUSE]	15	10	3	5	364	41
Q9QUR6	Prolyl endopeptidase OS=Mus musculus GN=Prep PE=2 SV=1 - [PPCE_MOUSE]	7	4	2	2	710	81
P61290	Proteasome activator complex subunit 3 OS=Mus musculus GN=Psme3 PE=1 SV=1 - [PSME3_MOUSE]	7	4	1	2	254	29
Q08189	Protein-glutamine gamma-glutamyltransferase E OS=Mus musculus GN=Tgm3 PE=1 SV=2 - [TGM3_MOUSE]	7	2	1	2	693	77
O70250	Phosphoglycerate mutase 2 OS=Mus musculus GN=Pgam2 PE=2 SV=3 - [PGAM2_MOUSE]	6	10	2	2	253	29
P08032	Spectrin alpha chain, erythrocytic 1 OS=Mus musculus GN=Spta1 PE=2 SV=3 - [SPTA1_MOUSE]	6	1	2	2	2415	280
P08752	Guanine nucleotide-binding protein G(i) subunit alpha-2 OS=Mus musculus GN=Gnai2 PE=1 SV=5 - [GNAI2_MOUSE]	6	7	2	2	355	40
Q9R0Q6	Actin-related protein 2/3 complex subunit 1A OS=Mus musculus GN=Arpc1a PE=1 SV=1 - [ARC1A_MOUSE]	6	7	2	2	370	42
Q922D4-2	Serine/threonine-protein phosphatase 6 regulatory subunit 3 OS=Mus musculus GN=Ppp6r3 - [PP6R3_MOUSE]	6	2	1	2	827	93
Q8C6G8	WD repeat-containing protein 26 OS=Mus musculus GN=Wdr26 PE=2 SV=3 - [WDR26_MOUSE]	6	4	2	2	641	70
Q4KML4	Costars family protein ABRACL OS=Mus musculus GN=Abracl PE=3 SV=1 - [ABRAL_MOUSE]	6	16	1	2	81	9
P10853	Histone H2B type 1-F/J/L OS=Mus musculus GN=Hist1h2bf PE=1 SV=2 - [H2B1F_MOUSE]	5	7	1	2	126	14
Q8C0D7-4	Isoform 4 of Inhibitor of growth protein 4 OS=Mus musculus GN=Ing4 - [ING4_MOUSE]	5	8	1	2	166	19
P0C0S6	Histone H2A.Z OS=Mus musculus GN=H2afz PE=1 SV=2 - [H2AZ_MOUSE]	5	7	1	2	128	14
Q9D7X8	Gamma-glutamylcyclotransferase OS=Mus musculus GN=Ggct PE=1 SV=1 - [GGCT_MOUSE]	5	6	1	2	188	21
P04117	Fatty acid-binding protein, adipocyte OS=Mus musculus GN=Fabp4 PE=1 SV=3 - [FABP4_MOUSE]	4	9	1	1	132	15
Q8VBT0	Thioredoxin-related transmembrane protein 1 OS=Mus musculus GN=Tmx1 PE=1 SV=1 - [TMX1_MOUSE]	4	4	1	1	278	31
P15532	Nucleoside diphosphate kinase A OS=Mus musculus GN=Nme1 PE=1 SV=1 - [NDKA_MOUSE]	3	11	1	1	152	17
Q8BQP8	Rab11 family-interacting protein 4 OS=Mus musculus GN=Rab11fip4 PE=1 SV=1 - [RFIP4_MOUSE]	3	2	1	1	635	72
Q8BH74	Nuclear pore complex protein Nup107 OS=Mus musculus GN=Nup107 PE=2 SV=1 - [NU107_MOUSE]	3	1	1	1	926	107
O70456	14-3-3 protein sigma OS=Mus musculus GN=Sfn PE=1 SV=2 - [1433S_MOUSE]	3	4	1	1	248	28
O70435	Proteasome subunit alpha type-3 OS=Mus musculus GN=Psma3 PE=1 SV=3 - [PSA3_MOUSE]	3	5	1	1	255	28
O70166	Stathmin-3 OS=Mus musculus GN=Stmn3 PE=1 SV=1 - [STMN3_MOUSE]	3	9	1	1	180	21
Q9JM76	Actin-related protein 2/3 complex subunit 3 OS=Mus musculus GN=Arpc3 PE=1 SV=3 - [ARPC3_MOUSE]	3	6	1	1	178	21
P41969	ETS domain-containing protein Elk-1 OS=Mus musculus GN=Elk1 PE=2 SV=3 - [ELK1_MOUSE]	3	4	1	1	429	45
Q07643	Collagen alpha-2(IX) chain OS=Mus musculus GN=Col9a2 PE=2 SV=1 - [CO9A2_MOUSE]	3	2	1	1	688	65
P68510	14-3-3 protein eta OS=Mus musculus GN=Ywhah PE=1 SV=2 - [1433F_MOUSE]	3	6	1	1	246	28
O89094-2	Isoform 2 of Caspase-14 OS=Mus musculus GN=Casp14 - [CASPE_MOUSE]	3	5	1	1	159	18
Q61176	Arginase-1 OS=Mus musculus GN=Arg1 PE=1 SV=1 - [ARGI1_MOUSE]	3	3	1	1	323	35
Q91UZ4	Egl nine homolog 3 OS=Mus musculus GN=Egln3 PE=2 SV=1 - [EGLN3_MOUSE]	3	5	1	1	239	27
P97372	Proteasome activator complex subunit 2 OS=Mus musculus GN=Psme2 PE=2 SV=4 - [PSME2_MOUSE]	3	4	1	1	239	27
Q61941	NAD(P) transhydrogenase, mitochondrial OS=Mus musculus GN=Nnt PE=1 SV=2 - [NNTM_MOUSE]	3	1	1	1	1086	114
P50580-2	Isoform 2 of Proliferation-associated protein 2G4 OS=Mus musculus GN=Pa2g4 - [PA2G4_MOUSE]	3	3	1	1	340	38
Q64331	Unconventional myosin-VI OS=Mus musculus GN=Myo6 PE=1 SV=1 - [MYO6_MOUSE]	3	1	1	1	1265	146
Q91WD9	Secretagogin OS=Mus musculus GN=Scgn PE=2 SV=1 - [SEGN_MOUSE]	3	7	1	1	276	32
P31695	Neurogenic locus notch homolog protein 4 OS=Mus musculus GN=Notch4 PE=1 SV=2 - [NOTC4_MOUSE]	3	1	1	1	1964	207
O54818	Tumor protein D53 OS=Mus musculus GN=Tpd52l1 PE=2 SV=1 - [TPD53_MOUSE]	3	7	1	1	204	23
O35685	Nuclear migration protein nudC OS=Mus musculus GN=Nudc PE=1 SV=1 - [NUDC_MOUSE]	3	4	1	1	332	38
Q9CWP6-2	Isoform 2 of Motile sperm domain-containing protein 2 OS=Mus musculus GN=Mospd2 - [MSPD2_MOUSE]	3	3	1	1	481	55
Q7TSH3	Zinc finger protein 516 OS=Mus musculus GN=Znf516 PE=1 SV=1 - [ZN516_MOUSE]	3	2	1	1	1157	125
P30658	Chromobox protein homolog 2 OS=Mus musculus GN=Cbx2 PE=1 SV=2 - [CBX2_MOUSE]	3	3	1	1	519	55
P62305	Small nuclear ribonucleoprotein E OS=Mus musculus GN=Snrpe PE=3 SV=1 - [RUXE_MOUSE]	3	13	1	1	92	11
Q9WTX8	Mitotic spindle assembly checkpoint protein MAD1 OS=Mus musculus GN=Mad1l1 PE=2 SV=1 - [MD1L1_MOUSE]	3	2	1	1	717	83
Q791T5-2	Isoform 2 of Mitochondrial carrier homolog 1 OS=Mus musculus GN=Mtch1 - [MTCH1_MOUSE]	3	5	1	1	372	40
Q9CXF0	Kynureninase OS=Mus musculus GN=Kynu PE=2 SV=3 - [KYNU_MOUSE]	3	3	1	1	464	52
Q60770	Syntaxin-binding protein 3 OS=Mus musculus GN=Stxbp3 PE=1 SV=1 - [STXB3_MOUSE]	3	2	1	1	592	68
Q61687	Transcriptional regulator ATRX OS=Mus musculus GN=Atrx PE=1 SV=3 - [ATRX_MOUSE]	3	1	1	1	2476	278
Q8BKT8-2	Isoform 2 of HAUS augmin-like complex subunit 7 OS=Mus musculus GN=Haus7 - [HAUS7_MOUSE]	3	4	1	1	248	28
Q62407-2	Isoform 2 of Striated muscle-specific serine/threonine-protein kinase OS=Mus musculus GN=Speg - [SPEG_MOUSE]	3	2	1	1	861	94
O88508	DNA (cytosine-5)-methyltransferase 3A OS=Mus musculus GN=Dnmt3a PE=1 SV=2 - [DNM3A_MOUSE]	3	1	1	1	908	102
O54774	AP-3 complex subunit delta-1 OS=Mus musculus GN=Ap3d1 PE=1 SV=1 - [AP3D1_MOUSE]	3	1	1	1	1199	135
Q9WUM3	Coronin-1B OS=Mus musculus GN=Coro1b PE=1 SV=1 - [COR1B_MOUSE]	3	2	1	1	484	54
Q8BQM8	Echinoderm microtubule-associated protein-like 5 OS=Mus musculus GN=Eml5 PE=2 SV=2 - [EMAL5_MOUSE]	3	1	1	1	1977	220
Q9JIT1	RING finger protein 32 OS=Mus musculus GN=Rnf32 PE=2 SV=1 - [RNF32_MOUSE]	3	4	1	1	368	42
Q925I7-2	Isoform 2 of Platelet-derived growth factor D OS=Mus musculus GN=Pdgfd - [PDGFD_MOUSE]	3	3	1	1	364	42
Q5DU28-3	Isoform 3 of Pecanex-like protein 2 OS=Mus musculus GN=Pcnxl2 - [PCX2_MOUSE]	3	1	1	1	1121	121
P28741	Kinesin-like protein KIF3A OS=Mus musculus GN=Kif3a PE=1 SV=2 - [KIF3A_MOUSE]	3	1	1	1	701	80
Q99PL5-5	Isoform RRp16.8 of Ribosome-binding protein 1 OS=Mus musculus GN=Rrbp1 - [RRBP1_MOUSE]	3	3	1	1	398	42
Q6NZL8-3	Signal peptide, CUB and EGF-like domain-containing protein 1 OS=Mus musculus GN=Scube1 - [SCUB1_MOUSE]	3	1	1	1	961	104
A2AM05-2	Isoform 2 of Centlein OS=Mus musculus GN=Cntln - [CNTLN_MOUSE]	3	3	1	1	318	36
Q9D952	Envoplakin OS=Mus musculus GN=Evpl PE=2 SV=3 - [EVPL_MOUSE]	3	0	1	1	2035	232

Table S1 (related to figure 6).
Proteins interacting with FMRpolyG.

Proteins associated with HA-FLAG-tagged FMRpolyG expressed in mouse N2A neuronal cells were captured through consecutive anti-FLAG and anti-HA affinity purification steps and identified by orbitrap ion trap mass analyzer.

Supplemental videos 1 and 2 (related to figure 3).
Mice expressing FMRpolyG present gait instability.

Video recording of ledge test of 3 months old bigenic CMV-cre/full-length or mutant FMR1 5'UTR transgenic mice indicates that FMRpolyG-expressing mice present some gait instability and increase foot slippage compared to CGG RNA only-expressing mice.

Supplemental videos 3 to 6 (related to figure 5).
Time laps recoding of neurons expressing FMRpolyG-GFP.

24 hours recording of primary cultures of E18 mouse cortical neurons transfected with RFP and with control GFP or FMRpolyG-GFP full-length or deleted of its N- or C-terminus. Expression of either FMRpolyG-GFP or its C-terminus fused to the GFP leads to neuronal cell death. In contrast, expression of the polyglycine stretch deleted of FMRpolyG C-terminal part leads to aggregates formation with no overt cell toxicity.

Supplemental Experimental Procedures

Fibroblasts were successfully reprogrammed using retroviruses expressing Oct4, Sox2, Nanog and Lin28 (Jung et al., 2014). Briefly, primary dermal fibroblasts were maintained on gelatin-coated dishes in DMEM 1 g/l glucose with antibiotic, antimycotic and 0.1 mM non-essential amino acids (Invitrogen) and 10% FBS for 5 passages. On day 1, 1x10exp5 fibroblasts were transduced by lentivirus carrying cDNAs of Oct4, Sox2, Nanog and Lin28 with 8 mg/ml of polybrene (Sigma). On day 2, medium was replaced with fresh medium and on day 3, infected cells were transferred onto a 100-mm dish containing 1x10exp6 feeder cells (passage 3 mitomycin-C treated mouse embryonic fibroblasts). From day 6 to 9, fibroblast medium was progressively switched to human induced pluripotent stem cell medium (KO-DMEM, 20% KOSR, 2 mM L-glutamine, 0.1 mM non-essential amino acids, Penicillin-Streptomycin, 0.1 mMb-mercaptoethanol supplemented with 10 ng/ml of bFGF (R&D Systems). Human iPSC clones were picked at week 4 and expanded on matrigel coated 35 mm dishes (BD Biosciences) in mTeSR1 medium (Stemcell Technologies). For embryoid body formation, hiPSC were dissociated with dispase solution (1 mg/ml, Stemcell Technologies), resuspended in 1 ml of Aggrewell medium (Stemcell Technologies) containing 2 mM Y27632 (Stemgent), centrifuged in Aggrewell plates for 3 min at 80g and further incubated at 37°C for 24 h. The next day, embryoid bodies were transferred in 3 ml of Aggrewell medium. The following days, medium was progressively switched to KO-DMEM, 20% FBS, 2 mM L-glutamine, 0.1 mM Non-Essential Amino Acids, penicillin–streptomycin and after 30 days, embryoid bodies were collected. Of interest, expanded CGG repeats were stable with no contraction or expansion in iPS clones compared to fibroblasts. karyotype analyses were normal for all cell lines and retroviral silencing as well as stemness and pluripotency were confirmed by classic RT-qPCR and teratoma assays. For karyotype analysis, hiPSC cells were treated with colchicine (Sigma) for 4 h and cells were shocked with hypotonic KCl 0.075 M solution for 20 min at 37°C. Cells were fixed in methanol:acetic acid solution (3:1) and conventional cytogenetic was performed applying RHG banding on metaphase chromosomes. Expression of pluripotent surface markers of hIPSC was analyzed by FACS using anti-Tra-I-60, anti-Tra-I-81 and anti-SSEA4 antibodies (Millipore). Expression of 90 validated genes associated with stem cell pluripotency and differentiation to all three germ layers of hiPSC and corresponding embryoid bodies were analyzed using the Human Stem Cell Pluripotency Array (Applied Biosystems) according to manufacturer instructions. For in vivo teratoma formation, cells from one Matrigel coated 60 mm-dish were collected by dispase treatment and resuspended in 75 ml of KO-DMEM, mixed with 75 ml of Matrigel (BD Biosciences) and the hiPSC-Matrigel mixture was injected subcutaneously in 8-week-old NOD/SCID female mice (Charles River Laboratory, 2 mices injected for each hiPSC clone). After two to three months, teratomas were dissected and fixed in formalin, embedded in paraffin and processed with hematoxylin and eosin staining at the histology laboratory of the Institute Clinique de la Souris (ICS).

Differentiation of iPSC in neurons was performed according Marteyn et al., 2011 and Boissart et al., 2013. Biefly, for differentiation in Neuronal Stem cell (NSC), one B6 dish of IPS (60-80 % confluence) was washed with NFS medium (N2B27 supplemented with FGF2 5 ng/ml (Peprotech, AF-100-18B), hNoggin 260 ng/ml (R&D120-10C), SB431542 8,7 µg/ml (Tocris, 1614) containing ROCK inhibitor 3,5 µg/ml (Y-27632, Calbiochem, 688000) before to cut clumps. Clumps were collected and incubated overnight at 37°C in B3 UltraLow Attachment Dish (Corning). The next day, the clumps were transferred to a dish pre-coated with poly-ornithine 0,1% (Sigma, p4957) and laminine at 1 mg/ml (Sigma, L2020) and maintained with medium. After 24 hours, the medium was changed to NFS medium without rock inhibitor and medium was changed every two days during 8-10 days. After the appearance of neural rosette, the medium was replaced with NSC (N2B27 supplemented with FGF2 10 ng/ml, EGF 10 ng/ml (R&D, 263-EG-200RD), hBDNF 20 ng/ml (R&D, 248-BD-025CF). At confluence, cells were passaged 1:3 in NSC medium. For differentiation of NSC in neurons, confluent cells were dissociated with trypsin and plated on pre-coated with polyornithine (Sigma) and laminin (Sigma) in 24-well plate (50000 cells/well) in neuron medium (N2B27 supplemented with hBDNF 20 ng/ml and laminine 2 µg/ml). Media change was performed every 2 days and cells were differentiated during 55 days.

Mouse models, genotyping and phenotyping.

All mouse procedures were done according protocols approved by the Committee on Animal Resources of the ICS animal facility and under the French and European authority guidelines.

For transgene construction, human 5'UTR FMR1 fused in the glycine frame to the GFP (Addgene) was cloned between the FseI and SmaI sites of the Rosa26 5' arm – CAG promoter – LOXP - SV40 polyA 3x - LOXP – Rosa26 3' arm vector Ai2 (Addgene). Deleted transgene for the ACG near-cognate codon was constructed by deleting 100 nucleotides from FseI to KasI sites within the FMR1 5'UTR plasmid. Both constructs are driven by the strong chimeric ubiquitous CAG promoter, and inserted by homologous recombination within the neutral Rosa26 mouse locus. To control expression of the transgenes, three upstream SV40-polyadeneylation sites bordered by loxP sites limit transcription of the expanded CGG repeats. Hence, expression and potential pathogenicity of the expanded CGG repeats is permitted only in offspring of the transgenic CGG mice crossed with mice expressing Cre recombinase.

Both mutant mouse lines were established at the ICS (Mouse Clinical Institute; http://www-mci.u-strasbg.fr). Both linearized constructs were electroporated separately in C57Bl/l6N mouse embryonic stem (ES) cells. After G418 selection, targeted clones were identified by long-range PCR using external primers and further confirmed by Southern blot with an internal probe against Neomycin and 5' external probe against Rosa26. We confirmed by southern blot and PCR the correct insertion of the transgenes into the Rosa26 locus, the absence of concatamerization at the locus and the presence of 99 CGG repeats, which were stably transmitted with no obvious contraction or expansion. Two positive ES clones for each future transgenic mice were injected into Balb/cN blastocysts. Resulting male chimeras were bred with wild type C57Bl/6N females to obtain germline transmission. Deletion of the floxed STOP cassette were performed by breeding F1 males with CMV-cre deleter females (Birling et al., 2012) or Nestin-cre delete mice (Isaka et al., 1999). Genotyping across the expanded CGG repeats was performed using the Expand High Fidelity PCR System (Roche, 11-732 -650 001) according manufacturer instructions and supplemented with 2,5 M Betaine (B0300 Sigma, 12,5 µl of 5 M Betaine for a final PCR volume of 25 µl) with one denaturation step at 94 °C for 2 min, 30 cycles of amplification 94 °C for 1 min, 60 °C for 1 min, 72 °C for 2 min and a final step at 72 °C for 5 min using the forward primer 5'-TCGACCTGCAGCCCAAGCTAGATCG and the reverse primer 5'-TCCTTGAAGAAGATGGTGCGCTCC. Rotarod test (Bioseb, Chaville, France) was performed with three testing trials during which the rotation speed accelerated from 4 to 40 rpm in 5 min. Trials were separated by 5-10 min interval. The average latency was used as index of motor coordination performance. Grip test: this test measures the maximal muscle strength (g) using an isometric dynamometer connected to a grid (Bioseb). Mice were allowed to grip the grid with all its paws then they were pulled backwards until they released it. Each mouse was submitted to 3 consecutive trials immediately after the modified SHIRPA procedure. The maximal strength developed by the mouse before releasing the grid was recorded and the average value of the three trials adjusted to body weight. The string test consisted of 3 trials separated by 5-10 min interval. On each trial the forepaws of the animal were placed on the thread that is a wire stretched horizontally 40 cm above the bench. The latency the animal took to catch the wire with its hindpaws was recorded. Open field test: mice were tested in automated open fields (Panlab, Barcelona, Spain), each virtually divided into central and peripheral regions. The open fields were placed in a room homogeneously illuminated at 150 Lux. Each mouse was placed in the periphery of the open field and allowed to explore freely the apparatus for 30 min, with the experimenter out of the animal's sight. The distance traveled, the number of rears, and time spent in the central and peripheral regions were recorded over the test session. The number of entries and the percent time spent in center area are used as index of emotionality/anxiety

***Drosophila* models of FMRpolyG.**

All Drosophila lines were maintained on standard culture and food conditions at 25oC, while all crosses and experiments were performed at 29oC. Control and driver lines used in this study are w1118 (control) from Bloomington, Actin5C-GAL4/CyO driver (ubiquitous driver line) as a gift from Zhe Han's lab, and RU486-inducible Geneswitch tubulin driver line Tub5-GAL4 (ubiquitous expression) as a gift from Scott Pletcher's lab. DNA fragments containing FMRpolyG-GFP or polyG-GFP without the C-terminal sequence were PCR amplified from counterpart of mammalian transfection vectors described elsewhere and inserted to a pUAST vector between EcoRI and XbaI sites. All constructs were sequence verified by Sanger sequencing and transgenic flies with these constructs were made via standard p-element insertion (BestGene, CA). Transgene expression levels of GFP gene were analyzed 3 days after induction with RU486 in flies from the individual lines crossed to the Tub5-GAL4 driver, and those with comparable RNA expression levels were used for this study. The fly eclosion assay has been described elsewhere (Todd et al, 2013). Briefly, homozygous UAS transgenic lines or control lines were crossed to a Actin5C-GAL4 ubiquitous driver line balanced over a marker chromosome (CyO), on standard food at 29oC, if the transgene elicited no toxicity, then the number of progeny bearing the GAL4 driver would be expected to be equivalent to those bearing the CyO marker. Over 100 flies of each genotype were scored from multiple crosses. The ratio of expected progeny carrying the transgene compared to those carrying the CyO marker was expressed as a percentage of the expected ratio of one. These percentages were then compared using a Fischer exact test to determine statistical significance. For the fly survival assay, The UAS transgenic lines or control lines were crossed to Tub5-Gal4 geneswitch driver flies on standard food absent of RU486 at 29oC. Adult offspring of the desired genotypes were collected 2-3 days after eclosion and transferred to standard fly food containing 200 µM RU486 without yeast granules. The flies were transferred to fresh food with drug every 2-3 days. Each genotype started with at least 4 vials of 25 flies/vial and the survival was determined daily or every other day for 3 weeks.

Monoclonal antibody production.

To generate monoclonal antibodies directed against FMRpolyG or FMRpolyA, 8 week old female BALB/c mice were injected intraperitoneally with KLH conjugated peptides (FMRpolyA C-ter 5FM: PRAPAAHLSGAGSRR, FMRpolyG N-ter 8FM: MEAPLPGGVRQRG or FMRpolyG C-ter 9FM: GGWASSARSPPLGGGLPALA) with 200ug of poly(I/C) as adjuvant. Three injections were performed at 2 weeks intervals and four days prior to hybridoma fusion, mice with positively reacting sera were re-injected. Spleen cells were fused with Sp2/0.Agl4 myeloma cells. Supernatants of hybridoma cultures were tested at day 10 by ELISA for cross-reaction with peptides. Positive supernatants were then tested by Immunofluorescence and western blot on transfected HeLa cells. Specific cultures were cloned twice on soft agar. Specific hybridomas were established and ascites fluid was prepared by injection of 2x106 hybridoma cells into Freund adjuvant-primed BALB/c mice. All animal experimental procedures were performed according to the French and European authority guidelines.

Cell cultures, viability assays and transfections.

Neuro2A, HEK293 or HeLa cells were plated in 6 well tissue culture plates in DMEM 1 g/l glucose with 5% FCS and gentamycin. After 24 hours, cells were transfected with plasmid DNA using Lipofectamine 2000 (Invitrogen) according to manufacturer instructions. Primary cortical neurons were prepared from C57Bl/6 mice embryos of day E18 and grown on polylysine coated 24-well plates in Neurobasal Medium (NBM) supplemented with 1xB27, 0.5 mM L-glutamine and 100 IU/ml penicillin/streptomycin at 37° C with 5% CO2. Neurons were transfected at day 3 with Lipofectamine 2000 (Invitrogen) in 400 µl NBM. Medium was replaced after 3h with a 1:1 (v:v) mixture of conditioned and fresh NBM. For cell viability, cells were detached by scraping and resuspended in PBS. TO-PRO-3 iodide (Fisher scientific, T-3605) was added at 20 nM to each sample and gently mix just prior to analysis on the FACS and 30 000 cells were counted.

Western blotting.
For the small FMRpolyG-FLAG tagged or endogenous FMRpolyG (<15kDa), 20 to 50 µg of proteins were resolved by 12% bis-Tris Gel (NuPAGE) and transferred onto PVDF 0,2 µm membrane. The membrane was blocked with 5% non-fat dry milk in TBS-Tween 1% and incubated with FLAG (rabbit PA1-984B), 8FM or 9FM antibody (1:100) overnight at 4°C. Membrane was washed 3 times and incubated with secondary peroxidase antibody (1:3000, Cell Signaling) 1 hour in TBS-Tween 1%, followed by washing and ECL chemoluminescence revelation (Amersham ECL Prime). Concerning FMRpolyG-GFP tagged (>30kDa) and other large proteins analyzes, 20 µg of proteins were homogenized in 1x laemmli sample loading buffer, denatured 3 min at 95°C, separated on 4-12% bis-Tris Gel (NuPAGE), transferred on nitrocellulose membranes (Whatman Protan), blocked with 5% non-fat dry milk in TBS-Tween 1% (Tris Buffer Saline buffer), incubated with anti-FMRpolyG (8FM or 9 FM, 1/100), Lap2b (BD Biosciences 611000), GFP (Abcam ab290), GAPDH (ab125247, Abcam), HA (ThermoFisher Scientific 26183) in TBS-Tween 1%, washed 3 times and incubated with anti-rabbit or mouse Peroxidase antibody (1:3000, Cell Signaling) 1 hour in TBS-Tween 1%, followed by washing and ECL chemoluminescence revelation (Amersham ECL Prime). Concerning human brain tissue preparation, small pieces of lyophilized frozen brain tissue were homogenized in 100 µl of Tris-SDS buffer (100 mM Tris pH 9, 5 % SDS 20%, 5 % β-mercaptoethanol), boiled at 100°C during 5 min then centrifuged at 13 000 rpm for 20 min at 4 °C. The supernatant was removed. The pellet was washed twice with water and homogenized in 20 µl of formic acid and incubated at 37°C during 30 min. Next, the homogenate was dried in speed-vac and resuspended in 40 µl Laemmli loading buffer prior to western blot analysis.

Lysostaphin treatment
3x10exp5 transfected HeLa cells were scrapped in PBS 1X and centrifuged during 10 min at 3000 rpm at 4°C. The pellet was resuspended in 800 µl of RIPA. 16 µl of cell extract was incubated with 1 µg of lysostaphin (Prospec, ENZ-269) during 10 to 30 minutes at 37°C. Laemmli buffer was add to the mix and proteins were analyze by western blot.

Immunofluorescence and immunohistochemistry.
Mouse or human brain sections were deparaffinized two times for 20 min in Histosol Plus (Shandon) and dehydrated as follows: twice in ethanol 100% (5 min), twice in ethanol 95% (5 min), once in ethanol 80% (5 min), once in ethanol 70% (5 min) and rinsed in PBS. Glass coverslips containing plated cells or brain sections treated as described above were fixed in PFA during 10 min and washed three times with PBS. The coverslips or slides were incubated for 10 min in PBS plus 0.5% Triton X-100 and washed three times with PBS before incubation during 1 hours with primary antibody against FMRpolyA (5FM, 1/100), FMRpolyG (8FM or 9FM, 1/50 to 1/100), ubiquitin (DAKO, Z0458), GFP (Abcam ab1218), Lap2 (Millipore 06-1002; Abcam ab185718; Abcam ab189993), Lamin B1 (Abcam ab16048). Slides or coverslips were washed twice with PBS before incubation with a goat anti-rabbit or goat anti-mouse secondary antibody conjugated with Cyanine 3 (1/500 dilution; Fisher) for 60 min; incubated for 2 min in PBS 1X-DAPI (1/10 000 dilution) and rinsed twice with PBS 1X before mounting in Pro-Long media (Molecular Probes). Slides were examined using a fluorescence microscope (Leica). For immunochemistry, brain sections were deparaffinized followed by antigen retrieval using microwave treatment in 0.01 m sodium citrate and treatment with 10 µg/ml protein kinase for 20 min at 37°c. Endogenous peroxidase activity was blocked, and immunostaining was performed overnight at 4°C using antibody against Iba1 (Abcam ab15690), Gfap (Abcam ab7260), ubiquitin (Dako Z0458; 1:250), FMRpolyG (8FM or 9FM, 1:10 to 1/50) or Lap2 (Millipore 06-1002; Abcam ab185718; Abcam ab189993). Antigen–antibody complexes were visualized by incubation with DAB substrate (Dako) and slides were counterstained with hematoxylin and eosin.

RNA FISH coupled to immunofluorescence.
Mouse brain sections were deparaffinized and rehydrated. Coverslips containing primary culture of E18 mouse cortical neurons cells or brain sections were fixed in PFA during 10 min and washed three times with PBS. The coverslips or slides were incubated for 10 min in PBS plus 0.5% Triton X-100 and washed three times with PBS before pre-hybridization in 40% DMSO, 40% formamide, 10% BSA (10 mg/ml), 2 × SCC for 30 min. The coverslips or slides were hybridized for 2 h in 40% formamide, 10% DMSO, 2 × SCC, 2 mM vanadyl ribonucleoside, 60 µg/ml tRNA, 30 µg/ml BSA plus 0.75 µg (CCG)8x-Cy3 DNA oligonucleotide probe (Sigma). The coverslips or slides were washed twice in 2 × SCC/50% formamide and twice in 2 × SCC. The coverslips were incubated for 2 min in 2 × SCC/DAPI (1/10 000 dilution) and rinsed twice in 2 × SSC before mounting in Pro-Long media (Molecular Probes). Coverslips were examined using a fluorescence microscope (Leica). For FISH followed by immunofluorescence, after 2 × SCC wash, the slide were washed twice in PBS 1X. The slides were incubated 1 hour with primary antibody against FMRpolyGly antibody (8FM, 1/50). Slides were washed twice with PBS before incubation with a goat anti-mouse secondary antibody conjugated with cyanine-3 (1/500 dilution; Fisher) for 60 min; incubated for 2 min in PBS 1X-DAPI (1/10 000 dilution) and rinsed twice in PBS 1X before mounting in Pro-Long media (Molecular Probes). Slides were examined using a fluorescence microscope (Leica).

Mass spectrometry analysis of FMRpolyG interactant and FMRpolyG N-terminus.
5x10exp6 HeLa cell were transfected with 18 µg of HA-FLAG tagged plasmid using Fugen HD (Promega) for 24h hours. Proteins were purified by HA-FLAG tandem purification kit according to the manufacturer's instruction (Sigma-Aldrich), separated on 4-12% bis-Tris Gel (NuPAGE) and visualized by silver staining (SilverQuest, Invitrogen). Gel bands were excised and subjected to manual in-gel reduction in 10 mM DTT in 100 mM NH4HCO3 (Sigma Aldrich) for 1 h at 57°C, alkylated for 45 min in the dark with 55 mM iodoacetamide in 100 mM NH4HCO3 (Sigma Aldrich), washed in 25 mM NH4HCO3, dehydrated with acetonitrile and dried in SpeedVac 5301 Concentrator (Eppendorf). Then the gel pieces were rehydrated with 12.5 ng/µL trypsin or LysC solution (Promega) in 50 mM NH4HCO3 and incubated overnight at 37°C. The peptides were extracted twice with acetonitrile/water/formic acid-45/45/10-v/v/v followed by a final extraction with acetonitrile /formic acid (FA)-95/05-v/v. Extracted peptides were then analyzed using an Ultimate 3000 nano-RSLC (Thermo Scientific) coupled in line with an Orbitrap ELITE (Thermo Scientific). Peptides were separated on a C18 nano-column with a linear gradient of acetonitrile and analyzed with in a Top 20 CID (Collision-induced dissociation) data-dependent mass spectrometry with an inclusion list. Data were processed by database searching using SequestHT (Thermo Fisher Scientific) with Proteome Discoverer 1.4 software (Thermo Fisher Scientific) against a homemade database of all potential three frames translated proteins or peptides from the 5'UTR of FMR1. Precursor and fragment mass tolerance were set at 7 ppm and 0.5 Da respectively. Oxidation (M) and Nterminal Acetylation were set as variable modification, and Carbamidomethylation (C) as fixed modification. Peptides were filtered with the Fixed value node of Proteome Discoverer 1.4. Similarly, for identification of protein interactants, 5x10exp6 Neuro2A cell were transfected with 15 µg of FMRpolyG FLAG-HA double-tagged plasmid using Fugen HD (Promega) for 24h hours. Proteins were purified by HA-FLAG tandem purification kit according to the manufacturer's instruction (Sigma-Aldrich). Proteins were visualized by silver staining (SilverQuest, Invitrogen) after separation of 4-12% bis-Tris Gel (NuPAGE) and identified using NanoESI_Ion Trap (LTQ XL Thermo Fisher).

Constructs.
PCMV6 containing C-terminally FLAG-tagged human cDNAs of LAP2ß was purchased from OriGene. Plasmids containing the 5'UTR of human FMR1 fused in the glycine frame with the FLAG tag or GFP are deposited at Addgene (Plasmid #63089, #63090 and #63091). The FLAG tag is fused to the C-terminal end of FMRpolyG with a two amino acids (glycine valine) linker, and the GFP tag is fused to the C-terminus of FMRpolyG with one amino acid (glycine) linker. Mutations of the ACG into ATG or deletions of the 5'UTR of FMR1 were achieved by oligonucleotide ligations. To insure stability of the expanded CGG repeats, all CGG plasmids were transformed into STBL3 bacterial strain (Invitrogen) and growth at room temperature (22°c).

Quantitative real time RT-PCR.
Total RNAs from mouse tissues or cells were isolated by TriReagent (Molecular Research Center). cDNAs were generated using the Transcriptor High Fidelity cDNA synthesis kit (Roche Diagnostics) for quantification of mRNAs. qPCR of mRNAs were realized using the LightCycler 480 SYBR Green I Master (Roche) in a Lightcycler 480 (Roche) with 15 min at 94°C followed by 50 cycles of 15 sec at 94°C, 20 sec at 58°C and 20 sec at 72°C. RPLPO mRNA was used as standard and data were analyzed using the Lightcycler 480 analysis software (2ΔCt method).

Subcellular fractionation and PCR.
Cells were scraped in PBS, cells were pelleted by centrifugation at 3000 rpm for 10 min minutes. The pellet was resuspended in Dautry Buffer (Tris HCl pH 7,8, 10 mM; NaCl 140 mM, MgCl2 1,5 mM; EDTA 10 mM, NP40 0,5%) and kept on ice 5 minutes. The homogenate was centrifuged at 3000 rpm for 15 minutes during 5 min at 4°C, pellet corresponding to nuclear fraction and supernatant to cytosolic fraction. Cytosolic fraction was centrifuged at 13000 rpm for 15 minutes to remove potential nucleus and 1 ml TriReagent (Molecular Research Center) was added. The pellet was washed with 400 µl of Dautry Buffer, centrifuged at 3000 rpm for 5 minutes at 4°C. Supernatant was removed and the pellet was homogenized in 400 µl of Dautry Buffer and 1 ml TriReagent (Molecular Research Center) was added. Total RNA from nuclear or cytosolic fraction was isolated as described in the manufacturer's protocol of TriReagent (Molecular Research Center). cDNAs were generated using the Transcriptor High Fidelity cDNA synthesis kit (Roche Diagnostics) for quantification of mRNAs. PCR was performed with Taq polymerase (Roche), one denaturation step at 94 °C for 2 min, 25 cycles of amplification 94 °C for 1 min, 60 °C for 1 min, 72 °C for 2 min and a final step at 72 °C for 5 min using the primer described below. The PCR products were precipitated, analyzed by electrophoresis on a 6.5% polyacrylamide gel, stained with ethidium bromide and quantified with a Typhoon scanner.

Oligonucleotides.

RPLO_FW	GAAGTCACTGTGCCAGCCCA
RPLO_REV	GAAGGTGTAATCCGTCTCCA
U6_FW	CTCGCTTCGGCAGCACATATA
U6_REV	GGAACGCTTCACGAATTTGCG
FMR1_FW	GAAAACAACTGGCAGCCTGA
FMR1_REV	AGCTAACCACCAACAGCAAG
GFP_FW	ACGTAAACGGCCACAAGTTC
GFP_REV	AAGTCGTGCTGCTTCATGTG
(CGG)60x_FW	GAACCCACTGCTTACTGGCTTA
(CGG)60x_REV	AACGCTAGCCAGTTGGGTC
Transgene mouse FMRpolyG_FW	GCAAGCTGACCCTGAAGTTC
Transgene mouse FMRpolyG_REV	GTCTTGTAGTTGCCGTCGTC

Supplemental references

Birling MC, Dierich A, Jacquot S, Hérault Y, Pavlovic G. (2012). Highly-efficient, fluorescent, locus directed cre and FlpO deleter mice on a pure C57BL/6N genetic background. Genesis. 50(6):482-9.

Boissart C, Poulet A, Georges P, Darville H, Julita E, Delorme R, Bourgeron T, Peschanski M, Benchoua A. (2013). Differentiation from human pluripotent stem cells of cortical neurons of the superficial layers amenable to psychiatric disease modeling and high-throughput drug screening. Transl Psychiatry. Aug 20;3:e294.

Jung L, Tropel P, Moal Y, Teletin M, Jeandidier E, Gayon R, Himmelspach C, Bello F, André C, Tosch A, et al. (2014). ONSL and OSKM cocktails act synergistically in reprogramming human somatic cells into induced pluripotent stem cells. Mol Hum Reprod. 20(6):538-49.

Isaka F, Ishibashi M, Taki W, Hashimoto N, Nakanishi S, Kageyama R. (1999). Ectopic expression of the bHLH gene Math1 disturbs neural development. Eur J Neurosci. 11:2582-8

Marteyn A, Maury Y, Gauthier MM, Lecuyer C, Vernet R, Denis JA, Pietu G, Peschanski M, Martinat C. (2011). Mutant human embryonic stem cells reveal neurite and synapse formation defects in type 1 myotonic dystrophy. Cell Stem Cell. 8(4):434-44.