

**Repeat Associated Non-AUG initiated Translation mediates  
neurodegeneration in a Drosophila models of Fragile X-associated  
Tremor Ataxia Syndrome.**

Michelle Frazer

Todd Lab

University of Michigan

Honors Thesis

Department of Cell, Molecular, and Developmental Biology

## **Abstract:**

Fragile X-associated Tremor Ataxia Syndrome (FXTAS) is a neurodegenerative disease that results from a CGG repeat expansion in the 5'UTR of FMR1. Pathogenesis in FXTAS is thought to involve a dominant RNA gain of function mechanism, whereby the CGG repeat mRNA binds to and sequesters specific RNA binding proteins. However, our group has recently discovered that the repeats are also capable of eliciting aberrant translation initiation in the 5'UTR in the absence of an AUG start codon (RAN translation), leading to the production of a polyglycine-containing protein that forms ubiquitinated aggregates in cells and animal models. A critical question that emerges from this work is whether this polyglycine protein contributes directly to toxicity, or whether the neurodegeneration is mediated strictly via RNA toxic mechanisms. To investigate this question, we created strains of drosophila that decouple the potential CGG RNA and protein mediated toxic products. This was achieved by placing the CGG repeat in either the 5'UTR or 3'UTR of a heterologous gene, eGFP. Placement in the 3'UTR precludes RAN translation. To enhance the protein mediated toxic effects, we have inserted an AUG start codon 5' to the repeat, which leads to increased production of the polyglycine protein. As previously reported, expression of a (CGG)<sub>100</sub> repeat in the 5'UTR of eGFP leads to a modest rough eye phenotype with isolated oomatidial expression and a decrease in viability with ubiquitous expression compared to control flies. In lines where the CGG repeat is in the 3'UTR of eGFP, there is very little overt oomatidial degeneration and no effect on viability. In contrast, flies with an ATG codon inserted 5' to the repeat, exhibit an enhanced degenerative eye phenotype and further reduced viability compared to flies lacking this ATG. These studies support a model where aberrant translation of a polyglycine protein in FXTAS contributes significantly to disease pathogenesis.

## **Acknowledgements:**

I would like to extend an enormous thank you to Dr. Peter Todd for giving me the opportunity to work in his lab these past two years, as well as for all of the advice, support, and assistance he's given me over that time.

To Dr. Fang He, another huge thank you, for allowing me to work with you on your projects, and for being so informative, patient, and generally wonderful to work alongside.

I would also like to thank all of the rest of the Todd lab, specifically Amy Krans for being so knowledgeable and helpful; Abbie Renoux for all of her help during my first project in the lab two years ago and then some; and Seok Yoon Oh for all of his assistance and support as well.

Another thank you to Dr. Collins, my co-sponsor, for her advice throughout this project, as well as for being a wonderful and engaging professor.

Thank you as well to Dr. Buttitta, for taking the time to be a reader of this thesis.

Finally, thank you to all of my family and friends for your support.

## **Introduction:**

### **Repeat Expansion Neurodegenerative Disorders**

At least 22 neurodegenerative diseases can be linked back to aberrant expansions within stretches of DNA, despite the relatively recent discovery of microsatellite regions within DNA. Repeat expansion diseases are currently classified into three main categories by their mechanism of action: diseases in which the expansion elicits a loss of expression; those in which exonic repeat expansions result in altered protein function; and a final category in which non-coding RNA elicit a toxic gain of function.

Repeat expansions that trigger a loss of function by eliciting changes in the local chromatin structure surrounding the repeat, leading to transcriptional silencing of the associated gene. Two examples of this subset are Fragile X Syndrome and Friedreich's Ataxia (FRDA); FXS results from the loss of FMRP, and FRDA from the reduced expression of frataxin (Campuzano, 1996). The loss of these proteins as the cause of disease is supported by rare cases of FXS and FRDA that are caused by point mutations rather than repeat expansions (Wang, 1997; McCormack, 2000). In FXS, the expansion of 200 or more repeats leads to hypermethylation at the site of *FMR1*, which then leads to transcriptional silencing and greatly lessened production of FMRP (Sutcliff, 1992). Similarly, FRDA results from a GAA repeat expansion in an intronic region of *FXN*, a gene that encodes the protein frataxin (Campuzano, 1997). In both of these disorders, the resulting symptoms can be traced to the loss of the target protein--cognitive dysfunction seen in FXS is resultant of loss of the synaptically-expressed FMRP, and the degeneration in Friedreich's ataxia is due to the disruption of frataxin's interaction with mitochondria and iron-sulfur clusters. Ultimately, while the mechanism of transcriptional silencing is similar between

diseases in this subset, the symptoms are unique and completely dependent on loss of the affected protein.

In contrast, exonic repeat expansions are inherited in a dominant fashion, suggesting that they elicit a dominant gain of function toxicity rather than haploinsufficiency. These repeats are coded into proteins homopolymeric proteins, typically polyglutamine. There are 9 polyglutamine expansion diseases, including Huntington's disease, Kennedy's disease (spinal and bulbar muscular atrophy), Dentatorubropallidoluysian Atrophy (DRPLA) and six different spinocerebellar ataxias. An additional repeat expansion disorder, oculopharyngeal muscular dystrophy (OPMD) results from a polyalanine expansion in the TATA binding protein. These homopolymeric protein expansions can elicit a toxic gain of function through a number of different mechanisms; the build-up of protein aggregates, transcriptional dysregulation, and neuronal processes such as mitochondrial dysfunction have all been implicated in the pathogenesis of these diseases. Polyglutamine proteins are very aggregate-prone, leading to the formation of insoluble inclusions, the contribution of which to disease is not fully determined (Scherzinger 1997). Studies have been done correlating aggregate formation to increased susceptibility of cell death, through mechanisms of oxidative stress or impairment of the UPS (Hackman 1998); however, other studies have indicated that aggregates could potentially be a protective mechanism, as some results show an inverse correlation between cell survival and propensity to form aggregates (Bowman 2005).

Transcriptional dysregulation is also a potential mechanism for polyglutamine disease toxicity, postulating that the proteins, both in aggregate form and in soluble monomeric form, could interact with transcription factors and transcription machinery in a way that disrupts transcriptional stability of the cells (Gatchel 2005). For example, in Huntington's disease,

mutant huntingtin has been shown to recruit important transcription factors such as TBP into inclusions both *in vitro* and in post mortem samples; likewise, CBP has been found in inclusions in SMBA cell models (Cha 2000). Aberrant proteins are also implicated in a disruption of normal cellular processes such as mitochondrial dysfunction and apoptosis. For example, mutant huntingtin disrupts calcium homeostasis and mitochondrial function, and the cleavage of polyglutamine proteins by calpain and caspases can produce toxic protein fragments (Tang 2005).

However, not all repeats fit easily into these two categories. A third class of repeat expansion disorders contain repeats in putatively non-exonic gene regions but are inherited in a dominant fashion. After early studies failed to demonstrate a role for haploinsufficiency in pathogenesis, it became clear that the repeat expansions are capable of eliciting a toxic gain of function as RNA. A number of degenerative diseases, such as DM1, have been explained by this RNA gain of function theory, which suggests that the non-coding RNA disrupts cellular metabolism through sequestration of critical proteins within the cell. The sequestration of these RNA binding proteins leads to a loss of their normal function, causing dysregulation of RNA splicing and maturation events, and improperly produced downstream proteins. In addition to aberrant splicing, repeat RNA is also found within inclusions in these diseases, along with protein degradation machinery, another possible mechanism of toxicity. There is a good deal of evidence to support alternate splicing as the disease-causing mechanism in myotonic dystrophy. Evidence supporting this includes the finding that mouse lines in which MBNL has been knocked out develop myotonia at around 6 weeks old, and that a number of mRNAs are alternatively spliced, including troponin, insulin receptor, and CIC-1 (Kanadia, 2003). This same group then found that the overexpression of MBNL in a poly(CUG) model of myotonic

dystrophy lead to the reversal of myotonia and missplicing (Kanadia, 2006). This data gives support to the theory that myotonic dystrophy arises from sequestration of MBNL and the resulting aberrant splice patterns of its target RNA. Some evidence suggests that FXTAS pathogenesis is analogous to that of myotonic dystrophy. For example, evidence suggests that Sam68, an RNA binding protein, associates with CGG repeat inclusions, resulting in aberrant splicing of several RNAs (Sellier, 2010). These aberrantly spliced RNAs include *Bcl-xl*, *SMN2*, and *ATP11B*, which have been found to be aberrantly spliced in FXTAS patients as well (Sellier, 2010).

### **Fragile X Tremor/Ataxia Syndrome**

Fragile X Tremor/Ataxia Syndrome (FXTAS) is a neurodegenerative disease related to Fragile X Syndrome (FXS), the most prevalent cause of inherited cognitive impairment, causing severe intellectual disability in roughly 1 in every 1 500 males (Rousseau 1991). Fragile X Syndrome results from the loss of Fragile X Mental Retardation Protein (FMRP), when the gene becomes hypermethylated due to a CGG expansion in the 5'-UTR region of *FMR1*. FMRP is an RNA binding protein that is produced at synapses, and its down-regulation leads to mental retardation, emotional disorders, and some stereotypic physical features. Individuals with FXS typically have a CGG expansion that is greater than 200 repeats (Hagerman RJ, 1999).

In contrast to FXS, FXTAS result from a premutation--a CGG expansion over 55 repeats long, but under the 200 repeats (RJ Hagerman 2001). Rather than impacting an individual developmentally, FXTAS operates as a neurodegenerative disease, affecting approximately 1 in 3000 men over the age of 50 worldwide (Jacquemont 2004). Due to its location on the X-chromosome, FXTAS is primarily found in males and tends to be more severe in males, as the

presence of another X-chromosome with a normal repeat expansion will produce the correct amount of FMRP (Jin 2000).

FXTAS pathology is characterized by a number of neurologic and cognitive abnormalities; most commonly, a progressive intention tremor, gait ataxia, and dementia. Less frequently, patients can experience parkinsonism, peripheral neuropathy, autonomic dysfunction, and cognitive and emotional disturbances (Greco 2006). These clinical symptoms are associated with neuropathologic changes including cortical and subcortical atrophy, and spongiosis of cerebellar and cerebral white matter (Greco 2002, Brunberg 2002). Another constant feature of FXTAS is the presence of ubiquitin positive inclusions in neurons and astrocytes throughout the cortex and deep cerebellum; interestingly, the number of these inclusions correlates with repeat length(Greco 2006).

### **Molecular Pathology of FXTAS**

The differences in phenotype between FXS and FXTAS suggest that, though related, the two diseases must operate under two distinct molecular mechanisms of pathogenesis. The initial pathological cause of FXTAS is the presence of a CGG premutation that is not methylated, and does not initially cause significant mental impairment developmentally. The CGG premutation is unstable, and is prone to expand inter-generationally with maternal transmission; indeed, this expansion is how FXS is passed down from a mother with a premutation to her son (Jin 2000). It is thought that the CGG expansion is able to form a hairpin structure.

In contrast to FXS, pre-mutation repeat expansions do not silence the *FMR1* gene--rather, they actually trigger an over-production of FMR1 mRNA, resulting from the non-methylated repeat and chromatin alterations. Despite this overproduction of FMR1 mRNA, there are still normal to

slightly reduced levels of FMRP, suggesting that FXTAS pathogenesis is not entirely due to insufficient amounts of FMRP (Garcia-Arocena 2010).

The elevated levels of FMR1 mRNA suggested early on that FXTAS might result from a CGG repeat RNA gain of function toxicity as the primary cause of neurodegeneration. Evidence of toxicity caused by the repeat in isolation also suggests inherent toxicity of the CGG repeat, even outside of the FMR1 gene. Handa et al founds that expressing a CGG repeat without FMRP still is toxic to human cell lines (Handa, 2005). Similarly, expressing repeat CGG RNA is ommatidial cells leads to degeneration in *drosophila* eyes, with dsFMRP still being expressed at regular levels (Jin, 2003). A number of proteins have been implicated as potentially sequestered by CGG RNA repeats. The evidence for their having a role in toxicity is thus: A number of specific proteins bind avidly to the RNA repeats. Two in particular, Pur  $\alpha$  and heterogeneous nuclear ribonucleoprotein A2/B1 (hnRNP A2/B1), are capable of partially rescuing CGG repeat associated toxicity when overexpressed in the same model system. Other proteins, including the CUG triplet repeat RNA-binding protein 1 (CUGBP1), and Src-associated substrate during mitosis of 68 kDa (Sam68), do not interact directly with the RNA but are found in inclusions in patients (Garcia-Arocena 2010 and Sellier 2010). The overexpression of proteins such as Pur  $\alpha$  and hnRNP A2 leading to rescue suggest that the loss of function of specific RNA binding proteins is the main component of toxicity of FXTAS, through mechanisms such as loss of chaperone activity or altered RNA maturation due to altered RNA splicing activity (Jin 2007 and Sofola 2007).

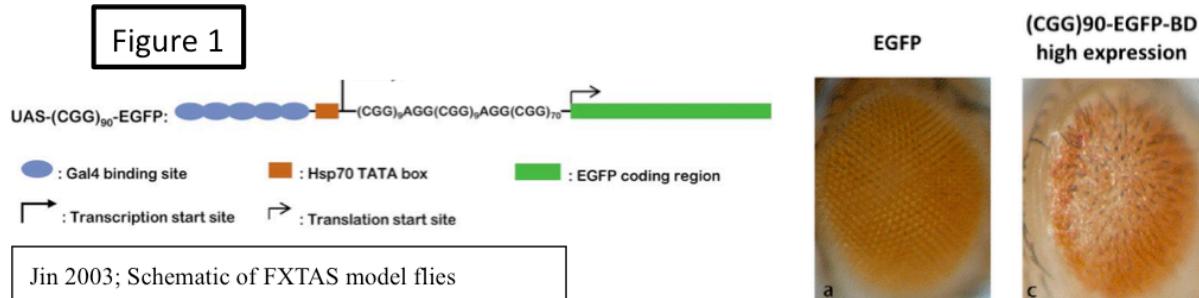
The caveat to this evidence in favor of sequestration, however, is the fact that many of the proteins that successfully rescue FXTAS phenotype when overexpressed also have general neuroprotective effects, and could thus be compensating for some other mechanism. For example

the overexpression of Hsp70 and Pur a contribute significantly to rescue when present in abundance in the cells (Jin 2003). This could be evidence that, rather than sequestration, the CGG RNA could serve as a trigger for some other mechanism that benefits from the presence of heightened cellular regulation. However, despite these alternative mechanisms, the RNA gain of function hypothesis is the prevailing model for FXTAS pathogenesis.

### ***Drosophila* Models of FXTAS**

Our lab uses a *Drosophila* model of CGG repeat mediated neurodegeneration developed by the Jin lab, and characterized in their 2003 *Neuron* paper. Flies were generated by placing a human *FMR1* DNA fragment from a premutation carrier upstream of EGFP. This sequence contains 90 CGG repeats, with AGG triplets at repeat 10 and 20, followed by a roughly 200 base pair flanking sequence. Importantly, the transcriptional start site occurs before the CGG repeat, and the only AUG translational start site is downstream of the repeat. These flies were initially generated to determine whether the repeat in isolation from the rest of the *FMR1* transcript could elicit toxicity.

**Figure 1**



### **Figure 1 Legend**

A. Adapted from Jin et al, 2003. Schematic of the FXTAS model line transgene. A CGG repeat, interspersed with AGG codons, is placed directly upstream of a GFP coding sequence. There is a transcriptional start site directly before the CGG repeat stretch, and a translational start site before the GFP coding region. Gal4 binding sites are upstream of the CGG repeat, and the transgene is under the control of the UAS promoter.

B. A comparison in eye morphology between *drosophila* expressing the EGFP transgene versus those expressing the CGG-GFP high expression gene. There are significant differences in the two eyes, with clear degeneration in the CGG transgene, and little degeneration in the fly expressing only GFP.

In order to observe the impact of the CGG repeat on neuronal viability, we and other researchers have utilized the GAL4/UAS system to direct expression of the repeats exclusively in ommatidial cells in the fly eye. The CGG-GFP sequence was placed under control of the UAS promoter. These flies were then crossed to a second line of flies that express Gal4 under a GMR promoter that is specific to retinal cells; thus, the CGG repeat was only expressed within the retinal cells and neurodegeneration would be visible in the animals' eyes. Lines were also created in which the repeat expression was weaker, as well as only 60 CGG repeats rather than 90. In addition, the repeat was also expressed under the control of a number of other promoters, including the ELAV promoter, in which expression is pan-neuronal; the Act5C promoter, resulting in ubiquitous expression; and the dpp promoter, which expressed the transgene around the boundaries of the imaginal discs and epithelial cells.

The researchers found that the rCGG transgene elicited neurodegeneration that was dosage and repeat length dependent. For moderately expressed CGG60, no toxic effects were observed, whereas in strongly expressed CGG60, a mild rough eye phenotype when expression was limited to the oommatidia and reduced viability when expressed ubiquitously. Moderately expressed CGG90 also resulted in a rough eye phenotype and reduced viability, and the strongly expressed CGG90 showed a very severe rough eye phenotype and larval lethality.

In flies where the transgene was activated after eclosion the CGG repeat elicited progressive degeneration, especially at higher temperatures. They also accumulated ubiquitinated inclusions, much like in human cases of FXTAS; these inclusions were found to co-localize with Hsp70, a chaperone protein. From this experiment, the authors concluded that the CGG RNA was the toxic force driving FXTAS pathogenesis. Because they saw neurodegeneration through this model, they concluded that RNA enough was sufficient to induce toxicity. They saw this as support for an RNA gain of function mechanisms, in which inclusions, such as the Hsp70 aggregates that they characterized, are generated by binding to RNA, which sequesters the proteins with which it interacts.

### **Repeat Associated Non-ATG Translation**

Translational initiation is highly organized event whereby an 80S ribosomal subunit is assembled over a specific start codon (usually AUG) on a messenger RNA (Sonnenberg 2009). In eukaryotes, initiation factors (eIFs) bind to the mRNA's m<sup>7</sup>G cap, thus activating the mRNA for the preinitiation complex (PIC) binding (Sonnenberg, 2009). The PIC is then able to scan the mRNA for an AUG codon in order to begin translation (Sonnenberg, 2009). While there have

been cases discovered of non-AUG initiated translation, in which codons similar to AUG (such as ACG, CUG, or GUG) are able to begin translation by being confused for an AUG, such occurrences are still dependent upon the binding of a methionine tRNA to begin transcription (Touriol 2003).

Recently, however, exciting new evidence has emerged suggesting that, in cases of repeat expansions, translation can be initiated by the repeat itself, through a process called repeat-associated non-ATG (RAN) translation. This process was first characterized in a paper by Zu et al., in which they discovered that mutating the ATG initiation codon just before the 5' CAG expansion did not affect the production of the aberrant polyglutamine protein (Zu 2011). To study this further, they inserted a 6X STOP codon cassette (two stops in each frame) just upstream of a CAG repeat and placed three different tags C-terminal to the repeat to monitor protein expression in all frames. Expression of this construct led to production of proteins in all three reading frames in the absence of an AUG start codon. These products were not present in untransfected cells or in cells treated with cyclohexamide, suggesting that translation is required for the production of these proteins (Zu 2011). Immunofluorescence indicated that all three aberrant proteins could be produced in the same cell (Zu 2011). The group used a variety of methods, including mass spectrometry, generation of novel antibodies, immunoprecipitation, and C-terminal epitope tags, to confirm the identity and presence of the homopolymeric proteins (Zu 2011). Mass spectrometry confirmed that no proteins containing the N-terminal methionine were detected, suggesting either that translation occurs without incorporation of the initiating methionine, or that it is rapidly removed following translation. Transcripts from these constructs co-sedimented with ribosomes, and they confirmed that no splicing occurred on these transcripts to insert an AUG start. Variation of CAG expansion length showed that a certain

length was required to elicit RAN translation and the comparison between CAG expansions and CAA expansions indicate that RAN translation is facilitated by hairpin structures, as CAG forms a hairpin and participates in RAN translation, but CAA, which does not form a hairpin structure, does not result in RAN translation products.

To begin to test the relevance of this mechanism to disease pathogenesis, Zu et al looked at whether expressing the homopolymeric proteins elicited co-expression of apoptosis proteins in transfected cells. In cells expressing all three RAN translation products, they found that significant increases in annexin-V, indicating an upregulation of apoptosis in these cells. Through using an antibody made to detect the predicted RAN translation products, researchers found significant staining in *in vivo* models of SCA8 and DM1 and in DM1 patient derived tissues. Ultimately, this research suggests the possibility that these proteins could participate in disease pathogenesis of disorders commonly thought to be solely determined by an RNA gain of function mechanism.

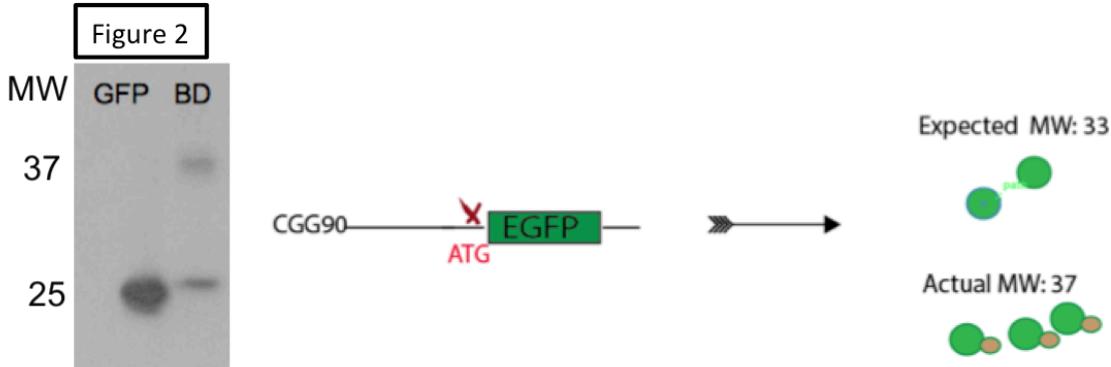
Recently, two further studies were published extended these initial findings by Zu et al to other repeat expansion diseases. The newly discovered C9ORF72 GGGGCC hexanucleotide repeat, implicated in FTD and ALS, was investigated as possibly undergoing RAN translation. As hairpins formed from repeats are more likely to undergo RAN translation, investigators modeled the secondary structure of the hexanucleotide repeat, finding that it is energetically favorable for this repeat to form a hairpin structure (Ash 2013, and Mori 2013). This lead the researchers to generate antibodies for all of the possible RAN translation products from such a repeat, which would produce three different two amino acid alternating copolymers--glycine-alanine, glycine-arginine, and glycine-proline (Ash 2013). Impressively, these antibodies detected RAN translation products through slot blot and immunohistochemical analysis in

samples from C9orf72 ALS and FTD patients, but not from other ALS/FTD patients or controls. While the contribution of these RAN translation products to toxicity is not yet completely defined, the presence of these aberrant protein products does seem to be a novel neuropathological feature in this disease. This study raises a number of interesting questions, particularly in regards to how toxic these products are, and how frequently they occur in pathogenesis of repeat expansion diseases. As FXTAS is a disorder caused by a repeat expansion that is able to form a strong hairpin in mRNA form, it is possible that RAN translation could be occurring within those with FXTAS.

**Unpublished data from our group indicating the production of a RAN translation product in association with CGG repeats.**

*Presence of a Higher Molecular Weight GFP product in CGG-GFP flies*

Shortly after obtaining the CGG90-GFP fly model from the Jin Lab, Dr. Todd noted that the GFP expressed in these flies formed aggregates when expressed. This unexpected finding led to a series of studies demonstrating that CGG repeats can trigger RAN translation. First, lysates from CGG-GFP flies demonstrate the presence of a higher molecular weight GFP product. While the expected molecular weight of GFP protein is around 25 kDa, in the FXTAS model flies there was a species observed through immunoblotting against GFP that was of a higher molecular weight, at about 37 kDa. Mass spectrometry performed on this higher MW GFP band indicated that it was a fusion protein of GFP and sequences within the FMR1 5'UTR. The reading frame and size of the band indicated that the repeat was being translated into a polyglycine protein, strongly suggesting that it is a RAN translation product, as there are no upstream translational starts in frame with polyglycine.



## Figure 2 Legend

A. A western blot probing for GFP, using protein lysates from *drosophila* with either a GFP transgene or Jin's CGG-GFP construct. While the GFP transgene produces only GFP protein at the expected size, in the CGG-GFP flies, a distinct band at a higher molecular weight is produced.

B. Our lab's model to explain these results. While there should only be GFP transgene produced because there is no translational start codon before the CGG repeat, we hypothesize that RAN translation results in some portion of the produced protein to be in the form of a higher molecular weight species, in which the GFP is fused to a polyglycine protein.

Further experiments were done to show that without the CGG repeat expansion, the FMRpolyG protein does not form. Deletions of repeat stretches in the 5'UTR severely decreased and even abolished production of the RAN translation product (data not shown). This indicates that translation of this protein initiates at a non-AUG codon within the CGG repeat, possibly through ribosomal scanning and stalling.



### Figure 3 Legend

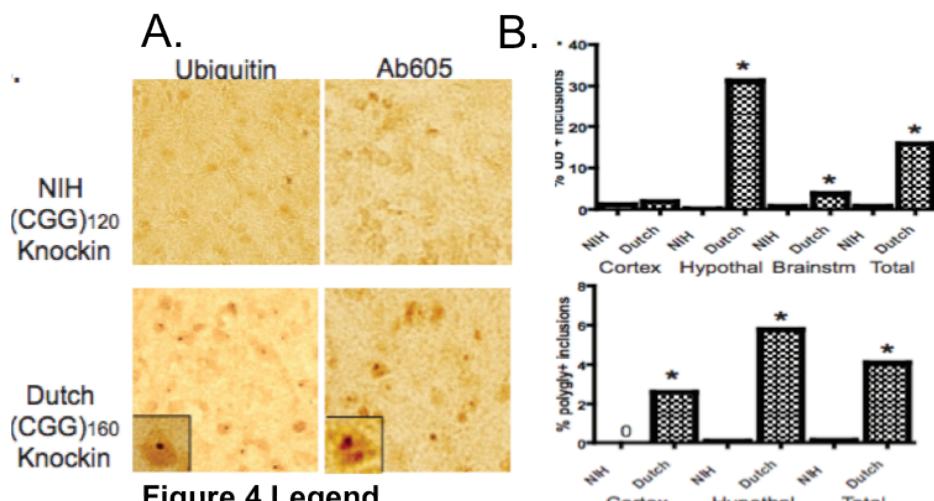
Mass spectrometry data of the observed higher molecular weight product. The data shows a polyglycine protein fused to GFP coding sequence.

### *Presence of aberrant translation product in FXTAS model mice*

In mouse model systems, our lab uses two strains to model FXTAS--a mouse line we refer to as the NIH mice, and another we call the Dutch line. Each line replicates certain behavior and molecular aspects of FXTAS, but interestingly only the Dutch line produces the RAN translation product. When probed by another antibody created to recognize the higher molecular weight product (which we named FMRpolyG) sections of cortex from the two different mouse lines show very different staining patterns, with the Dutch line showing intranuclear inclusions, but with little staining in the NIH mouse line.

This discrepancy between the two models of the same disease can be explained through the aberrant translation product hypothesis, through examining the sequences of the two lines. While the both lines exhibit a CGG of premutation length that has been knocked in to the mouse FMR1 gene, the NIH line begins with a translational stop codon (TAA) that is 18 base pairs upstream of the beginning of the repeat sequence, which is predicted to abolish the production of the HMW product. Indeed, when the NIH mouse sequence just proximal to the repeat is placed upstream of the GGC repeat sequence in cell culture lines, it abolishes the production of the polyglycine protein. This finding provides evidence for the presence of a RAN translation product that is produced in FXTAS.

**Figure 4**



**Figure 4 Legend**

A. Histology of mouse hippocampal slices of the two different FXTAS model mice, using both an antibody against ubiquitin and the predicted polyglycine sequence. In the NIH mice, there is not significant staining in either antibody; however, in the Dutch mice, there are distinct intranuclear inclusions using both antibodies. This suggests the presence of ubiquitin positive inclusions, partly made up of polyglycine proteins, that are not present in the NIH mouse line. B. Quantification of intranuclear inclusions present in different areas of the NIH and Dutch mice. Clearly, there are significantly more inclusions present in the Dutch mice than the NIH.

**Figure 5**

**A. NIH 5'UTR seq:** TAACGGCCCAAACGGCCC (GGC)<sub>120</sub>

**Dutch 5'UTR seq:** GCGTGCAGCGCGGC (GGC)<sub>160</sub>

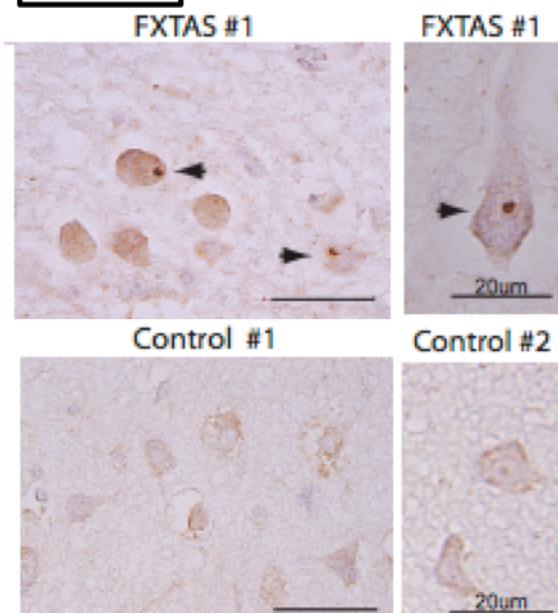
**Figure 5 Legend**

The 5' sequences of the NIH and Dutch mice. The most noticeable difference between the two is the presence of the TAA stop codon directly before the beginning of the repeat, present in the NIH mice, but not the Dutch. We believe that this prevents RAN translation, and explains the difference in molecular pathology between the two mice.

### *Aberrant protein present in FXTAS patients and FXTAS model systems*

To determine whether or not the aberrant polyglycine protein is produced in human cases of FXTAS, a monoclonal antibody (2C13) was developed based on the predicted sequence of the human aberrant product. This antibody, which was tested for specificity based on its ability to recognize a recombinant protein generated in bacteria, was tested on western blots of cerebellar lysates from FXTAS and control brain samples. The predicted protein product from the aberrantly translated product is predicted to be approximately 11.5 kDa, with an N-terminal polyglycine stretch and a 42 amino acid stretch out of frame with the FMR1 protein coding region (Todd, unpublished data). In FXTAS brain lysates, a 12 kDa protein was observed by the 2C13 antibody that was not observed in the control lysates. When cerebellar sections were immunostained with this antibody, the FXTAS sections exhibited perinuclear aggregates, while control tissue did not. FXTAS sections also consistently exhibited ubiquitinated intranuclear inclusions in the hippocampus, consistent with previously published data.

**Figure 6**



**Figure 6 Legend**

Representative samples of FXTAS and control hippocampus and frontal cortex, stained with an antibody against the FMRpolyG. In the FXTAS samples there are large intranuclear inclusions, not present in the control sample.

**Experimental goal**

The work that the Todd lab and I have done thus far indicate the presence of a RAN translation protein product in various models of FXTAS. In flies, lysates taken from ommatidial cells expressing the CGG<sub>100</sub>-GFP transgene show the presence of a higher molecular weight GFP product when probed for GFP protein, as well as significant GFP inclusion formation in ommatidial cells. The mass spectrometry data for this protein indicates that it is a fusion protein between GFP and a polyglycine (GGC) protein. We found that staining FXTAS mouse and human tissue with antibodies created against this FMRpolyG protein showed the presence of inclusions positive for this protein, which were not present within control tissue.

Given the presence of this protein within FXTAS model systems and patients but not in control tissue, it would seem that it is part of FXTAS pathogenesis. The central question of this thesis is whether this aberrantly produced FMRpolyG protein is able to produce toxicity.

Based on the prevalence of this protein across all different models of FXTAS, as well as the numerous other neurodegenerative diseases with protein-induced toxicity, I hypothesized that production of this protein does lead to toxicity. To test this, I used a number of parameters to characterize fly lines expressing CGG RNA and the FMRpolyG protein at different levels, in order to determine FMRpolyG's ability to modulate toxicity.

# Methods

## Fly crosses

*Drosophila Melanogaster* lines were crossed and maintained at 25 degrees on standard food unless otherwise indicated. All crosses set up between the transgene and the GMR-GAL4 driver were done at 30 degrees.

## Generation of fly lines

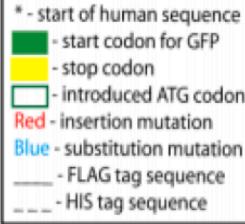
All fly constructs were derived from cell culture plasmids with the replacement of (CGG)88 with (CGG)100 from FXTAS patient fibroblast cell genomic DNA, and were inserted into restriction sites BglII and XbaI of a pUAST vector. Transgenic flies were made by standard p-element insertion (The Bestgene Inc., Chino Hills, CA), and were mapped and balanced with standard genetic techniques. Briefly, flies were crossed to double balancer (Cyo/+; TM3/+) flies first, and flies with both of the balancers and transgene were selected to breed. Following this cross, chromosomal location of the transgene was determined based on which balancer it migrated. Flies were maintained with standard food at 25 degrees (Todd, 2013). Transgenes used are as follows (also see Table 1):

- 1) 5'UTR: CGG<sub>100</sub> repeat followed by GFP open reading frame. The GFP translational start codon is in-frame with the GGC polyglycine.
- 2) 5'ATG: An introduced ATG codon followed by the CGG<sub>100</sub> repeat and in-frame GFP open reading frame.
- 3) 5' Stop: A translational stop in the 5'UTR just before the pure CGG repeat. Following this is the GFP reading frame.

- 4) 3'UTR: The coding sequence for GFP, followed by a translational stop. After the stop, in the 3'UTR, is the CGG<sub>100</sub> repeat.

Table 1

**Sequence of *Drosophila* constructs (glycine frame)**

		High MW GFP?
5'UTR (UAS (CGG) <sub>n</sub> +1 GFP)	*CTCAGCTCGTTCGGTTCACTTCCGTGGAGGGCCGCCTTGAGCGGGCGGCCGACGGCGAGCGCGGGCGGCCGTTGACGGAGGCCG	
	CTGCCAGGGGGCGTGCAGCGCG-(CGG) <sub>n</sub> -CTGGGCTCGAGCGCCCGAGCCCACCTCTCGGGGGCGGCTCCGGCGTAGCAGGGCTGAAG	PRESENT
	AGACCATG	
5'ATG (UAS ATG FLAG (CGG) <sub>n</sub> +1 GFP)	ATG GACTACAAAGACGATGACAGGCCTCA*...GCCGACGGCGAGCGCGGGCGGCCGCGTGA CGGAGGCGCCGCTGCCAGGGGGCGTCGGCAGC	
	GCG-(CGG) <sub>n</sub> ...GAAGGGCTGAAGAGACCATG	PRESENT
5'stop (UAS stop @-12 (CGG) <sub>n</sub> +1 GFP)	*...GCCGACGGCGAGCGCGGGCGGCCGCGTGA CGGAGGCGCCGCTGCCAGGGGGCTAGCGGCAGCGCG-(CGG) <sub>n</sub> ...GAAGGGCTGAAGAGACCATG	ABSENT
3'UTR (UAS GFP stop (CGG) <sub>n</sub> +1 6xHIS)	ATG (GFP seq)TAA...*...GCCGACGGCGAGCGCGGGCGGCCGCGTGA CGGAGGCGCCGCTGCCAGGGGGCGTCGGCAGCGCG-(CGG) <sub>n</sub> -CTGGGCTCGAGCCATCATCACCATCACCACTAG	ABSENT
		

**Table 1 Legend**

The sequences of each of the transgenic fly lines, as well as whether or not they produce the aberrant protein product.

## RNA Extraction

Groups of twenty 1-2 day old adult flies were anesthetized using CO<sub>2</sub> and decapitated. Total RNA was then extracted from the fly heads through a previously described Trizol protocol (Todd 2010). Briefly, tissues were homogenized in 1 mL of Trizol per 50 to 100 mg of tissue. 0.2 mL of chloroform was then added per 1 mL Trizol used for the phase separation step. Samples were shaken vigorously for 20 seconds and then incubated at room temperature for 2 to 3 minutes. Samples were centrifuged at 12,000 x g for 15 minutes at 4 degrees. Following centrifugation, the solution separates into a bottom red, phenol-chloroform phase, an interphase, and an upper, clear aqueous phase, which contains the RNA. The aqueous phase was then removed and placed

into another centrifugation tube. To precipitate, the aqueous phase was mixed with 0.5 mL isopropanol per 1 mL Trizol reagent used. This solution was incubated for 10 minutes at room temperature, followed by centrifugation at 4 degrees, producing an RNA pellet. Finally, the pellet was washed with 1 mL 75% ethanol, dried to isolate pellet, and then the RNA was resuspended and quantified.

## RT PCR

Following RNA extraction, RT PCR was performed in order to create cDNA to be used in QPCR reactions. For this, the iScript (Bio-Rad) cDNA synthesis protocol was used. Briefly, 500ng of RNA was mixed with 4 uL of 5X mix, 2uL of Oligo dT primer, a 1 uL of iScript reverse transcriptase and DEPC treated water to 20 uL total. This was then mixed lightly and run through the iScript cDNA synthesis PCR program. The iScript cDNA PCR protocol is as follows: 5 minutes at 25 degrees, 42 degrees for 30 minutes, then 5 minutes at 85 degrees for enzyme inactivation.

## QPCR analysis

The resulting cDNA was subjected to real-time PCR analysis with gene specific primers as follows: eGFP: (Forward: 5'-TCTTCTTCAAGGACGACGGCAACTAC-3', Reverse: 5'-GTACTCCAGCTTGCCCCAGGATGT-3'), RPL32: (Forward: 5'-GTTGTGCACCAGGAACCTTCTTGAATCCG-3', Reverse: 5'-CTTCCAGCTTCAAGATGACCATCCGC-3'). PCR analysis was performed using the iQ SYBR Green Supermix in a myiQ Single Color RTPCR system (BioRad). All runs included a standard dilution curve representing at least 2x and 0.01X the RNA concentration utilized for all primer sets to ensure linearity. Further, equivalent efficiency of individual primer sets was

confirmed prior to data analysis. The levels of (CGG)100-containing eGFP mRNA were normalized to RPL32 mRNA for each sample run and expressed as a ratio of levels found to individual lines (fold control expression). All samples were run in triplicate and all data represent at least three independent experiments (Todd, 2013).

### **Phenotype assessment**

Phenotype severity was assessed using a validated scale of neurodegeneration seen in fly eyes (Pandey, 2007). The scale assigns one point based on each disturbance found in ommatidial structure and organization, bristle organization, ommatidial fusion and pitting, de-pigmentation, and scarring (See Table 2). When degeneration occurred in more than 5% of the eye, 2 extra points were given; when present in over 50% of the eye, 4 points were given. Typically >50 flies were scored, and averages were created of all of the genotypes so that they could be graphed and compared. Data was analyzed using a Kruskal Wallis ANOVA and Mann-Whitney U test for non-parametric comparisons.

<b>Feature</b>	<b>Absent</b>	<b>Present</b>
Supernumerary IOB	0	1
Abnormal bristle orientation	0	1
Ommatidial fusion	0	1
Ommatidial pitting	0	1
Disorganization of ommatidial array	0	1
Retinal collapse	0	1
Scarring	0	1

**Table 2 Legend**

The criteria used for assigning eye phenotype scores to the different lines.

Adapted from Pandey et al. 2007.

Table 2

## Fly Eye Imaging

Young flies were anesthetized with CO<sub>2</sub> or frozen. At 80X magnification, pictures of eyes were taken with a Leica MZ APO microscope, and captured with Leica DFL320 digital camera. For fluorescent image capture of GFP aggregates, flies were frozen and placed on a glass coverslip. An inverted Olympus 1X71 scope was used to take images by epifluorescence, and images were then processed in Slidebook 4.0 software.

## Fly Histology/Immunohistology

1-2 day old flies were taken, anesthetized and decapitated, and their heads were immediately frozen in OCT media. They were then cryosectioned transversely. The sections were post-fixed

for 15 minutes in 4% paraformaldehyde, permeabilized with 0.1% triton-X, and processed for immunohistochemistry as previously described (Pandey 2007). Antibodies were not necessary to visualize GFP in the CGG90-EGFP flies. For ubiquitin visualization, a Millipore antibody was used at 1:250 dilution.

### **Western Blot Analysis**

Western blots were prepared as previously described and developed on film (Todd and Mack 2000). Briefly, protein was extracted by homogenizing fly heads in RIPA buffer with protease inhibitors. Equal amounts were then boiled with 4X SDS Dye and run on a 12% SDS polyacrylamide gel. After PVDF transfer at 30 amps for 12 hours at room temperature, blots were incubated with antibodies to GFP(1:1000) and tubulin (1:5000).

### **Viability assessments**

Analysis of relative eclosion rates were performed as previously described (Lanson et al., 2011). Each transgene was balanced over a marker chromosome (CyO). If the transgene elicited no toxicity, then 50% of progeny should have the CyO marker and 50% express the transgene. >200 flies of each genotype were scored over multiple crosses. The relative % progeny carrying the transgene were expressed as a % of total eclosed flies. These numbers were then compared using a Fischer exact test to determine statistical significance.

# Results

## Creation of novel FXTAS fly lines

Four new fly lines were generated with the goal of decoupling the RNA and protein toxicity from one another. A first line was created in which the GGC<sub>100</sub> stretch was followed by the coding region for EGFP, with the translational start site directly preceding the EGFP ORF; this strain should produce both a GFP protein and a RAN translated polyglycine-GFP fusion protein. The second transgene was created to produce high translational expression of the polyglycine protein; its sequence is ATG-FLAG-GCC<sub>100</sub>-GFP, with the ATG driving translation of the GGC repeat in frame with the GFP. Next, a line was created in which the GGC<sub>100</sub> stretch is placed in the 3'UTR of GFP; GFP-STOP-GGC<sub>100</sub>. Finally, a stop codon was inserted 5' of the GGC repeat, to prevent RAN translation of the polyglycine product. The purpose of these final lines is to ensure that, while the CGG repeat mRNA is present, its protein product will not be produced, thus providing a test for whether or not the aberrant protein product can be implicated in FXTAS toxicity. To demonstrate this empirically, we transfected cells with a construct expressing GFP-STOP-GGC<sub>100</sub> with a FLAG tag after the repeat. Expression of this construct was compared to vectors expressing the other constructs described above. Serial probing of these blots with antibodies to GFP, FLAG and Actin revealed the presence of a FLAG positive band only in constructs expressing the ATG-FLAG CGG-GFP construct, despite high level expression of the gene in all cell lines as measured by GFP production.

Figure 7

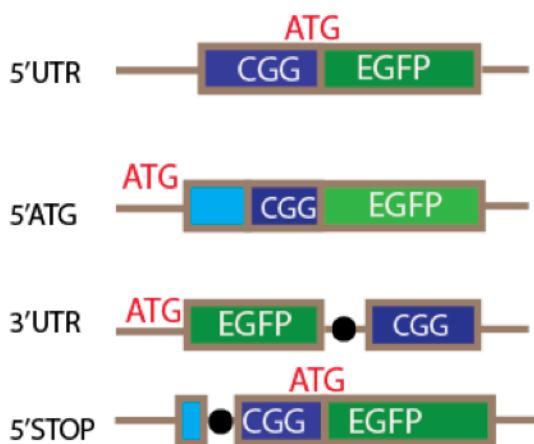


Figure 7 Legend

A schematic of the three new fly lines. The first is comparable to the original *drosophila* model of FXTAS, with a transcriptional start followed by a CGG repeat, a translational start, and then the GFP coding sequence. The next line has a 5'-ATG prior to the CGG repeat, driving the expression of the GFP fusion protein. The final line has the repeat in the 3' UTR, preceded by a translational stop, preventing the RAN translation fusion protein.

Figure 8

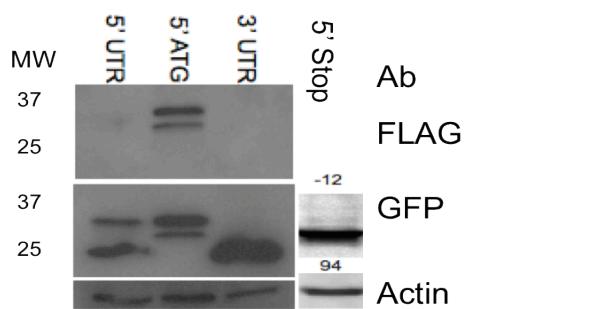


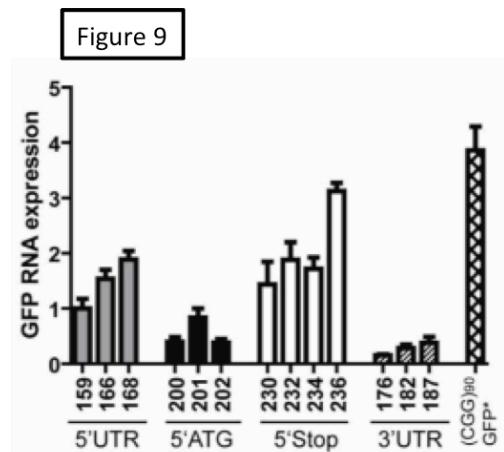
Figure 8 Legend

Western blot probing for the presence of GFP, FLAG, and Actin in cell lines transfected with a GFP-STOP-GGC<sub>100</sub> with a FLAG tag after the repeat, as well as the other transgenic lines described above. FLAG was only detected in the 5'ATG, despite high expression of the GFP mRNA in all lines.

In order to measure neurodegeneration in these lines, we took advantage of the UAS/GAL4 system. The transgenes described above were placed under the control of the UAS promoter, and their expression was directed to the by placing GAL4 under the GMR promoter, which is retinal-cell specific. This allows us to visualize degeneration through retinal cell death and disorganization. The fly lines were bred and maintained at 30 degrees.

### Quantification of GFP mRNA

To determine the levels of GFP mRNA transcript produced by each of the three lines, cDNA was synthesized via RT-PCR, and the levels of transcript were quantified through QPCR. While placing the CGG expansion in the 5' UTR moderately increased transcript, we chose lines to analyze to ensure that the level of transcript production was as close as possible. While the placement of the CGG repeat in the 5' UTR produced more transcript, the two lines with the repeat located in the 5' UTR produced less translated GFP.



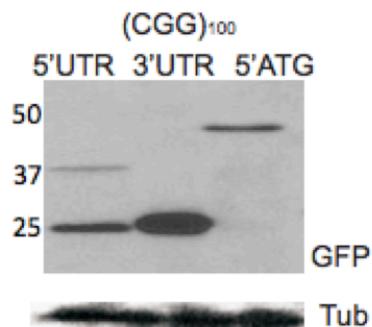
**Figure 9 Legend**

Quantification of GFP mRNA in each of the drosophila lines. The 5'Stop produces GFP transcript at levels more comparable to the other lines than does the 3'UTR, controlling for that difference.

### **RAN translation product is present in FXTAS model lines**

To assess whether or not these new lines were undergoing RAN translation, we probed the flies for the presence of a higher molecular weight (HMW) GFP fusion protein found in cell lines. For this, we performed a western blot analysis on proteins extracted from fly heads. The blot was serially probed first with an antibody for GFP and then secondarily with an antibody to Tubulin and developed using standard film. As anticipated based on previous studies, CGG<sub>100</sub>-EGFP expressing flies demonstrated two products by western blot- one correlating with translational initiation at the AUG for GFP and one correlating with RAN translational initiation 5' to the repeat. In contrast, in flies where the CGG repeat was placed in the 3'UTR, Only the band correlating with AUG initiation at GFP was observed. In flies where an AUG in an appropriate Kozak sequence was inserted upstream of the repeat, we observe production of a larger GFP-polyglycine fusion protein, consistent with the inclusion of the FLAG tag and some additional sequence preceding the repeat. Interestingly, usage of this upstream AUG largely precludes initiation at the repeat by RAN translation or initiation at the canonical AUG for GFP. These results indicate that aberrant translation is supported in *drosophila* and suggests that it is at least partially dependent on a ribosomal scanning initiation mechanism.

**Figure 10**



**Figure 10 Legend**

The western blot of the lysates from the three new lines. The 5'ATG line produces only the higher molecular weight fusion protein (at a slightly higher weight due to the FLAG tag). The 3' UTR line produces only the pure GFP at its normal size. The 5'UTR line produces predominantly GFP, but also produces a modest amount of HMW GFP fusion protein.

#### **Transgenic lines producing the aberrant fusion protein showed neurodegeneration**

Because the transgene was directed to retinal cells specifically, neurodegeneration was quantifiable through direct observation of the fly eyes. Each eye was assigned a score based on the levels of cell death, ommatidial structure, and bristle arrangement, with higher scores indicating more degeneration and toxicity.

The 5'ATG lines exhibited significant degeneration of the retinal cells, indicating that the high expression of the fusion protein was highly toxic. The 5'UTR line showed moderate degeneration, while the 5'Stop and 3'UTR showed little evidence of degeneration. At least 50 flies from each line were assigned scores, which were averaged to reflect a typical score for each line. These scores reflect the levels of degeneration observed visually. These results suggest that production of the polyglycine protein correlates significantly with toxicity.

Figure 11

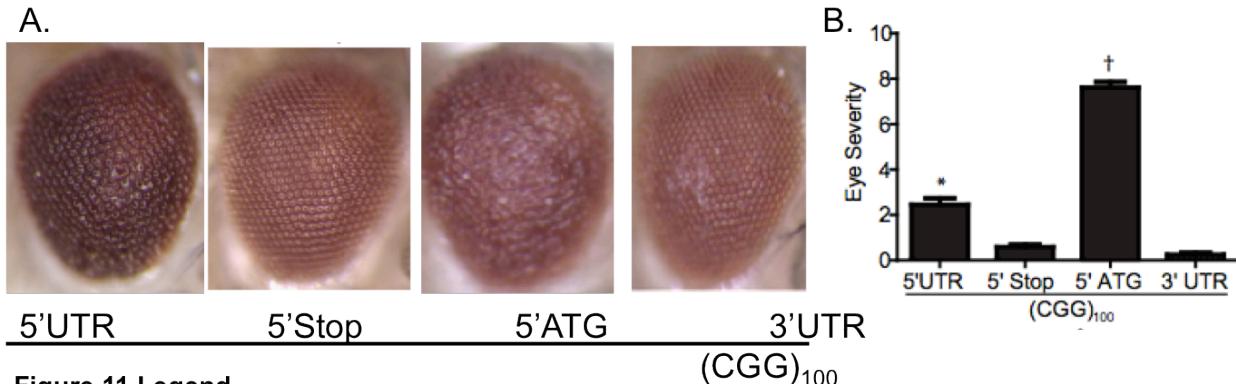


Figure 11 Legend

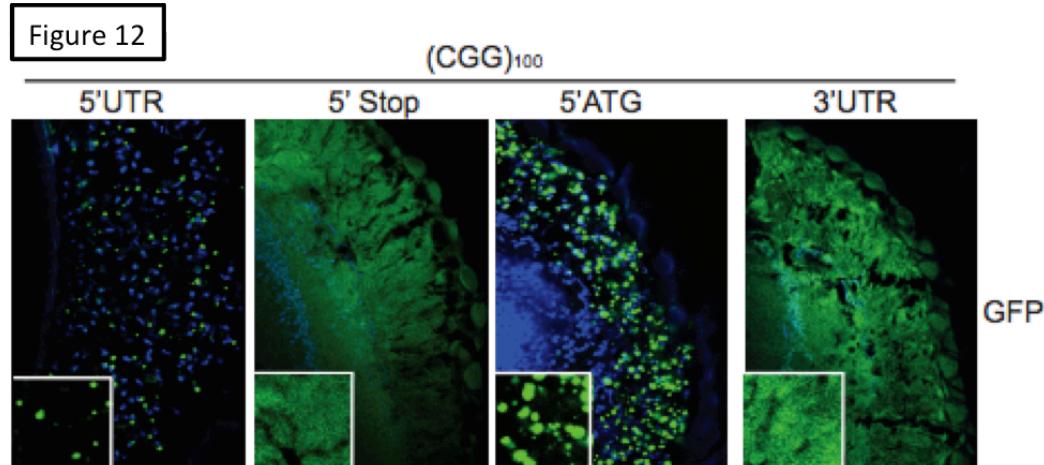
A. A representative image from each of the three lines. All flies were kept at 30 degrees, and imaged shortly after eclosion. The 5'ATG flies exhibit the most severe degeneration, and the 5' CGG flies have a moderate rough eye phenotype. The 3'UTR flies do not exhibit significant degeneration.

B. A graphical representation of the average phenotype score for each genotype. Individual flies were assigned an eye score based on a standardized set of criteria described in table 2. The pattern observed corroborates the eye imaging data and pattern discussed above.

## Inclusion formation

A hallmark of human FXTAS and animal models of FXTAS is the presence of intranuclear inclusions within affected areas. To determine whether or not these new fly lines replicated this aspect of the disease, we examined the eyes for GFP inclusions both through transverse sections and confocal imaging, and whole-eye imaging with epifluorescence. Epifluorescence imaging showed that the number of GFP + inclusions directly correlated with the amount of HMW produced. The 5'ATG line produced a large number of GFP punctate inclusions, and 5'UTR produced a moderate amount, while the 5'Stop and 3'UTR produced none at all. In order to determine whether or not these GFP inclusions colocalize with ubiquitin, similar to the ubiquitin-positive inclusions in FXTAS, we performed a colocalization with a ubiquitin antibody on transverse slices of retinal tissue, and analyzed them through confocal imaging. In both 5'ATG and 5'UTR, the vast majority of GFP inclusions colocalized with ubiquitin. In the 3'UTR and

5' Stop lines, the diffuse GFP does not aggregate and form inclusions. This suggests that the GFP itself does not aggregate; rather, the addition made to it to produce the HMW GFP species contains is what has the propensity for aggregation. We believe that this aggregate-prone species is the translated CGG repeat, which results in a polyglycine protein when it is in-frame with GFP.



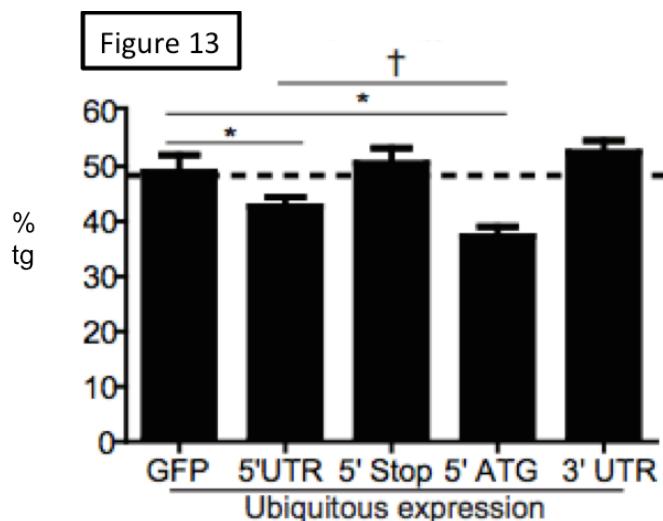
### Figure 12 Legend

Confocal imaging of cross sections from each line. GFP aggregation increases with increasing expression of the FMRpolyG protein.

### Viability Assessment

As a second readout for repeat associated toxicity, we determined whether or not ubiquitous expression of the three different transgenes had an impact on fly viability. We crossed flies with transgenes under control of the UAS promoter to those expressing act5-GAL4, which results in ubiquitous expression of the transgene. For each line, the transgene was balanced over the CyO balancer. If the transgene elicits no toxicity, 50% of the resulting progeny should have the CyO marker, and 50% express the transgene. Through this, we found that placing the CGG repeat in the 5' UTR resulted in a decrease in the percentage of progeny produced with the transgene,

indicating that the transgene conferred increased toxicity to the flies and thus decreased viability. Placing the ATG in front of the CGG repeat in the 5' UTR enhanced toxicity and decreased viability, while the line in which the CGG repeat was in the 3' UTR following a translational stop experienced no decreased viability. This provides more evidence that producing the polyglycine protein results in enhanced toxicity of the CGG repeat.



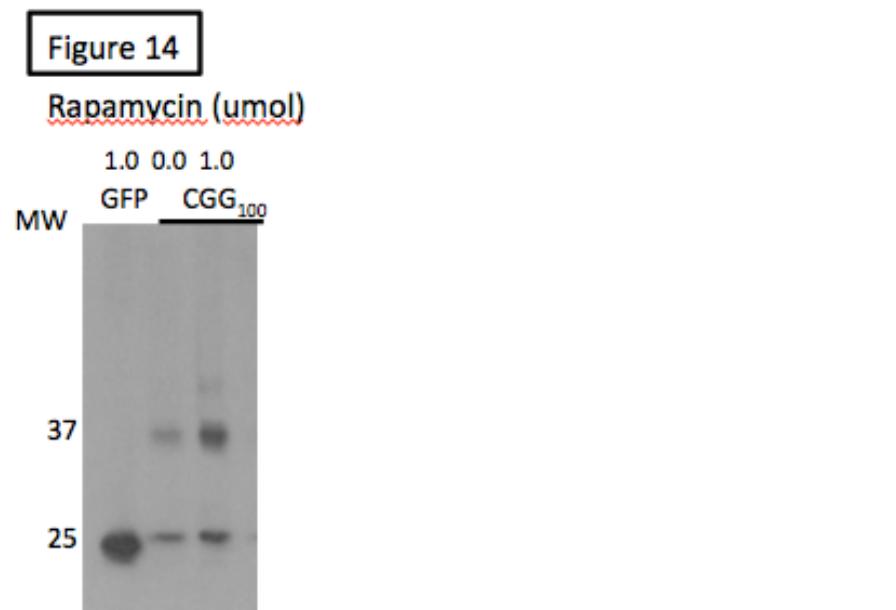
### Figure 13 Legend

Viability data from each transgenic line, as determined by percentage of progeny containing the transgene. The 5'UTR and 5'ATG have reduced viability at a significant level.

### Rapamycin treatment is ineffective in clearance of the FMRpolyG protein

Numerous studies have suggested that rapamycin has neuroprotective effects in a number of neurodegenerative disease models, through its ability to inhibit the mTOR pathway (Berger, 2006). The prevailing theory regarding its ability to rescue neurodegeneration is that it stimulates the autophagy pathway, thus clearing toxic protein aggregates. We thus decided to test rapamycin's effect on flies with a CGG<sub>100</sub>-GFP transgene (analogous to the 5'UTR line).

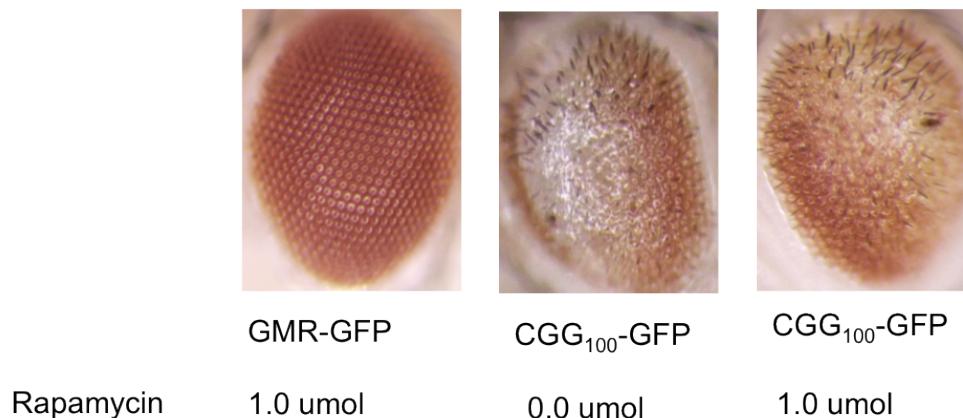
CGG<sub>100</sub>-GFP and GMR-GFP flies were bred and eclosed on food with or without rapamycin at 1.0 uM concentration at 25 degrees. Protein lysates were probed for the presence of the higher molecular weight GFP protein, and fly eyes were imaged and scored as described above. From the western blot, we found that there was no decrease in higher molecular weight product as we had expected. There was also no significant difference in eye phenotype between CGG<sub>100</sub>-GFP lines with or without rapamycin. Thus, rapamycin seemed ineffective in reducing aberrant protein production and aggregation, as well as ineffective on rescuing the degenerative phenotype.



**Figure 14 Legend**

Western blot probing for GFP protein from lysates of GMR-GFP flies and CGG<sub>100</sub>-GFP flies that have been bred and eclosed on different amounts of rapamycin. Quantification of this data (not shown) does not show a significant difference in production of the higher molecular weight protein depending on rapamycin presence.

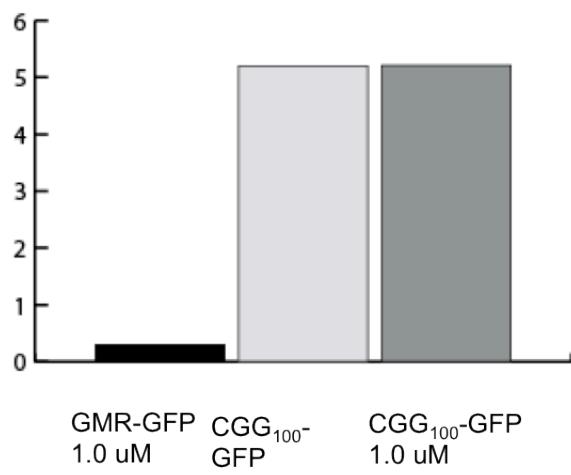
Figure 15



**Figure 15 Legend**

Representative images from the different genotypes and conditions, showing no change in CGG-GFP toxicity due to rapamycin.

Figure 16

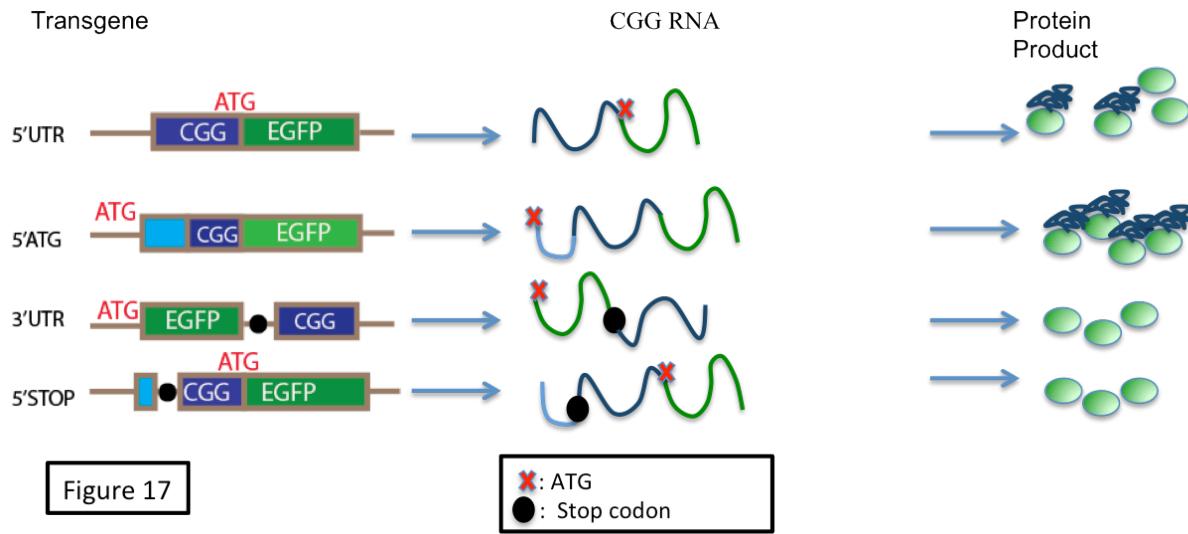


**Figure 16 Legend**

Quantification of eye phenotypes from each group, using the scoring system described above.

## **Discussion**

The results provided above challenge the predominant view of FXTAS toxicity; namely, that it is strictly caused by an RNA gain of function mechanism. The fly lines we generated successfully decouple the production of CGG RNA from the production of CGG RAN translated protein, as well as provide a model in which the polyglycine fusion protein is definitively expressed. Our model of the three lines suggests that the difference in sequence leads to the production of different proteins--in the case of the line with the repeat in the 3'UTR directly after a translational stop, no polyglycine-GFP fusion protein is produced. When the ATG drives production of the CGG repeat, virtually the only species produced is the higher molecular weight fusion protein. In the more traditional FXTAS model, however, in which the CGG is placed in the 5' UTR followed by an ATG translational start, we propose that a mix of these two species is present. This last model seems to be relatively representative of the way in which FXTAS works in mammalian cells.



**Figure 17 Legend**

A schematic of each of the transgene's protein production. In all cases, mRNA is produced, but in only some of the cases a fusion FMRpolyG protein is produced.

Importantly, through characterization of these flies, we discovered striking differences between toxicity of the different lines. The differential levels of toxicity followed the predicted levels of HMW fusion protein produced, with high expression resulting in severe degeneration, modest expression leading to moderate degeneration, and no expression resulting in no degeneration. This shows that, at least in *drosophila*, the aberrantly expressed protein is able to significantly mediate toxicity.

Similarly, it also seems that production of CGG mRNA in these lines is not as crucial to toxicity as the RNA gain of function model would suggest. In addition to the differential levels of degeneration, viability data also indicates that somehow the protein produced enhances the toxicity of the repeat, as CGG RNA expressed ubiquitously did not significantly decrease viability, while ubiquitous expression of the RAN translation product did. This data is the first to provide evidence of a direct role for RAN translation in FXTAS pathogenesis, as well as dissociate CGG repeat RNA from possible polyglycine protein production.

Following the establishment of the toxicity of this protein, one of the next tasks at hand is to determine what about this protein is toxic. In many neurodegenerative disorders that involve the formation of aggregations that appear to be toxic, impairment of the UPS is implicated in toxicity, along with cellular/oxidative stress. As neurons are particularly vulnerable to cellular stress and disruption to homeostasis, mechanisms that aim to maintain normalcy within the cells, such as autophagy and the UPS, are crucial, and often implicated in neuronal disease and dysfunction (Wong 2010). With this in mind, it would be valuable to further investigate the role that protein quality control and degradation pathways play in FXTAS. While in many diseases upregulation of protein degradation systems, such as upregulation of autophagy through treatment with rapamycin, has been shown to be effective in attenuating the toxicity of aggregative species in animal models. Indeed, previous studies and work in our own lab demonstrates that overexpression of specific protein chaperones, such as HSP-70, partially alleviates CGG repeat associated toxicity. This appears to be a Ubiquitin proteasome specific effect, however, as our preliminary data in FXTAS suggests that genetic impairment of autophagic pathways does not exacerbate CGG associated toxicity and Rapamycin treatment is not effective in reducing toxicity (Todd 2010). Further investigation into this area might be productive in determining how the RAN translation products produce toxicity. If there is somehow an impairment to the autophagy system, induction of autophagy through rapamycin would not have an effect on inclusion toxicity, and may even increase it, as an increase in autophagosome formation without clearance results in further accumulation of proteins and cellular stress (Bove 2011). Investigating whether or not the presence of these aberrantly translated proteins somehow impair one of the many pathways neurons use to maintain homeostasis could be a fruitful area for discovering the mechanism of toxicity.

Determining whether or not the presence of the protein causes an increase in cellular stress may also provide insight into its toxicity. There are many ways in which a cell can respond to stress, and observing cells expressing these repeat proteins for specific signs of cellular stress may help determine which pathways are disrupted or activated by the presence of the polyglycine/polyalanine proteins. For example, looking for an upregulation of apoptotic markers or heat shock proteins could help determine exactly what mechanisms are at work to cause degeneration in these flies. Comparing changes in our 5'ATG line and the 3'UTR line could be particularly insightful, as it would help to elucidate what mechanisms are at work when proteins are produced in comparison simply to repeat RNA. These comparisons are also useful in that they could help to determine the relative contribution of repeat RNA to FXTAS; while this data provides evidence of the toxic nature of an aberrant RAN translation product, there is still a significant amount of data discussed in the introduction that implicates CGG RNA in a gain of function toxicity mechanism. It has been shown that this mRNA is able to sequester proteins that will very likely have downstream impacts on normal cellular processing. Thus, there are still many reasons to believe that RNA plays some part in toxicity; the difficult part will be delineating the roles of the two toxic species, and the extent to which both of them contribute to toxicity.

The concept of RAN translation itself is a very new one, only brought to the attention of the scientific community in 2010 by Zu et al. Because of its newness, few researchers studying trinucleotide expansion diseases have really explored it as an option as a possible aspect of pathogenesis. There are a wide range of repeat expansion diseases for which this mechanism might be applicable, and future research will need to be done to explore this as a possibility. It would be extremely interesting if this method of translation turns out to be a common occurrence

in neurodegenerative diseases, especially if the proteins produced from it play a role in toxicity in other diseases as well. Perhaps in the future, treatments could be proposed based on inhibiting this mechanism. As RAN translation appears to be dependent on the formation of an mRNA hairpin, this may prove to be a target of therapeutic research. In addition to being an important aspect of disease, it is a distinct possibility that RAN translation may have other roles within cells even in non-disease states, which may provide further insight into the way in which ribosomes and other translational machinery works.

Ultimately, this data presents a new aspect of FXTAS pathology, and is only a small aspect of our lab's larger body of work that suggests that the canonical description of FXTAS as solely an RNA gain of function disorder is only one aspect of the disease. We will continue to examine this aspect of the disease and how it contributes to our understanding of FXTAS and neurodegenerative disorders as a whole. Hopefully, the research that has been done and will be done in the future by our lab will be informative to the neurodegenerative field as a whole, and pave the way for new treatments and solutions to the problems they pose.

## References:

1. Mandel, J. L. Instability of a 550Base Metylation Abnormal Pair in DNA Fragile X Syndrome. **252**, 1097–1102 (2013).
2. Sutcliffe, J. S. *et al.* DNA methylation represses FMR-1 transcription in fragile X syndrome. *Human molecular genetics* **1**, 397–400 (1992).
3. Campuzano, V. *et al.* Friedreich's Ataxia: Autosomal Recessive Disease Caused by an Intronic GAA Triplet Repeat Expansion. *Science* **271**, 1423–1427 (1996).
4. Bové, J., Martínez-Vicente, M. & Vila, M. Fighting neurodegeneration with rapamycin: mechanistic insights. *Nature reviews. Neuroscience* **12**, 437–52 (2011).
5. Wong, E. & Cuervo, A. M. Autophagy gone awry in neurodegenerative diseases. *Nature neuroscience* **13**, 805–11 (2010).
6. Ash, P. E. a *et al.* Unconventional Translation of C9ORF72 GGGGCC Expansion Generates Insoluble Polypeptides Specific to c9FTD/ALS. *Neuron* **77**, 639–46 (2013).
7. Brunberg, J. a *et al.* Fragile X premutation carriers: characteristic MR imaging findings of adult male patients with progressive cerebellar and cognitive dysfunction. *AJNR. American journal of neuroradiology* **23**, 1757–66 (2002).
8. Gatchel, J. R. & Zoghbi, H. Y. Diseases of unstable repeat expansion: mechanisms and common principles. *Nature reviews. Genetics* **6**, 743–55 (2005).
9. Tassone, F. *et al.* FMRP expression as a potential prognostic indicator in fragile X syndrome. *American journal of medical genetics* **84**, 250–61 (1999).
10. Campuzano, V. *et al.* Frataxin is reduced in Friedreich ataxia patients and is associated with mitochondrial membranes. *Human molecular genetics* **6**, 1771–80 (1997).
11. Manuscript, A. NIH Public Access. **57**, 825–829 (2010).
12. Greco, C. M. *et al.* Neuropathology of fragile X-associated tremor/ataxia syndrome (FXTAS). *Brain : a journal of neurology* **129**, 243–55 (2006).
13. Wang, Y. C., Lin, M. L., Lin, S. J., Li, Y. C. & Li, S. Y. Novel point mutation within intron 10 of FMR-1 gene causing fragile X syndrome. *Human mutation* **10**, 393–9 (1997).
14. McCormack, M. L. *et al.* Frataxin point mutations in two patients with Friedreich ' s ataxia and unusual clinical features. 661–664 (2000).

15. Handa, V. *et al.* Long CGG-repeat tracts are toxic to human cells: implications for carriers of Fragile X premutation alleles. *FEBS letters* **579**, 2702–8 (2005).
16. Sellier, C. *et al.* Sam68 sequestration and partial loss of function are associated with splicing alterations in FXTAS patients. *The EMBO journal* **29**, 1248–61 (2010).
17. Kanadia, R. N. *et al.* A muscleblind knockout model for myotonic dystrophy. *Science (New York, N.Y.)* **302**, 1978–80 (2003).
18. Trost, B. M. *et al.* PNAS. 20001 (2006).
19. Kanadia, R. N. *et al.* Reversal of RNA missplicing and myotonia after muscleblind overexpression in a mouse poly(CUG) model for myotonic dystrophy. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 11748–53 (2006).
20. Davies, S. W. *et al.* Formation of neuronal intranuclear inclusions underlies the neurological dysfunction in mice transgenic for the HD mutation. *Cell* **90**, 537–48 (1997).
21. Scherzinger, E. *et al.* Huntingtin-encoded polyglutamine expansions form amyloid-like protein aggregates in vitro and in vivo. *Cell* **90**, 549–58 (1997).
22. Hackam, a S. *et al.* The influence of huntingtin protein size on nuclear localization and cellular toxicity. *The Journal of cell biology* **141**, 1097–105 (1998).
23. Bowman, A. B., Yoo, S.-Y., Dantuma, N. P. & Zoghbi, H. Y. Neuronal dysfunction in a polyglutamine disease model occurs in the absence of ubiquitin-proteasome system impairment and inversely correlates with the degree of nuclear inclusion formation. *Human molecular genetics* **14**, 679–91 (2005).
24. Cha, J.-H. J. Transcriptional dysregulation in Huntington’s disease. *Trends in Neurosciences* **23**, 387–392 (2000).
25. Tang, T. *et al.* Disturbed Ca 2<sup>+</sup> signaling and apoptosis of medium spiny neurons in Huntington’s disease. (2005).
26. Rousseau, F., et al. Direct Diagnosis by DNA analysis of the Fragile X Syndrome of mental retardation. *The New England Journal of Medicine* **325**, 1673–1681 (1991).
27. Population, P. C. *et al.* Penetrance of the Fragile X – Associated Tremor / Ataxia Syndrome in a. **291**, (2013).
28. Jin, P. & Warren, S. T. Understanding the molecular basis of fragile X syndrome. *Human molecular genetics* **9**, 901–8 (2000).
29. Greco, C. M. *et al.* Neuronal intranuclear inclusions in a new cerebellar tremor/ataxia syndrome among fragile X carriers. *Brain : a journal of neurology* **125**, 1760–71 (2002).

30. Garcia-Arocena, D. & Hagerman, P. J. Advances in understanding the molecular basis of FXTAS. *Human molecular genetics* **19**, R83–9 (2010).
31. Jin, P. *et al.* RNA-mediated neurodegeneration caused by the fragile X premutation rCGG repeats in Drosophila. *Neuron* **39**, 739–47 (2003).
32. Jin, P. *et al.* Pur alpha binds to rCGG repeats and modulates repeat-mediated neurodegeneration in a Drosophila model of fragile X tremor/ataxia syndrome. *Neuron* **55**, 556–64 (2007).
33. Sofola, O. a *et al.* RNA-binding proteins hnRNP A2/B1 and CUGBP1 suppress fragile X CGG premutation repeat-induced neurodegeneration in a Drosophila model of FXTAS. *Neuron* **55**, 565–71 (2007).
34. Sonenberg, N. & Hinnebusch, A. G. Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell* **136**, 731–45 (2009).
35. Touriol, C. *et al.* Generation of protein isoform diversity by alternative initiation of translation at non-AUG codons. *Biology of the Cell* **95**, 169–178 (2003).
36. Zu, T. *et al.* Non-ATG – initiated translation directed by microsatellite expansions. 1–6 (2010).doi:10.1073/pnas.1013343108/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1013343108
37. Mori, K. *et al.* The C9orf72 GGGGCC Repeat Is Translated into Aggregating Dipeptide-Repeat Proteins in FTLD/ALS. *Science (New York, N.Y.)* **1335**, (2013).
38. Pandey, U. B. *et al.* HDAC6 rescues neurodegeneration and provides an essential link between autophagy and the UPS. *Nature* **447**, 859–63 (2007).
39. Todd, P. K. & Mack, K. J. Sensory stimulation increases cortical expression of the fragile X mental retardation protein in vivo. *Brain research. Molecular brain research* **80**, 17–25 (2000).
40. Berger, Z. *et al.* Rapamycin alleviates toxicity of different aggregate-prone proteins. *Human molecular genetics* **15**, 433–42 (2006).
41. Todd, P. K. *et al.* Histone deacetylases suppress CGG repeat-induced neurodegeneration via transcriptional silencing in models of fragile X tremor ataxia syndrome. *PLoS genetics* **6**, e1001240 (2010).
42. Hagerman, R. J. *et al.* Intention tremor, parkinsonism, and generalized brain atrophy in male carriers of fragile X. *Neurology* **57**, 127–130 (2001).

