In silico, in vitro, and *in vivo* Approaches to Identify Molecular Players in Fragile X Tremor and Ataxia Syndrome

Saif N. Haify[1], Teresa Botta-Orfila[2], Renate K. Hukema[1] and Gian Gaetano Tartaglia[3,4,5,6,7]*

[1] Department of Clinical Genetics, Erasmus MC, Rotterdam, Netherlands, [2] Biological Fluids Bank of the Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain, [3] Centre for Genomic Regulation (CRG), The Barcelona Institute for Science and Technology, Barcelona, Spain, [4] Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain, [5] Department of Biology 'Charles Darwin', Sapienza University of Rome, Rome, Italy, [6] Department of Neuroscience and Brain Technologies, Istituto Italiano di Tecnologia, Genoa, Italy, [7] Universitat Pompeu Fabra (UPF), Barcelona, Spain

Fragile X-associated tremor/ataxia syndrome (FXTAS) is a late-onset neurodegenerative monogenetic disorder affecting carriers of premutation (PM) forms of the *FMR1* gene, resulting in a progressive development of tremors, ataxia, and neuropsychological problems. This highly disabling disease is quite common in the general population with an estimation of about 20 million PM carriers worldwide. The chances of developing FXTAS increase dramatically with age, with about 45% of male carriers over the age of 50 being affected. Both the gene and pathogenic trigger, a mutant expansion of CGG RNA, causing FXTAS are known. This makes it an interesting disease to develop targeted therapeutic interventions for. Yet, no such interventions are available at this moment. Here we discuss *in silico*, *in vitro,* and *in vivo* approaches and how they have been used to identify the molecular determinants of FXTAS pathology. These approaches have yielded substantial information about FXTAS pathology and, consequently, many markers have emerged to play a key role in understanding the disease mechanism. Integration of the different approaches is expected to provide crucial information about the value of these markers as either therapeutic target or biomarker, essential to monitor therapeutic interventions in the future.

Keywords: RNA, Fragile X associated tremor ataxia syndrome (FXTAS), mouse model, computational modeling, protein network

INTRODUCTION

Fragile X-associated Tremor and Ataxia syndrome (FXTAS) is a late onset neurodegenerative disease that affects Fragile X premutation (PM) carriers. The Fragile X gene (*FMR1*) responsible for the disease finds its origin on the X chromosome and codes for the Fragile X Mental Retardation Protein called FMRP. This protein is essential in the neuronal development and synaptic plasticity of the brain. The 5′ untranslated region (UTR) of the *FMR1* gene carries a variable number of CGG repeats, which is between 4 and 55 in healthy individuals. Due to the instability of the repeat the CGG sequence can expand to 55–200 repeats over generations (Leehey, 2009; Berman et al., 2014; Foote et al., 2016; Glineburg et al., 2018). These individuals are referred to as PM carriers.

FXTAS is characterized by several clinical features such as intention tremors and cerebellar gait ataxia, which are proposed to be major clinical diagnostic criteria for FXTAS disease pathology.

Other, more minor clinical criteria are Parkinsonism, working memory deficit and executive function deficit (Berry-Kravis et al., 2007; Leehey, 2009; Berman et al., 2014). Furthermore the neurodegeneration is characterized by brain atrophy, neuropsychiatric features and cognitive impairments or dementia (Foote et al., 2016). It is estimated that in the general population 1:110-250 females and 1:260-800 males are PM carriers with an interesting side note that not all carriers develop FXTAS due to incomplete penetrance (Tassone et al., 2012; Hunter et al., 2014; Foote et al., 2016; Glineburg et al., 2018). FXTAS penetrance is about 40% in male carriers and only 11–18% in female carriers (Berry-Kravis et al., 2007; Foote et al., 2016; Glineburg et al., 2018). A major hallmark of the disease is the presence of ubiquitin-positive intranuclear inclusions throughout the brains of FXTAS patients (**Figure 1**).

To date there is no cure available for PM carriers or FXTAS patients which is why development and validation of relevant models is necessary (Leehey, 2009). This will allow us to better understand this complex neurological disorder and also to be used for drug development in the future. At present, the onset and development of FXTAS is explained by two main mechanisms (Botta-Orfila et al., 2016): (i) RNA-mediated sequestration, and subsequent inactivation, of proteins attracted by the CGG trinucleotide repeats in the 5′ UTR region of *FMR1* RNA and (ii) toxic aggregation of Repeat-Associated Non-AUG (RAN) polyglycine peptides translated from the *FMR1* 5′ UTR (FMRpolyG; **Figure 1**). Previous work indicates that *FMR1* RNA forms aggregates containing specific proteins such as HNRNP A2/B1, MBNL1, LMNA, and INA (Iwahashi et al., 2006). Also the FMRpolyG peptide (Sellier et al., 2017) was found in the aggregates, together with CUGBP1, KHDRBS1, and DGCR8 being involved in splicing regulation, mRNA transport and regulation of microRNAs (Sellier et al., 2010).

Here we discuss *in silico*, *in vitro,* and *in vivo* approaches and how they have been used to identify the molecular determinants of FXTAS pathology. Our aim is to provide a comprehensive review for future research in the area. We believe that synergetic approaches from different research areas are necessary for FXTAS, because its pathological substrate is still under debate and there is still insufficient knowledge of targets for the development of a therapeutic intervention.

IN SILICO APPROACHES TO PREDICT MOLECULAR INTERACTIONS OCCURRING IN FXTAS

Difficulties in the biochemical purification of protein–RNA assemblies, which are extremely labile, make it extremely hard to extract the aggregates and identify the molecules that are crucial in disease spreading (Tartaglia, 2016). The use of computational methods aids in characterizing RNA-binding proteins (RBPs) with CGG repeats. We refer the reader to advanced reviews for more details related to some sophisticated algorithms (Cirillo et al., 2014).

PREDICTIONS OF PROTEIN–RNA INTERACTIONS OCCURRING IN FXTAS

The ability of the first *FMR1* exon (containing 79 CGG repeats in the PM range) to interact with specific RBPs. was assessed through the *cat*RAPID algorithm (Agostini et al., 2013). *cat*RAPID estimates the binding potential through van der Waals, hydrogen bonding and secondary structure propensities of both protein and RNA sequences allowing identification of binding partners (Bellucci et al., 2011). The library employed in the study was composed of 3340 DNA-binding, RNA-binding, and structurally disordered proteins (Livi et al., 2016).

*cat*RAPID (**Figure 2**) identified known CGG-binding proteins such as ROA (1, 2, and 3), SRSF (1, 4, 5, 6, 7, and 10), HNRNP (C, D, M) as well as MBNL1 and KHDRBS3 (Cid-Samper et al., 2018) (see section "*In vitro* Approaches to Predict Molecular Interactions Occurring in FXTAS"). In addition, strong binding propensities were found for an additional set of 92 proteins that are known to aggregate in stress granules (Jain et al., 2016) and a group of 37 RBPs that have strong potential to aggregate, as predicted by the *cat*GRANULE algorithm (Bolognesi et al., 2016). Among the RBPs identified, there are 25 splicing factors, including LSM3, SFPQ and TRA2A, 24 RBPs involved in RNA metabolism, including FUS, RBM8A, AGO2 and stress granules RBPs such as TIA1, MBNL1, and DDX1 (Cid-Samper et al., 2018) as well as CIRBP and PTBP2 (Cirillo et al., 2013).
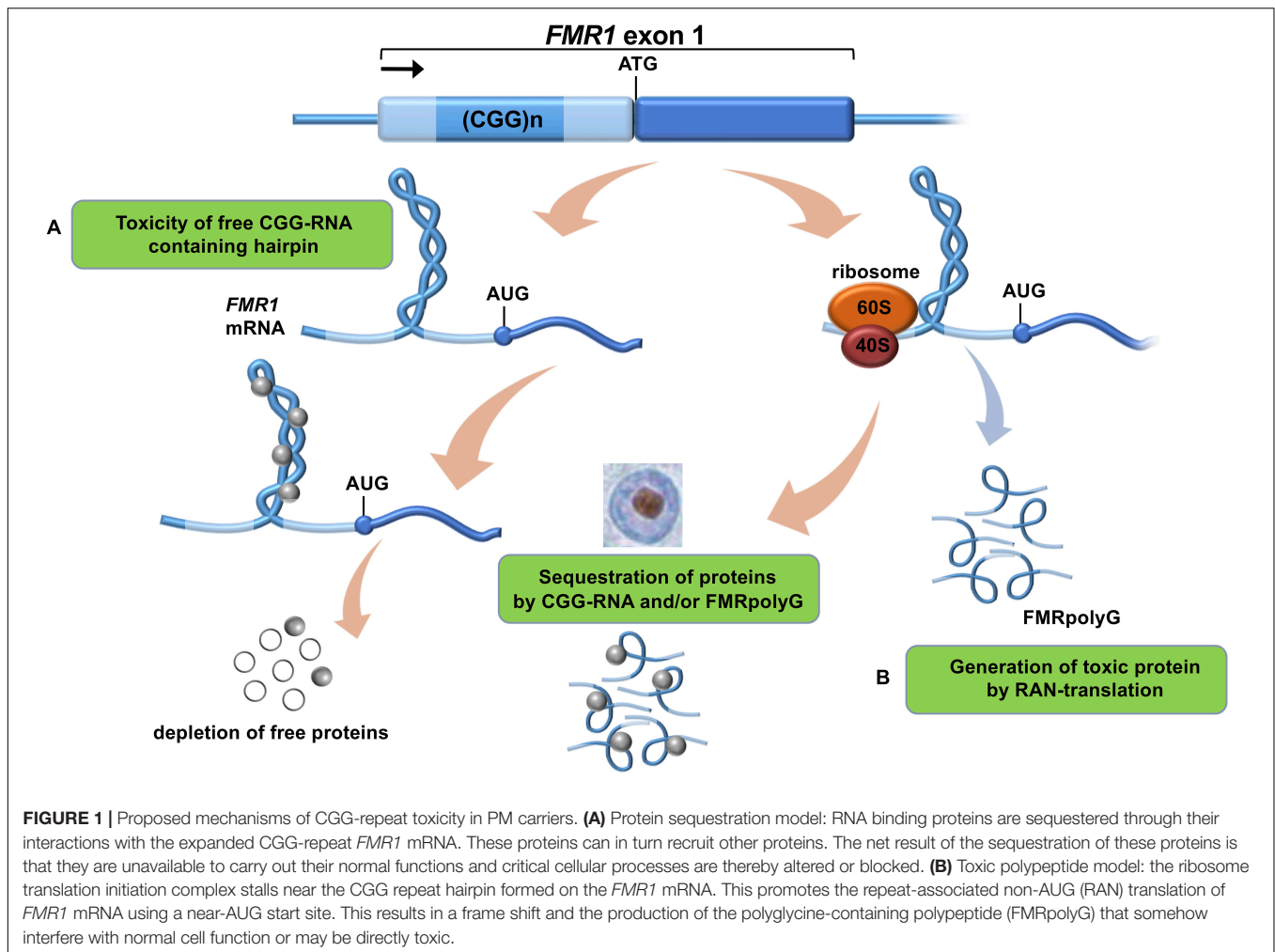
IN VITRO APPROACHES TO PREDICT MOLECULAR INTERACTIONS OCCURRING IN FXTAS

Protein interaction with CGG repeats can be determined by several experimental *in vitro* approaches. The success in discovering new findings is influenced by the technical capacity to preserve the natural characteristics of the protein–RNA partners, such as the secondary and tertiary structure, electrostatic and hydrophobic interactions, hydrogen bonding, rate of transcription of RNA and translation of protein, etc. We refer the reader to advanced reviews for details related to the experimental methods (Dasti et al., 2019).

CGG repeats form hairpins *in vitro* (Krzyzosiak et al., 2012). UV-monitored structure melting indicates that they are more stable than hairpins formed by CAG, CUG, or CCG repeats. Although the *in vivo* structure might differ, due to interactions with proteins and other molecules, crystallography supports the model that CGG repeats have intermolecular duplexes (Kiliszek et al., 2011).

RNA PULL-DOWN TO DETECT CGG REPEATS INTERACTIONS

The RNA pull-down consists on a selective extraction of a protein–RNA complex from a whole sample lysate, by using high-affinity tags, e.g., biotin on the RNA molecule. After tagging the known interaction partner, RNA, the RBPs complexed with it are

FIGURE 1 | Proposed mechanisms of CGG-repeat toxicity in PM carriers. **(A)** Protein sequestration model: RNA binding proteins are sequestered through their interactions with the expanded CGG-repeat *FMR1* mRNA. These proteins can in turn recruit other proteins. The net result of the sequestration of these proteins is that they are unavailable to carry out their normal functions and critical cellular processes are thereby altered or blocked. **(B)** Toxic polypeptide model: the ribosome translation initiation complex stalls near the CGG repeat hairpin formed on the *FMR1* mRNA. This promotes the repeat-associated non-AUG (RAN) translation of *FMR1* mRNA using a near-AUG start site. This results in a frame shift and the production of the polyglycine-containing polypeptide (FMRpolyG) that somehow interfere with normal cell function or may be directly toxic.

purified by using agarose or magnetic beads. To identify which proteins are associated with expanded CGG repeats, Sellier and colleagues adopted an *in vitro* approach. Proteins extracted from mouse brain and COS7 cell nuclei were captured on streptavidin resin coupled to biotinylated *in vitro*-transcribed RNA composed of 60CGG repeats, eluted, separated on SDS–PAGE gels and identified by MALDI-TOF analysis. More than 20 proteins were identified, including a heat-shock protein and several RBPs, such as SRSF (1, 4, 5, 6, 7, and 10), MBNL1 and HNRNP-G (Sellier et al., 2010). The list of interactors (**Figure 2**) included SPNR, HNRNP-A1, HNRNP-A2/B, HNRNP-C, HNRNP-D, HNRNP-E, and HNRNP-H (Sellier et al., 2010).

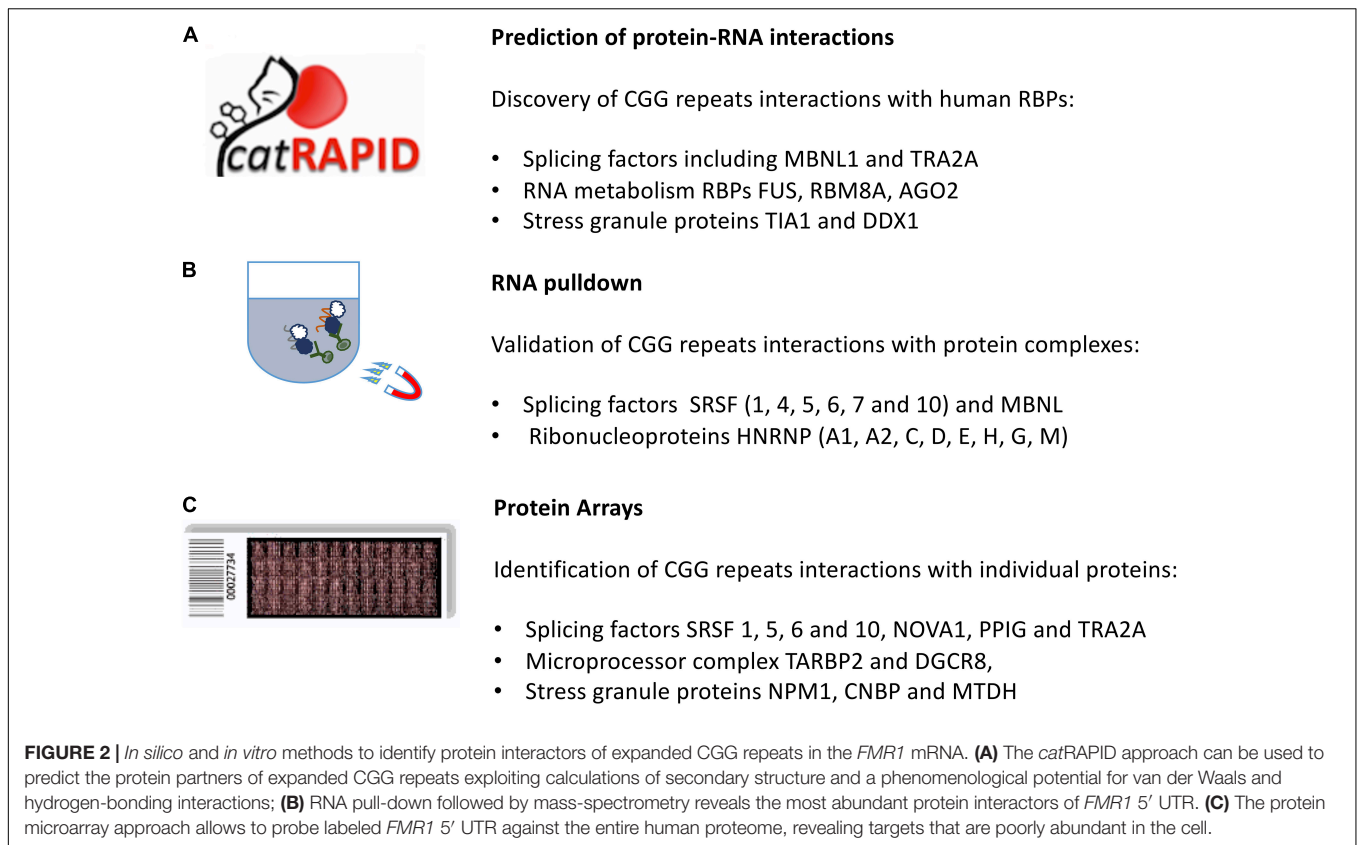## PROTEIN MICROARRAYS TO DETECT *FMR1* INTERACTIONS

Protein microarray technology was used to detect RBP interactions with the first *FMR1* exon (Cid-Samper et al., 2018). In this approach, individual human proteins, expressed in a eukaryotic system and subsequently purified, are isolated in separated nitrocellulose chambers and the RNA labeled with

Cy5 is used for probing (Marchese et al., 2017). Both expanded (79CGG, permutated range) and normal (21CGG) repeats were probed on independent replicas and the 3′ UTR of a similar length transcript, SNCA, was used as a control for the specificity of RBP interactions (Marchese et al., 2017).

Using fluorescence intensities to measure binding affinities (Cirillo et al., 2016), previously identified partners SRSF 1, 5, and 6 ranked in the top 1% of all interactions, followed by KHDRBS3 (2%) and MBNL1 (5%). The overall list included 85 RBPs showing an enrichment in Gene Ontology terms related to splicing activity, as reported by *clever*GO (Klus et al., 2015) and includes several SRSF proteins, PCBP 1 and 2, HNRNP-A0 and F, NOVA1, PPIG, and TRA2A (**Figure 2**).

RNA ELECTROPHORETIC MOBILITY SHIFT ASSAY (RNA-EMSA) TO VALIDATE INDIVIDUAL PROTEIN INTERACTIONS WITH CGG REPEATS

In the RNA-EMSA assay protein–RNA interactions are detected as migration differences in gel electrophoresis: the RNA probe

A

Prediction of protein-RNA interactions

Discovery of CGG repeats interactions with human RBPs:

- Splicing factors including MBNL1 and TRA2A
- RNA metabolism RBPs FUS, RBM8A, AGO2
- Stress granule proteins TIA1 and DDX1

B

RNA pulldown

Validation of CGG repeats interactions with protein complexes:

- Splicing factors SRSF (1, 4, 5, 6, 7 and 10) and MBNL
- Ribonucleoproteins HNRNP (A1, A2, C, D, E, H, G, M)

C

Protein Arrays

Identification of CGG repeats interactions with individual proteins:

- Splicing factors SRSF 1, 5, 6 and 10, NOVA1, PPIG and TRA2A
- Microprocessor complex TARBP2 and DGCR8,
- Stress granule proteins NPM1, CNBP and MTDH

FIGURE 2 | *In silico* and *in vitro* methods to identify protein interactors of expanded CGG repeats in the *FMR1* mRNA. **(A)** The *cat*RAPID approach can be used to predict the protein partners of expanded CGG repeats exploiting calculations of secondary structure and a phenomenological potential for van der Waals and hydrogen-bonding interactions; **(B)** RNA pull-down followed by mass-spectrometry reveals the most abundant protein interactors of *FMR1* 5′ UTR. **(C)** The protein microarray approach allows to probe labeled *FMR1* 5′ UTR against the entire human proteome, revealing targets that are poorly abundant in the cell.

is labeled by radioactivity of by fluorescent or chemiluminiscent dyes, and incubated with the whole protein extract from cell lysate, in different concentrations of the first one. In case the RNA is selectively bound to proteins, the electrophoretic band will differ from a negative control, as a migration shift. In this assay, an experimental secondary approach is performed, by assessing competitive binding: an excess of unlabeled RNA is incubated with the binding reaction and in case the shifted signal decreases, there is evidence for specificity of the binding.

Using RNA-EMSA, the binding of HNRNP-A2 to 105CGG repeats was studied in the presence of the brain cytoplasmic (BC1) RNA involved in neuronal translational (Muslimov et al., 2011). Differently from 105CCC repeats, the 105CGG repeats competed with binding of BC1 RNA to HNRNP-A2, which indicates impairment of neuronal function. Using 105CGG repeats at levels comparable to *FMR1* abundance in PM disease cells, the authors showed that distal dendritic delivery of BC1 RNA is significantly reduced while 105CCC repeats had no effect on dendritic BC1 targeting (Muslimov et al., 2011).

RNase H PROTECTION APPROACH TO CHARACTERIZE THE *FMR1* R-LOOP

The RNase H protection approach is used to detect DNA and RNA fragments in cell lysates. RNase H cleaves the target RNA molecule at a specific site hybridized with a DNA probe. In these

specific sites, if a protein is bound, the hybridization is blocked, therefore no cleavage by RNase H will occur.

It has been observed that transcription through the GC-rich *FMR1* 5′ UTR region favors formation of a three-stranded nucleic acid structure, composed of a DNA:RNA called R-loop formation, with the nascent RNA assembling with the template DNA strand (Loomis et al., 2014). Using DNA:RNA immunoprecipitation of genomic DNA from cultured human dermal fibroblasts with both normal and PM alleles, the authors reported for *FMR1* R-loop formation As expected for R-loop formation. Treatment with purified recombinant human RNases H1 and H2 eliminated DNA–RNA interaction (Loomis et al., 2014).
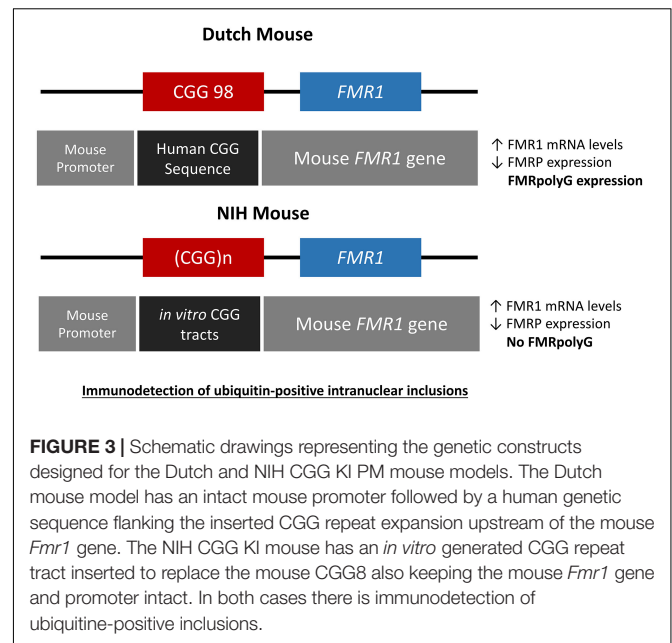
FLUORESCENT *IN SITU* HYBRIDIZATION CO-LOCALIZATION (FISH) TO LOCALIZE CGG REPEATS IN THE CELL

Fluorescent *in situ* hybridization (FISH) co-localization techniques require knowing in advance which RNA and protein interaction will be studied. This technique is commonly used in FXTAS research as it allows verifying where CGG repeats localize with other molecules. By means of FISH coupled to immunofluorescence it has been shown that CGG expansions and TRA2A significantly co-localize in COS7 cells (Cid-Samper et al., 2018). Similarly, co-localization of MBNL1, KHDRBS1 and HNRNP-G within CGG aggregates was observed in COS7

cells (Sellier et al., 2010). By contrast, other *in vitro* identified candidates such as SPNR, HNRNP-A1, HNRNP-A2/B, HNRNP-C, HNRNP-D, HNRNP-E, and HNRNP-H have been shown by FISH to have poor co-localization with CGG repeats (Sellier et al., 2010). Several proteins, including a number of heat-shock proteins, HNRNP-A2/B1, CUGBP1, lamin-A/C and MBP were found to localize with ubiquitin-positive inclusions in CGG-expressing *Drosophila*, KI mouse model and FXTAS patients. Some of the co-localizations are model dependent. Indeed, it should be noted that PURα co-localizes with cytoplasmic CGG repeats in flies (Jin et al., 2007) but not in mammalian cells, where it is found strictly nuclear (Sellier et al., 2010). So, given the propensity of PUR-α to interact with CGG repeats (Cirillo et al., 2013), it is possible that its subcellular localization prevents physical interaction with RNA.

IN VIVO APPROACHES TO PREDICT MOLECULAR INTERACTIONS OCCURRING IN FXTAS

The development of animal models has provided the field with important clinical but also molecular information regarding the pathology associated with CGG repeat expansions on *FMR1* Fragile X syndrome (FXS), PM carriers and FXTAS. PM carriers of the expanded CGG repeat have a 2 to 8-fold elevation of *FMR1* mRNA levels when compared to healthy CGG repeat individuals (Hagerman and Hagerman, 2002). Increased levels of *FMR1* mRNA with a CGG expansion are proposed to be toxic due to sequestration of specific RNA binding proteins that are crucial for normal cell function, partly resulting in FXTAS pathology. This theory is supported by inclusions isolated from FXTAS post-mortem brain tissue containing over 30 critical proteins, such as Lamin A/C, KHDRBS1, Drosha, and HNRNP-A2 (Glineburg et al., 2018). The most prominent neuropathological hallmark of FXTAS is the presence of eosinophilic, ubiquitin-positive intranuclear inclusions in neurons and astroglia in the entire brain upon post-mortem histological analysis (Greco et al., 2002, 2006, 2007, 2008; Gokden et al., 2009). Additional neuropathological features found in FXTAS are reduced number of Purkinje cells, Bergmann gliosis, axonal swelling in the granular cell layer of the cerebellum, and prominent cortical and subcortical white matter pathology (Greco et al., 2002, 2006). A second mechanism more recently found suggests an additional model for toxicity in FXTAS called repeat-associated non-AUG (RAN) translation. This mechanism proposes that the expanded CGG repeat is translated in absence of a canonical start codon resulting in pathogenicity through translation of a toxic polyglycine (FMRpolyG) peptide (Berman et al., 2014; Glineburg et al., 2018). Over the past 10 to 15 years many well-characterized mouse models have shown that aforementioned theories could indeed contribute to disease pathology. These mouse models also allowed us to characterize FXTAS disease pathogenesis and progression together with the underlying neurobiological changes and to develop and test highly potential targeted therapeutic interventions. This part of the review will describe the current available mouse models for FXTAS with their

FIGURE 3 | Schematic drawings representing the genetic constructs designed for the Dutch and NIH CGG KI PM mouse models. The Dutch mouse model has an intact mouse promoter followed by a human genetic sequence flanking the inserted CGG repeat expansion upstream of the mouse *Fmr1* gene. The NIH CGG KI mouse has an *in vitro* generated CGG repeat tract inserted to replace the mouse CGG8 also keeping the mouse *Fmr1* gene and promoter intact. In both cases there is immunodetection of ubiquitine-positive inclusions.

specific aspects, advantages and limitations, and what insights they have provided over the past years into disease mechanism.

THE DUTCH MOUSE

The first mouse model generated to exhibit much of the pathology seen in affected PM carriers and in FXTAS pathology at the genetic level but as well as histopathological and molecular level was the Dutch mouse (CGG$_{dut}$ KI) model. This mouse model was developed in Rotterdam at the Erasmus MC in the Netherlands. This model was generated by replacing the endogenous murine *Fmr1* 8CGG repeat with a human 98CGG repeat containing the human *FMR1* flanking regions. The *Fmr1* mouse promoter was left unchanged. This was done by homologous recombination in embryonic stem (ES) cells (Bontekoe et al., 2001; Foote et al., 2016). Upon paternal and maternal transmission the CGG$_{dut}$ KI mouse showed mild instability of the CGG repeat with both short expansions and contractions present (Bontekoe et al., 2001; Willemsen et al., 2003; Brouwer et al., 2007). The CGG$_{dut}$ KI mice (**Figure 3**) have been bred into a C57BL/6J and FVB background over several generations to establish lines with expanded alleles greater than 450CGGs (Bontekoe et al., 2001; Berman and Willemsen, 2009). Although repeat lengths of more than 450CGGs were found, no increased methylation of the *Fmr1* gene has been reported. When examining these mice at the histopathological level clear ubiquitin-positive intranuclear inclusions could be shown similar to what is seen in FXTAS post-mortem patient brain. As mentioned before FXTAS patients are characterized with elevated *FMR1* mRNA and slightly decreased FMRP protein. Although FXTAS is considered to be a late-onset neurodegenerative disorder, some phenotypes of the disease in patients and in the CGG$_{dut}$ KI mouse could suggest that

FXTAS might also have features of a neurodevelopmental disorder. Developmental abnormalities during the PM stage such as altered learning and memory may contribute to the late manifestation of FXTAS. For example PM carriers have a smaller hippocampus that correlates with impaired performance in standardized tests of memory (Hagerman, 2013). Hippocampal neurons with an abnormal dendritic morphology have been observed in FXTAS neurons at a time of development when nuclear inclusions are not detectable yet. When observing hippocampal neurons from the CGG_{dut} KI mouse researchers found cortical migration to be affected in these mice and that these neurons upon culturing display shorter dendrites and have a reduced dendritic complexity (Hunsaker et al., 2009, 2012; Chen et al., 2010; Cunningham et al., 2011; Berman et al., 2014). Recently, the impact of elevated *Fmr1* mRNA levels on the morphology of dendrites and axons was studied more in depth (Drozd et al., 2019). Indeed, these morphological phenotypes are associated with increased levels of *Fmr1* mRNA because upon treatment with shRNAs specifically targeting the *Fmr1* mRNA these phenotypes are rescued. In addition, proteomic analysis in the CGG_{dut} KI mouse showed that upon rescue of *FMR1* mRNA levels a large number of important RBPs such as Tia1, HNRNPll, and ROAA could be normalized. Other rescued proteins are Rab-GTPases, which are critical for synaptic function in neurons in brain developmental disorders. Also Aldh4a1/P5CDH and Samm50, which are involved in mitochondrial dysfunction, were found to be deregulated in these mice confirming previous hypotheses indicating the important role of mitochondria in FXTAS pathology. All these proteins provide possible future pharmacological targetable molecules for early therapeutic intervention for FXTAS. More research focusing on other aspects of neurodevelopment in FXTAS such as the role of FMRP in the PM stage is necessary before one can categorize FXTAS to also be a neurodevelopmental disorder. All together, the CGG_{dut} KI mouse model nicely recapitulates the histopathology and molecular changes observed in patients (Berman and Willemsen, 2009; Foote et al., 2016).

THE NIH MOUSE

The National Institutes of Health developed a second KI mouse model (CGG_{nih} KI) with a CGG repeat length between 118-120CGGs by generating a CGG repeat flanked by *Sfi*I sites and then ligating this CGG repeat into the target construct with exon 1 of the mouse *Fmr1* gene in the correct orientation minimalizing changes to the mouse regions flanking the CGG repeat, the gene itself and the promoter (Entezam et al., 2007; Hoffman et al., 2012; Foote et al., 2016). This resulted in these mice having a translational TAA stop codon upstream of the CGG repeat. This TAA stop codon is present in the endogenous mouse *Fmr1* gene but not in the human *FMR1* gene. These mice, similar to the CGG_{dut} KI mouse show moderate intergenerational expansions with elevated *Fmr1* mRNA levels and decreased FMRP levels at the molecular level without any methylation of the *Fmr1* gene. Importantly, histopathological analysis revealed less ubiquitin-positive intranuclear inclusions being present in these
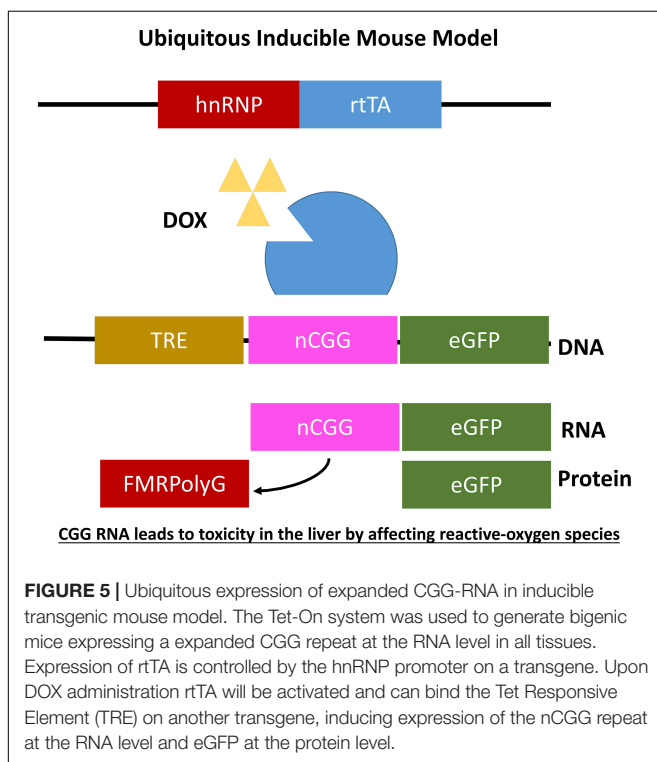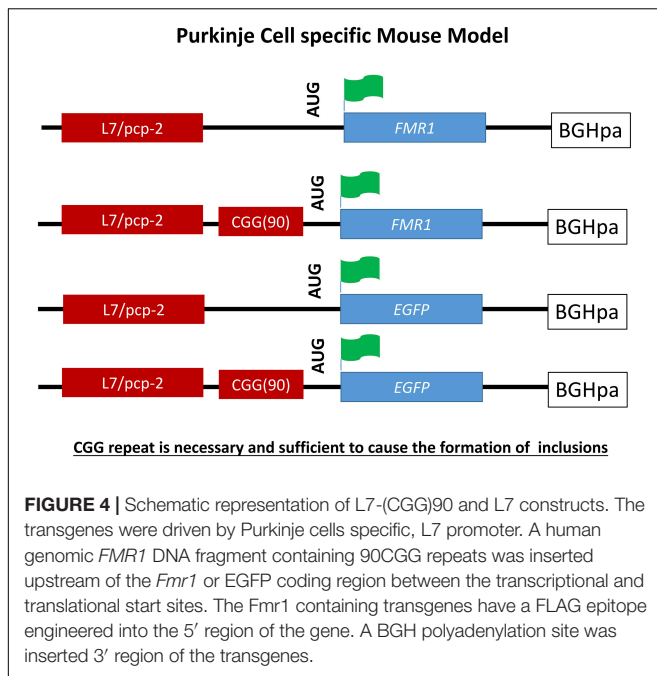
mice compared to the CGG_{dut} KI mouse (Entezam et al., 2007; Foote et al., 2016).

PURKINJE CELL SPECIFIC MOUSE MODEL

A Purkinje cell specific mouse model was generated to provide evidence that the expanded CGG repeat is necessary to cause FXTAS pathology similar to human and distinguish these effects seen from possible alterations in the *Fmr1* gene. To do so transgenic mice were generated expressing a CGG repeat in the context of *Fmr1* (L7-CGG90-*Fmr1*) or in context of the enhanced green fluorescent protein (L7-CGG90-EGFP) specifically in Purkinje neurons using the pcp2/L7 promoter in the cerebellum (Hashem et al., 2009). With these two lines it would be possible to determine whether ectopic expression of a 90CGG expanded repeat would cause neurodegeneration in the cerebellum or not. A significant Purkinje cell loss was observed in both the L7-CGG90-*Fmr1* and L7-CGG90-EGFP mice (**Figure 4**). Ubiquitin-positive intranuclear inclusions being the hallmark of FXTAS could be found in Purkinje cells of both the lines with the expanded CGG repeat but not in the control mice suggesting an essential role for the expanded CGG repeat and RNA in inclusion formation. It is assumed that the proteasome degradation pathway is involved in FXTAS disease progression in humans since essential proteins involved in this pathway are found in ubiquitin-positive inclusions (Willemsen et al., 2003). Inclusions in the Purkinje cell specific mice contained as well the 20S core complex of the proteasome, heat-shock protein HSP40 and RAD23B protein, which is a known protein of the ubiquitin-mediated proteasome degradation pathway. Behavioral examination revealed that mice expressing the expanded CGG mRNA in the Purkinje cells had an impaired motor performance on the rotarod test. These neuropathological and behavioral observations provide evidence that expression of the CGG repeat mRNA is sufficient to cause Purkinje cell dysfunction and loss of neurons similar to that reported in FXTAS patients (Greco et al., 2002). Although there might be a connection between intranuclear inclusion formation and in this case Purkinje cell death, such a conclusion could only be made upon understanding the role of other essential proteins present in these inclusions.

FMRpolyG AND CGG-RNA TRANSGENIC MICE

To further elaborate on the two hypotheses regarding the pathogenic mechanism in PM and FXTAS pathology based on expression of mutant *FMR1* mRNA bearing an expanded CGG repeat a new mouse model was suggested. How *FMR1* mRNA containing expanded CGG repeats is pathogenic was still unclear. To study this, the lab of Charlet Berguerand at the Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC) in France created two transgenic mouse models (**Figure 5**). The first model contains the full human 5′ UTR of *FMR1* with expanded 99CGG repeats that express both the CGG-RNA as

FIGURE 4 | Schematic representation of L7-(CGG)90 and L7 constructs. The transgenes were driven by Purkinje cells specific, L7 promoter. A human genomic *FMR1* DNA fragment containing 90CGG repeats was inserted upstream of the *Fmr1* or EGFP coding region between the transcriptional and translational start sites. The Fmr1 containing transgenes have a FLAG epitope engineered into the 5′ region of the gene. A BGH polyadenylation site was inserted 3′ region of the transgenes.

FIGURE 5 | Ubiquitous expression of expanded CGG-RNA in inducible transgenic mouse model. The Tet-On system was used to generate bigenic mice expressing a expanded CGG repeat at the RNA level in all tissues. Expression of rtTA is controlled by the hnRNP promoter on a transgene. Upon DOX administration rtTA will be activated and can bind the Tet Responsive Element (TRE) on another transgene, inducing expression of the nCGG repeat at the RNA level and eGFP at the protein level.

well as the FMRpolyG protein. The second line referred to as the mutant line also expresses 99CGG repeats, but without the non-canonical ACG start-codon and the surrounding 5′ UTR sequence region only expressing the CGG-RNA and no protein (Sellier et al., 2017). Both transgenic mouse lines have a CAG promoter inserted within the *Rosa26* locus and expression is controlled by a loxP-polyadeneylation cassette (Sellier et al.,

2017). Using a ubiquitously and embryonically expressed Cre recombinase results in deletion of the loxP cassette, which leads to high expression of transgene RNA throughout the brain, heart and liver, with less expression in skeletal muscle and the kidneys. Transgene RNA expression is the same in both transgenic mouse lines. The full human *FMR1* 5′ UTR transgenic mice showed upon histological analysis FMRpolyG protein nuclear aggregates that co-localize with ubiquitin in the brain (Sellier et al., 2017). The mutant *FMR1* 5′ UTR mice did not show these aggregates. The occurrence of CGG RNA *foci* in the brain of these mice (**Figure 5**) is slightly lower compared to other previously reported mouse models (Sellier et al., 2010, 2017). FMRpolyG protein also accumulated in other tissue than the brain, which is similar to what our group reported regarding FMRpolyG aggregates in non-CNS tissues in FXTAS patients (Buijsen et al., 2014). Other sense proteins such as the polyalanine (FMRpolyA) and polyarginine (FMRpolyR) peptides from other reading frames were not observed in these mice. There is to some extent neuronal cell death and little loss of Purkinje neurons in the mice expressing CGG-RNA and the FMRpolyG protein but not in the mice expressing only CGG-RNA (Sellier et al., 2017).

FMR1 OVEREXPRESSING MICE

FMR1 mRNA toxicity is mentioned several times before to be one of the hypothesis causing FXTAS pathology and is supported by elevated levels of *FMR1* mRNA with an expanded CGG repeat (Hagerman and Hagerman, 2002). It could be that toxicity occurs due to expression of an expanded CGG repeat or because of just elevated *FMR1* mRNA levels independent of the CGG repeat being present. A *Drosophila* model shows that repeat length is a direct cause of pathology since longer repeats show extreme retina pathology while short repeats result in little pathology (Jin et al., 2003). To confirm these findings in mice, transgenic mice were generated overexpressing *FMR1* mRNA with a normal length CGG repeat consisting of 29CGGs (Fernandez et al., 2012). Human *FMR1* cDNA with 29CGG repeats under control of a SV40/T7 promoter was injected in oocytes to generate these transgenic 29CGG mice with 20–100 fold increase in *FMR1* mRNA in tissues such as liver, cerebral cortex and cerebellum. Although these mice express extreme levels of *FMR1* mRNA, there was no significant difference in general activity or anxiety-related behaviors in open-field tests, which suggests that expression of *FMR1* mRNA alone is not enough to induce pathology but that a CGG repeat expansion is necessary for pathology to occur.

A yeast artificial chromosome (YAC) containing the full-length human *FMR1* gene was used to generate transgenic mice overexpressing *FMR1* mRNA to study the toxic effect of *FMR1* mRNA as well as to study CGG repeat instability (Peier et al., 2000; Peier and Nelson, 2002). These mice were generated by direct microinjection of purified YAC DNA containing a 92CGG repeat isolated from an adult male permutation carrier into the pro-nuclei of fertilized FVB/N mouse oocytes and then transferred into foster mice which eventually led to identifying a mouse line having a repeat containing 90CGGs (Peier and
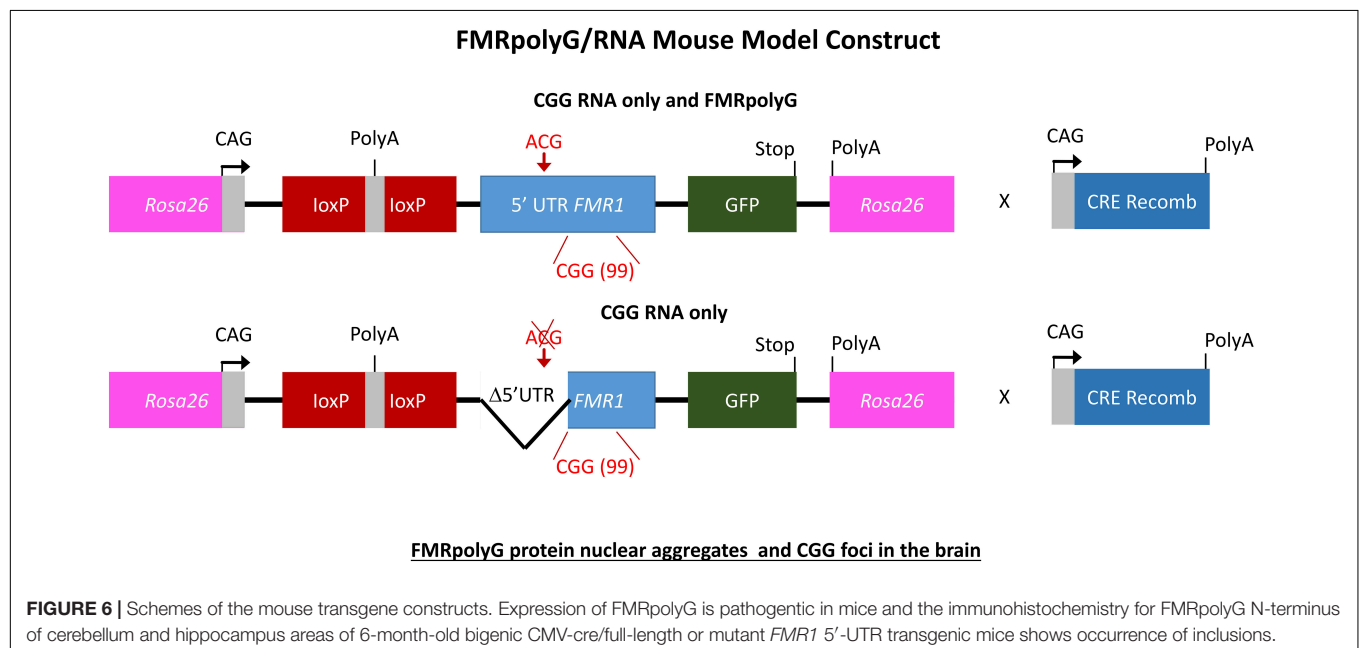
Nelson, 2002). These YAC mice show a 2–3 fold increase in expression of *FMR1* mRNA and a 10–15 fold increase in FMRP protein production (Peier et al., 2000; Spencer et al., 2008). Importantly to mention is that histological analysis presented no changes in overall brain morphology due to overexpression of *FMR1* mRNA or FMRP. When YAC mice were cross-bred with *Fmr1* KO mice lacking FMRP protein, some of the pathological features attributed to the absence of FMRP could be reversed and even overcorrected which also resulted in some abnormal behaviors. The authors attributed phenotype to overexpression of FMRP but the high levels of *FMR1* mRNA could also have contributed to these behavioral effects (Spencer et al., 2008).

INDUCIBLE MOUSE MODELS

As mentioned before it was not yet clear whether overexpression of RNA bearing normal CGG repeat length or an expanded CGG repeat is sufficient to induce toxicity in animal models and how to translate results in mice to the PM and FXTAS pathology. To further elaborate on this question two additional doxycycline inducible transgenic mouse models were generated. Both mouse models consisted of two transgenic mouse lines differing in the promoter used. The main advantage of both models is that expression of the transgene can be switched on and off using doxycycline. The first mouse line expresses the tetracycline response element (TRE) and the other line expressing the reverse tetracycline-controlled transactivator protein (rtTA), which can be activated using doxycycline or doxycycline-derivative in drinking water or food. The TRE-element was coupled to an expanded CGG repeat in frame with eGFP while the rtTA was coupled to an hnRNP promoter. Both mouse lines were crossbred generating double transgenic offspring with ubiquitous expression of 90CGG repeat RNA

(hnRNP-rtTA/TRE-90CGG-eGFP; **Figure 6**). Strange enough these mice died within 5 days after induction of the transgene using doxycycline. After histopathological analysis of the liver and molecular analysis two markers, cytochrome C and glutathione oxidase (GPX1), could be identified being affected in the livers with increased steatosis, mitochondrial dysfunction and subsequent apoptosis. Due to the early death of these mice no neuropathology could be observed. In summary this mouse model taught us that *in vivo* expression of expanded CGG RNA leads to severe toxicity in the liver by affecting reactive-oxygen species (ROS) signaling (Hukema et al., 2014).

Since FXTAS pathology is considered to be a neurodegenerative disease, it was clear that the next step would be to generate a brain specific doxycycline inducible mouse model expressing RNA containing an expanded CGG repeat. For this a brain specific prion protein (PrP)-rtTA driver was used to activate expression of normal length CGG RNA or expanded CGG repeat RNA fused to eGFP (Hukema et al., 2015; Castro et al., 2017). The PrP promoter drives expression in glial cells and neurons in the central nervous system (CNS). Compared to the hnRNP-rtTA mouse line, these mice could be treated for several weeks with doxycycline without any apparent liver toxicity. The RNA bearing an expanded CGG repeat was expressed the strongest in the cerebellum, hippocampus and striatum. Examining brain sections under the microscope showed ubiquitin-positive intranuclear inclusions co-localizing with FMRpolyG-positive intranuclear inclusions. Inclusion formation was followed in time showing increasing numbers and size when mice were longer treated with doxycycline. Other FXTAS-related proteins such as 20S core complex of the proteasome, HSP40 and RAD23B could also be found co-localizing with the ubiquitin-positive aggregates. Our group could show that

FIGURE 6 | Schemes of the mouse transgene constructs. Expression of FMRpolyG is pathogentic in mice and the immunohistochemistry for FMRpolyG N-terminus of cerebellum and hippocampus areas of 6-month-old bigenic CMV-cre/full-length or mutant *FMR1* 5'-UTR transgenic mice shows occurrence of inclusions.

FIGURE 7 | Diagram of DNA fragment used for pronuclear injection with either an 11CGG or 99CGG trinucleotide repeat expansion on exon. Astroglial-specific expression of nCGG repeat RNA yields RAN translation and formation of ubiquitin-positive and FMRpolyG-positive inclusions in astroglia and neurons.

introducing a wash-out period at an early stage was enough to halt and even reverse inclusion body formation. When repeating the wash-out step at a later point in time disease pathology could only be halted but not reversed (Hukema et al., 2015; Castro et al., 2017). Upon doxycycline induction the brain specific inducible mice had deficits in the compensatory eye movements. By turning off expression of the RNA with an expanded repeat this functional phenotype could be stopped (Hukema et al., 2015). After doxycycline induction double transgenic brain specific inducible mice performed poorly on the rotarod test indicating that motor performance is affected, more specifically motor coordination and motor learning which was also in line with the high number of intranuclear inclusions observed in lobule X of the cerebellum. When examined in the open-field test for emotional disturbance these mice were more anxious but had no deficits in emotional learning nor memory impairments correlating with intranuclear inclusions found in the amygdala and the hippocampus (Hukema et al., 2015). Behavioral phenotype and rescue were paralleled by the intranuclear inclusions formed in several brain regions such as lobule X of the cerebellum, the central amygdala and basolateral nuclei region of the amygdala, and hippocampal sub-regions such as the dentate gyrus and third cornu ammonis region of the hippocampus (Castro et al., 2017). Turning off transgene expression could halt functional phenotype and even reverse neuropathology if intervened at an early stage potentially suggesting that brain region specific therapeutic intervention might be beneficial for FXTAS patients in the future (Castro et al., 2017).

ASTROCYTE-SPECIFIC MOUSE MODEL

The FXTAS field has been provided with many interesting mouse models but there are still questions left unanswered. One interesting question in the field concerns the specific role of astrocytes in FXTAS pathology. What is the role of

astroglial cells in FXTAS? Is for example expression of an expanded CGG repeat enough to induce pathology? FXTAS patient brains have inclusions throughout the entire brain, in neurons but also in astroglial cells such as the Bergmann glia in the cerebellum (Willemsen et al., 2003). Although the CGG_{dut} KI mouse model of FXTAS recapitulates most aspects seen in FXTAS, these mice have relatively few numbers of astrocytes with intranuclear inclusions. Therefore an astroglial-specific mouse model was necessary. This transgenic mouse line with a C57BL/6j background was generated through pronuclear injection using a astrocyte-specific Gfa2 promoter to induce expression of an expanded 99CGG repeat fused to an eGFP marker gene only in astrocytes (**Figure 7**) (Wenzel et al., 2019). Immunocytochemical analysis of eGFP expression patterns show expression of 99CGG RNA was restricted to astroglia and Bergmann glia only and was not present in neurons, microglia or oligodendrocytes. Ubiquitin-positive and FMRpolyG-positive inclusions were observed in the nucleus of astroglia with inclusion bodies also being present in the cytoplasm of astrocyte processes. Although astrocytes with inclusion bodies were low in number, they were widely distributed in the brain such as in the neocortex, cerebellum and occasionally in subcortical regions in the hypothalamus and some brain stem nuclei. These mice had no sign of Purkinje neuronal dropout as well as no ubiquitin-positive inclusions in Purkinje cells could be found (Sellier et al., 2017; Wenzel et al., 2019). This study was the first of its kind to provide evidence of FXTAS related RAN-translation products being present in mouse astroglia. Surprisingly, these mice also had ubiquitin-positive intranuclear inclusions in neurons even though there was no leaky expression in neurons of the Gfa2-promoter and these cells did not express the expanded 99CGG RNA suggesting a prion-like spread of pathology, similar to what is seen in other models, from astrocytes to neurons by a cell-to-cell transfer mechanism (Westergard et al., 2016). Several behavioral tests were performed to examine neurological disease phenotypes associated with FXTAS pathology. These mice had difficulty with basic gait parameters such as stance time and range of motion. Also when placed on the ladder rung test their motor performance was clearly affected. These mice did not suffer from any emotional disturbance nor memory learning deficits. Although this model presents evidence of key hallmarks of FXTAS pathology, more work is necessary to further elaborate on the role of astroglia in FXTAS disease pathology.

CONCLUSION

Different experiments have been done to study the expanded CGG repeats in the *FMR1* gene. Despite many insights, the different *in silico, in vitro,* and *in vivo* approaches have not yet resulted in a targeted therapeutic intervention for FXTAS. Indeed, while we know the exact mutation leading to disease, we still could not identify the pathogenic trigger, or combination of triggers.

The fact that a number of proteins have been consistently found with different approaches is reassuring that the synergy

from different research areas give hints on their roles in pathology. For instance, studies based on RNA-pulldown (Sellier et al., 2010) and protein-microarray approaches (Cid-Samper et al., 2018) consistently show that CGG triplets, which are known to form stable hairpins *in vitro* (Krzyzosiak et al., 2012), interact physically with proteins (e.g., HNRNPs and SRSFs) that have domains annotated as preferably binding single-stranded regions (e.g., RRM domains). These results, supported by *cat*RAPID predictions (Cirillo et al., 2013), not only indicate that a number of interactions could occur in loop regions of the RNA but also suggest that the structural disorder, which is found enriched by *cat*GRANULE (Bolognesi et al., 2016) in protein partners such as TRA2A (Cid-Samper et al., 2018), might be involved in the binding, in addition to canonical domains (Castello et al., 2016). Thus, loop regions of the RNA or disordered regions of proteins could be the target for therapeutic intervention.

Candidates identified in *in silico* and *in vitro* approaches can serve as targets for intervention or as biomarkers to test the therapy can be determined by performing *in vivo* studies. For instance, the 9-hydroxy-5,11-dimethyl-2-(2-(pi-peridin-1-yl)ethyl)-6H-pyrido[4,3-b]carbazol-2-ium (also named *1a*) compound has been prioritized as a candidate for its ability to bind to loops produced by unpaired G-G nucleotides in CGG hairpins (Disney et al., 2012). The 1a compound has been used in cellular models to prevent interactions with RBPs such as TRA2A and consequent co-aggregation (Cid-Samper et al., 2018). In agreement with this finding, it has been reported that *1a* and structurally similar compounds reduce RAN translation of *FMR1* in cultured cells and patient-derived neurons (Su et al., 2014). Three small-molecule RAN inhibitors have been recently shown to interact with CGG and GGGGCC

repeat RNAs by circular dichroism and native gel analysis: BIX01294, CP-31398, and propidium iodide (Green et al., 2019). These small, bioactive compounds can selectively inhibit RAN translation across multiple disease-causing repeat expansion mutations and could therefore be employed in future treatments (Green et al., 2019).

AUTHOR CONTRIBUTIONS

GT and TB-O wrote the *in silico* and *in vitro* section. RH and SH wrote the *in vivo* section.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Agostini, F., Zanzoni, A., Klus, P., Marchese, D., Cirillo, D., and Tartaglia, G. G. (2013). catRAPID omics: a web server for large-scale prediction of protein-RNA interactions. *Bioinformatics* 29, 2928–2930. doi: 10.1093/bioinformatics/btt495

Bellucci, M., Agostini, F., Masin, M., and Tartaglia, G. G. (2011). Predicting protein associations with long noncoding RNAs. *Nat. Methods* 8, 444–445. doi: 10.1038/nmeth.1611

Berman, R. F., Buijsen, R. A., Usdin, K., Pintado, E., Kooy, F., Pretto, D., et al. (2014). Mouse models of the fragile X premutation and fragile X-associated tremor/ataxia syndrome. *J. Neurodev. Disord.* 6:25. doi: 10.1186/1866-1955-6-25

Berman, R. F., and Willemsen, R. (2009). Mouse models of fragile X-associated tremor ataxia. *J. Investig. Med.* 57, 837–841. doi: 10.2310/JIM.0b013e3181af59d6

Berry-Kravis, E., Abrams, L., Coffey, S. M., Hall, D. A., Greco, C., Gane, L. W., et al. (2007). Fragile X-associated tremor/ataxia syndrome: clinical features, genetics, and testing guidelines. *Mov. Disord*. 22, 2018–2030; quiz 2140.

Bolognesi, B., Lorenzo Gotor, N., Dhar, R., Cirillo, D., Baldrighi, M., Tartaglia, G. G., et al. (2016). A concentration-dependent liquid phase separation can cause toxicity upon increased protein expression. *Cell Rep.* 16, 222–231. doi: 10.1016/j.celrep.2016.05.076

Bontekoe, C. J., Bakker, C. E., Nieuwenhuizen, I. M., van der Linde, H., Lans, H., de Lange, D., et al. (2001). Instability of a (CGG)98 repeat in the Fmr1 promoter. *Hum. Mol. Genet.* 10, 1693–1699. doi: 10.1093/hmg/10.16.1693

Botta-Orfila, T., Tartaglia, G. G., and Michalon, A. (2016). Molecular pathophysiology of fragile X-associated tremor/ataxia syndrome

and perspectives for drug development. *Cerebellum* 15, 599–610. doi: 10.1007/s12311-016-0800-2

Brouwer, J. R., Mientjes, E. J., Bakker, C. E., Nieuwenhuizen, I. M., Severijnen, L. A., Van der Linde, H. C., et al. (2007). Elevated Fmr1 mRNA levels and reduced protein expression in a mouse model with an unmethylated fragile X full mutation. *Exp. Cell Res.* 313, 244–253. doi: 10.1016/j.yexcr.2006.10.002

Buijsen, R. A., Sellier, C., Severijnen, L. A., Oulad-Abdelghani, M., Verhagen, R. F., Berman, R. F., et al. (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. doi: 10.1186/s40478-014-0162-2

Castello, A., Horos, R., Strein, C., Fischer, B., Eichelbaum, K., Steinmetz, L. M., et al. (2016). Comprehensive identification of RNA-binding proteins by RNA interactome capture. *Methods Mol. Biol.* 1358, 131–139. doi: 10.1007/978-1-4939-3067-8_8

Castro, H., Kul, E., Buijsen, R. A. M., Severijnen, L., Willemsen, R., Hukema, R. K., et al. (2017). Selective rescue of heightened anxiety but not gait ataxia in a premutation 90CGG mouse model of fragile X-associated tremor/ataxia syndrome. *Hum. Mol. Genet.* 26, 2133–2145. doi: 10.1093/hmg/ddx108

Chen, Y., Tassone, F., Berman, R. F., Hagerman, P. J., Hagerman, R. J., Willemsen, R., et al. (2010). Murine hippocampal neurons expressing Fmr1 gene premutations show early developmental deficits and late degeneration. *Hum. Mol. Genet.* 19, 196–208. doi: 10.1093/hmg/ddp479

Cid-Samper, F., Gelabert-Baldrich, M., Lang, B., Lorenzo-Gotor, N., Ponti, R. D., Severijnen, L., et al. (2018). An integrative study of protein-RNA condensates identifies scaffolding RNAs and reveals players in fragile X-associated tremor/ataxia syndrome. *Cell Rep.* 25, 3422.e7–3434.e7.

Cirillo, D., Agostini, F., Klus, P., Marchese, D., Rodriguez, S., Bolognesi, B., et al. (2013). Neurodegenerative diseases: quantitative predictions of protein-RNA interactions. *RNA* 19, 129–140. doi: 10.1261/rna.034 777.112

Cirillo, D., Blanco, M., Armaos, A., Buness, A., Avner, P., Guttman, M., et al. (2016). Quantitative predictions of protein interactions with long noncoding RNAs. *Nat. Methods* 14, 5–6. doi: 10.1038/nmeth.4100

Cirillo, D., Livi, C. M., Agostini, F., and Tartaglia, G. G. (2014). Discovery of protein-RNA networks. *Mol. Biosyst.* 10, 1632–1642. doi: 10.1039/c4mb00099d

Cunningham, C. L., Martinez Cerdeno, V., Navarro Porras, E., Prakash, A. N., Angelastro, J. M., Willemsen, R., et al. (2011). Premutation CGG-repeat expansion of the Fmr1 gene impairs mouse neocortical development. *Hum. Mol. Genet.* 20, 64–79. doi: 10.1093/hmg/ddq432

Dasti, A., Cid-Samper, F., Bechara, E., and Tartaglia, G. G. (2019). RNA-centric approaches to study RNA-protein interactions in vitro and in silico. *Methods* doi: 10.1016/j.ymeth.2019.09.011 [Epub ahead of print].**

Disney, M. D., Liu, B., Yang, W.-Y., Sellier, C., Tran, T., Charlet-Berguerand, N., et al. (2012). A small molecule that targets r(CGG)(exp) and improves defects in fragile X-associated tremor ataxia syndrome. *ACS Chem. Biol.* 7, 1711–1718. doi: 10.1021/cb300135h

Drozd, M., Delhaye, S., Maurin, T., Castagnola, S., Grossi, M., Brau, F., et al. (2019). Reduction of Fmr1 mRNA levels rescues pathological features in cortical neurons in a model of FXTAS. *Mol. Ther. Nucleic Acids* 18, 546–553. doi: 10.1016/j.omtn.2019.09.018

Entezam, A., Biacsi, R., Orrison, B., Saha, T., Hoffman, G. E., Grabczyk, E., et al. (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. doi: 10.1016/j.gene.2007.02.026

Fernandez, J. J., Martinez, R., Andujar, E., Perez-Alegre, M., Costa, A., Bonilla-Henao, V., et al. (2012). Gene expression profiles in the cerebellum of transgenic mice over expressing the human FMR1 gene with CGG repeats in the normal range. *Genet. Mol. Res.* 11, 467–483. doi: 10.4238/2012.March.1.4

Foote, M., Arque, G., Berman, R. F., and Santos, M. (2016). Fragile X-associated tremor/ataxia syndrome (FXTAS) motor dysfunction modeled in mice. *Cerebellum* 15, 611–622. doi: 10.1007/s12311-016-0797-6

Glineburg, M. R., Todd, P. K., Charlet-Berguerand, N., and Sellier, C. (2018). Repeat-associated non-AUG (RAN) translation and other molecular mechanisms in fragile X tremor ataxia syndrome. *Brain Res.* 1693, 43–54. doi: 10.1016/j.brainres.2018.02.006

Gokden, M., Al-Hinti, J. T., and Harik, S. I. (2009). Peripheral nervous system pathology in fragile X tremor/ataxia syndrome (FXTAS). *Neuropathology* 29, 280–284. doi: 10.1111/j.1440-1789.2008.00948.x

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

Greco, C. M., Hagerman, R. J., Tassone, F., Chudley, A. E., Del Bigio, M. R., Jacquemont, S., et al. (2002). Neuronal intranuclear inclusions in a new cerebellar tremor/ataxia syndrome among fragile X carriers. *Brain* 125, 1760–1771. doi: 10.1093/brain/awf184

Greco, C. M., Soontrapornchai, K., Wirojanan, J., Gould, J. E., Hagerman, P. J., and Hagerman, R. J. (2007). Testicular and pituitary inclusion formation in fragile X associated tremor/ataxia syndrome. *J. Urol.* 177, 1434–1437. doi: 10.1016/j.juro.2006.11.097

Greco, C. M., Tassone, F., Garcia-Arocena, D., Tartaglia, N., Coffey, S. M., Vartanian, T. K., et al. (2008). Clinical and neuropathologic findings in a woman with the FMR1 premutation and multiple sclerosis. *Arch. Neurol.* 65, 1114–1116. doi: 10.1001/archneur.65.8.1114

Green, K. M., Sheth, U., Flores, B. N., Wright, S. E., Sutter, A., Kearse, M. G., et al. (2019). High-throughput screening yields several small-molecule inhibitors of repeat-associated non-AUG translation. *J. Biol. Chem.* 294, 18624–18638. doi: 10.1074/jbc.RA119.009951

Hagerman, P. (2013). Fragile X-associated tremor/ataxia syndrome (FXTAS): pathology and mechanisms. *Acta Neuropathol.* 126, 1–19. doi: 10.1007/s00401-013-1138-1

Hagerman, R. J., and Hagerman, P. J. (2002). The fragile X premutation: into the phenotypic fold. *Curr. Opin. Genet. Dev.* 12, 278–283. doi: 10.1016/s0959-437x(02)00299-x

Hashem, V., Galloway, J. N., Mori, M., Willemsen, R., Oostra, B. A., Paylor, R., et al. (2009). Ectopic expression of CGG containing mRNA is neurotoxic in mammals. *Hum. Mol. Genet.* 18, 2443–2451. doi: 10.1093/hmg/ddp182

Hoffman, G. E., Le, W. W., Entezam, A., Otsuka, N., Tong, Z. B., Nelson, L., et al. (2012). Ovarian abnormalities in a mouse model of fragile X primary ovarian insufficiency. *J. Histochem. Cytochem.* 60, 439–456. doi: 10.1369/0022155412441002

Hukema, R. K., Buijsen, R. A., Raske, C., Severijnen, L. A., Nieuwenhuizen-Bakker, I., Minneboo, M., et al. (2014). Induced expression of expanded CGG RNA causes mitochondrial dysfunction in vivo. *Cell Cycle* 13, 2600–2608. doi: 10.4161/15384101.2014.943112

Hukema, R. K., Buijsen, R. A., Schonewille, M., Raske, C., Severijnen, L. A., Nieuwenhuizen-Bakker, I., et al. (2015). Reversibility of neuropathology and motor deficits in an inducible mouse model for FXTAS. *Hum. Mol. Genet.* 24, 4948–4957. doi: 10.1093/hmg/ddv216

Hunsaker, M. R., Arque, G., Berman, R. F., Willemsen, R., and Hukema, R. K. (2012). Mouse models of the fragile X premutation and the fragile X associated tremor/ataxia syndrome. *Results Probl. Cell Differ.* 54, 255–269. doi: 10.1007/978-3-642-21649-7_14

Hunsaker, M. R., Wenzel, H. J., Willemsen, R., and Berman, R. F. (2009). Progressive spatial processing deficits in a mouse model of the fragile X premutation. *Behav. Neurosci.* 123, 1315–1324. doi: 10.1037/a0017616

Hunter, J., Rivero-Arias, O., Angelov, A., Kim, E., Fotheringham, I., and Leal, J. (2014). Epidemiology of fragile X syndrome: a systematic review and meta-analysis. *Am. J. Med. Genet. A* 164A, 1648–1658. doi: 10.1002/ajmg.a.36511

Iwahashi, C. K., Yasui, D. H., An, H. J., Greco, C. M., Tassone, F., Nannen, K., et al. (2006). Protein composition of the intranuclear inclusions of FXTAS. *Brain* 129, 256–271. doi: 10.1093/brain/awh650

Jain, S., Wheeler, J. R., Walters, R. W., Agrawal, A., Barsic, A., and Parker, R. (2016). ATPase-modulated stress granules contain a diverse proteome and substructure. *Cell* 164, 487–498. doi: 10.1016/j.cell.2015.12.038

Jin, P., Duan, R., Qurashi, A., Qin, Y., Tian, D., Rosser, T. C., et al. (2007). Pur α binds to rCGG repeats and modulates repeat-mediated neurodegeneration in a *Drosophila* model of fragile X tremor/ataxia syndrome. *Neuron* 55, 556–564. doi: 10.1016/j.neuron.2007.07.020

Jin, P., Zarnescu, D. C., Zhang, F., Pearson, C. E., Lucchesi, J. C., Moses, K., et al. (2003). RNA-mediated neurodegeneration caused by the fragile X premutation rCGG repeats in *Drosophila*. *Neuron* 39, 739–747. doi: 10.1016/s0896-6273(03)00533-6

Kiliszek, A., Kierzek, R., Krzyzosiak, W. J., and Rypniewski, W. (2011). Crystal structures of CGG RNA repeats with implications for fragile X-associated tremor ataxia syndrome. *Nucleic Acids Res.* 39, 7308–7315. doi: 10.1093/nar/gkr368

Klus, P., Ponti, R. D., Livi, C. M., and Tartaglia, G. G. (2015). Protein aggregation, structural disorder and RNA-binding ability: a new approach for physico-chemical and gene ontology classification of multiple datasets. *BMC Genomics* 16:1071. doi: 10.1186/s12864-015-2280-z

Krzyzosiak, W. J., Sobczak, K., Wojciechowska, M., Fiszer, A., Mykowska, A., and Kozlowski, P. (2012). Triplet repeat RNA structure and its role as pathogenic agent and therapeutic target. *Nucleic Acids Res.* 40, 11–26. doi: 10.1093/nar/gkr729

Leehey, M. A. (2009). Fragile X-associated tremor/ataxia syndrome: clinical phenotype, diagnosis, and treatment. *J. Investig. Med.* 57, 830–836. doi: 10.2310/JIM.0b013e3181af59c4

Livi, C. M., Klus, P., Delli Ponti, R., and Tartaglia, G. G. (2016). catRAPID signature: identification of ribonucleoproteins and RNA-binding regions. *Bioinformatics* 32, 773–775. doi: 10.1093/bioinformatics/btv629

Loomis, E. W., Sanz, L. A., Chedin, F., and Hagerman, P. J. (2014). Transcription-associated R-loop formation across the human FMR1 CGG-repeat region. *PLoS Genet.* 10:e1004294. doi: 10.1371/journal.pgen.1004294

Marchese, D., Botta-Orfila, T., Cirillo, D., Rodriguez, J. A., Livi, C. M., Fernandez-Santiago, R., et al. (2017). Discovering the 3' UTR-mediated regulation of alpha-synuclein. *Nucleic Acids Res.* 45, 12888–12903. doi: 10.1093/nar/gkx1048

Muslimov, I. A., Patel, M. V., Rose, A., and Tiedge, H. (2011). Spatial code recognition in neuronal RNA targeting: role of RNA-hnRNP A2 interactions. *J. Cell Biol.* 194, 441–457. doi: 10.1083/jcb.201010027

Peier, A. M., McIlwain, K. L., Kenneson, A., Warren, S. T., Paylor, R., and Nelson, D. L. (2000). (Over)correction of FMR1 deficiency with YAC transgenics:

behavioral and physical features. *Hum. Mol. Genet.* 9, 1145–1159. doi: 10.1093/hmg/9.8.1145

Peier, A. M., and Nelson, D. L. (2002). Instability of a premutation-sized CGG repeat in FMR1 YAC transgenic mice. *Genomics* 80, 423–432. doi: 10.1006/geno.2002.6849

Sellier, C., Buijsen, R. A. M., He, F., Natla, S., Jung, L., Tropel, P., et al. (2017). Translation of expanded CGG repeats into FMRpolyG is pathogenic and may contribute to fragile X tremor ataxia syndrome. *Neuron* 93, 331–347. doi: 10.1016/j.neuron.2016.12.016

Sellier, C., Rau, F., Liu, Y., Tassone, F., Hukema, R. K., Gattoni, R., et al. (2010). Sam68 sequestration and partial loss of function are associated with splicing alterations in FXTAS patients. *EMBO J.* 29, 1248–1261. doi: 10.1038/emboj.2010.21

Spencer, C. M., Graham, D. F., Yuva-Paylor, L. A., Nelson, D. L., and Paylor, R. (2008). Social behavior in Fmr1 knockout mice carrying a human FMR1 transgene. *Behav. Neurosci.* 122, 710–715. doi: 10.1037/0735-7044.122.3.710

Su, Z., Zhang, Y., Gendron, T. F., Bauer, P. O., Chew, J., Yang, W.-Y., et al. (2014). Discovery of a biomarker and lead small molecules to target r(GGGGCC)-associated defects in c9FTD/ALS. *Neuron* 83, 1043–1050. doi: 10.1016/j.neuron.2014.07.041

Tartaglia, G. G. (2016). The grand challenge of characterizing ribonucleoprotein networks. *Front. Mol. Biosci.* 3:24. doi: 10.3389/fmolb.2016.00024

Tassone, F., Iong, K. P., Tong, T. H., Lo, J., Gane, L. W., Berry-Kravis, E., et al. (2012). FMR1 CGG allele size and prevalence ascertained through newborn screening in the United States. *Genome Med.* 4:100. doi: 10.1186/gm401

Wenzel, H. J., Murray, K. D., Haify, S. N., Hunsaker, M. R., Schwartzer, J. J., Kim, K., et al. (2019). Astroglial-targeted expression of the fragile X CGG repeat premutation in mice yields RAN translation, motor deficits and possible evidence for cell-to-cell propagation of FXTAS pathology. *Acta Neuropathol. Commun.* 7:27. doi: 10.1186/s40478-019-0677-7

Westergard, T., Jensen, B. K., Wen, X., Cai, J., Kropf, E., Iacovitti, L., et al. (2016). Cell-to-cell transmission of dipeptide repeat proteins linked to C9orf72-ALS/FTD. *Cell Rep.* 17, 645–652. doi: 10.1016/j.celrep.2016.09.032

Willemsen, R., Hoogeveen-Westerveld, M., Reis, S., Holstege, J., Severijnen, L. A., Nieuwenhuizen, I. M., 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. doi: 10.1093/hmg/ddg114