Fmr1 translationally activates stress-sensitive mRNAs encoding large proteins in oocytes and neurons

Ethan J. Greenblatt[1,2], and Allan C. Spradling[1]

[1]*Howard Hughes Medical Institute Research Laboratories*

Department of Embryology, Carnegie Institution for Science

3520 San Martin Dr.

Baltimore, Maryland 21218 USA

[2]future address:

Department of Biochemistry and Molecular Biology

University of British Columbia

Vancouver, British Columbia, Canada

Author for correspondence (e-mail: spradling@carnegiescience.edu)

Running Head: FMRP compensates for size dependent translational inhibition

Keywords: Fmr1, FMRP, neuron, oocyte, ribosome profiling, RNP, translation,

Corresponding Author: Dr. Allan C. Spradling

Tel. 410-246-3015 Fax. 410-243-6311

Email: spradling@carnegiescience.edu

Abstract

Mutations in Fmr1 are the leading heritable cause of intellectual disability and autism spectrum disorder. We previously found that Fmr1 acts as a ~2-fold activator of translation of large proteins in Drosophila oocytes, in contrast to its proposed role as a repressor of translation elongation. Here, we show that genes associated with autism spectrum disorders tend to be dosage-sensitive and encode proteins that are larger than average. Reanalysis of Fmr1 KO mouse cortex ribosome profiling data demonstrates that autism-associated mRNAs encoding large proteins exhibit a concordant reduction in ribosome footprints, consistent with a general role for Fmr1 as a translational activator. We find no evidence that differential ribosomal pausing affects translational output in Fmr1-deficient Drosophila oocytes or mouse cortex. Furthermore, long Fmr1 target transcripts are preferentially enriched in stress granules upon acute stress. Our data thus identify a critical role for Fmr1 in promoting the translation of long, stress-sensitive, autism-associated mRNAs.

Introduction

Mutations in the Fmr1 gene, encoding the RNA binding protein FMRP, lead to fragile X syndrome (FXS) and fragile X primary ovarian insufficiency (FXPOI), causes of intellectual disability (ID), autism spectrum disorder (ASD), and primary ovarian insufficiency (Hagerman et al. 2017; Sullivan et al. 2011). Neurons and oocytes control the translation and transport of mRNAs using ribonucleoprotein particles (RNPs), a subset of which contain FMRP (Barbee et al. 2006; Rosario et al. 2016; Antar et al. 2005). FMRP binds to hundreds of neuronal mRNAs, including those that encode chromatin remodeling enzymes, microtubule adaptors, synaptic scaffolding proteins, and ubiquitin ligases (Darnell et al. 2011; Sawicka et al. 2019). Morever, FMRP targets are non-randomly enriched in genes associated with ID and ASD disorders (Darnell et al. 2011; Ascano et al. 2012). The association of FMRP with a subset of mRNAs

suggests that FMRP recognizes mRNAs containing sequence features that determine FMRP binding, which evidently include a significant number of ID/ASD genes. Binding determinants may include G-quadruplex structures (Vasilyev et al. 2015; Phan et al. 2011; Darnell et al. 2001), U-rich structures (Chen et al. 2003; Dolzhanskaya et al. 2003), or other specific sequence motifs (Ascano et al. 2012; Zalfa et al. 2005; Ray et al. 2013) (reviewed in (Anderson et al. 2016). However, a random set of neural transcripts matched to the abundance and size of FMRP target mRNAs was also significantly enriched in ID/ASD genes, consistent with relatively non-specific binding (Ouwenga and Dougherty, 2015). In contrast to the idea that FMRP acts nondiscriminately, *Drosophila* Fmr1 is required specifically for the efficient translation of genes encoding large proteins (Greenblatt and Spradling, 2018), and subsequent studies found that mammalian FMRP preferentially binds to long mRNAs (Sawicka et al. 2019; Li et al. 2020).

Why did a system using FMRP evolve to control the expression (Greenblatt and Spradling 2018; Das Sharma et al. 2019) and/or mRNA localization (Dictenberg et al. 2008) of ID/ASD and other genes? One possible explanation is that FMRP-bound transcripts are subject to a form of activity-dependent or homeostatic gene regulation that is essential for aspects of neurodevelopment and synaptic plasticity commonly perturbed in ID/ASD. FMRP associates with polysomes (Feng et al. 1997), and neural FMRP-bound transcripts exhibit reduced translation elongation rates when transferred into *in vitro* translation reactions (Darnell et al. 2011), suggesting a role for FMRP in the repression of translation elongation. FMRP-deficient neurons exhibit an exaggerated form of activity-dependent neuronal plasticity known as mGluR5-dependent long-term depression (LTD), which is thought to be due to a role for FMRP in the repression of activity-dependent protein production (Bear et al. 2004; Bear 2005). Inconsistent with this model, however, is the finding by multiple groups that blocking activity-dependent translation with small molecule inhibitors does not rescue exaggerated mGluR5-

dependent LTD in FMRP-deficient neurons (Hou et al. 2006; Park et al. 2008; Ronesi et al. 2012). Because activity-dependent translation is not required for FMRP-related LTD defects to appear, FMRP appears to affect neuronal function in ways that extend beyond the regulation of activity-dependent translation. FMRP neurons exhibit a large number of alterations, including changes in mRNA transport (Dictenberg et al. 2008), metabolic signaling (Sharma et al. 2010), dendritic spine maturation/pruning (Rudelli et al. 1985; Comery et al. 1997), and homeostatic plasticity (Zhang et al. 2018; Bülow et al. 2019).

The enrichment of a diverse set of ID/ASD genes among FMRP targets suggests that these targets share common properties that makes their post-transcriptional regulation unique in the RNA granule-rich cytoplasm of neurons and oocytes. (Li et al. 2020; Sawicka et al. 2019; Darnell et al. 2011). Emerging analyses of the contents of repressive RNPs, including stress granules and heat shock granules, have shown that these structures are significantly enriched for long mRNA transcripts (Padrón et al. 2019; Khong et al. 2017); the average length of stress granule-enriched transcripts is about ~2-fold longer than non-enriched transcripts (Khong et al. 2017). Stress granules form through multivalent interactions among and between proteins and RNA molecules (Molliex et al. 2015). Longer mRNAs contain more potential sites for promiscuous interactions with non-specific RNA binding proteins, including translational repressors, and other RNA molecules. The control of gene expression using an RNA granule based-system of gene regulation may, therefore, bias its effects increasingly towards genes encoding longer transcripts.

We hypothesized that cells relying on RNA granule-based post-transcriptional regulation use FMRP-mediated upregulation to minimize the effects of transcript length biases on gene regulation, in order to maintain translation from large mRNAs that encode critical homeostatic functions (Greenblatt and Spradling 2018). We show here that ASD associated-genes tend to encode mRNAs that are longer than typical mRNAs, even when compared to

other neuronally-expressed transcripts. By comparing ribosome profiling data from experiments conducted using FMRP-deficient *Drosophila* oocytes (Greenblatt and Spradling, 2018) or mouse cortex (Das Sharma et al. 2019), we find that FMRP has an ancient, conserved function to promote rather than repress translation from transcripts encoding large proteins. Dozens of genes encoding large proteins exhibit concordantly reduced ribosome footprints in FMRP-deficient cells including many that are implicated in ASD/ID syndromes. By analyzing the distribution of ribosomes along the length of FMRP target genes, we find that the reduction of ribosome footprints of FMRP targets is not due to an alleviation of ribosome stalling (Das Sharma et al, 2019) as would be expected from the FMRP repressor model. Together, these data show that FMRP activates the translation of bound transcripts, by a magnitude of ~2-fold or less, likely by increasing the rate of translation initiation.

Consistent with the model that FMRP-dependent translational activation offsets translational repression by stress granule-associated proteins, FMRP targets become highly enriched in stress granules upon acute stress of cultured cells. Finally, FMRP ASD/ID targets appear to be highly sensitive to gene dosage effects caused by haploinsufficiency, suggesting that weak translational activation of ASD/ID genes by FMRP is essential to ensure normal neural development and function. Understanding how FMRP activates ASD/ID-associated target translation will speed the development of pharmacological interventions that reverse the underlying symptoms of FXS and potentially other autism spectrum disorders.

Results

ASD/ID genes tend to encode large proteins

In order to determine whether characteristics of transcripts encoded by ASD/ID genes distinguish them from other neuronally-expressed genes, we examined general properties of all transcripts expressed in the juvenile mouse cortex (Das Sharma et al. 2019). We classified

SFARI Class I and Class II (Abrahams et al. 2013; Pereanu et al. 2017) autism genes as "ASD genes" and found that they differ from other neuronally-expressed genes in several ways. (1) ASD/ID genes as a group encode proteins much larger than average. The median coding sequence (CDS) length of ASD/ID genes is 2.2-fold greater than that of all neuronally-expressed genes as a whole (3003 vs. 1334 bps, Fig. 1A). The average ASD/ID gene encodes a protein of 1,234 amino acids, vs. 589 amino acids for all neuronally-expressed proteins averaged together. (2) ASD/ID genes contain large introns. The median sum intron length of ASD/ID genes is 3.8-fold greater than that of all neuronally-expressed genes (78.3 kbps vs. 20.6 kbps, Fig. 1B), consistent with prior studies showing that ASD/ID genes tend to encode long genes (King et al. 2013; Zhao et al. 2018). (3) The untranslated regions of ASD/ID genes are slightly longer than average; the median 5'UTR length of ASD/ID genes is 1.6-fold longer than average (250 bps vs. 160 bps, Fig. 1C) and the median 3'UTR length 1.6-fold longer (1703 vs. 1053 bps, Fig. 1D). In contrast, we found that the expression of ASD/ID genes was similar to other neural genes. We found that the mRNA levels (Fig. 1E) and translation levels (Fig. 1F) of ASD/ID genes in the mouse cortex were not significantly different from all neural genes as a whole.

Fmr1 has an ancient, conserved function to promote the translation of large proteins

We previously found that *Drosophila* Fmr1 functions to promote the translation of large proteins in oocytes (Greenblatt and Spradling, 2018) and the levels of a number of large proteins are reduced in Fmr1 deficient *Drosophila* oocytes and neurons. This includes the 5,322 amino acid E3 ubiquitin ligase Poe/Ubr4, the 2,053 amino acid snRNA 3' processing factor IntS1 in *Drosophila* oocytes (Greenblatt and Spradling 2018), and the 3,719 amino acid protein kinase A regulator rugose/NBEA in the *Drosophila* mushroom body (Sears and Broadie 2018). This finding differed from a previous model of FMRP function, which concluded from *in*

vitro reconstitution experiments that FMRP inhibits translation by promoting ribosome stalling (Darnell et al. 2011). Mammalian and *Drosophila* Fmr1 contain the same protein domains and appear to share a common function, as the expression of human Fmr1 rescues neural defects in *Drosophila* Fmr1 mutants (Coffee and Broadie et al. 2012). We hypothesized that two important differences may limit the ability of *in vitro* experiments to reconstitute the *in vivo* activity of Fmr1: (1) *in vitro* translation assays conducted in rabbit reticulocyte extracts may lack essential regulatory factors, such as RNA granule-associated proteins, and differ in their metabolic signaling status, thereby making them a poor model for Fmr1-mediated translational control *in vivo*, and (2) the addition of translation initiation inhibitors required for the measurement of elongation rates by ribosome run-off in *in vitro* assays will mask any differences in translation initiation rates between control and experimental samples. The initiation of translation is generally thought to be the rate-limiting step in most systems (12); thus, changes to translation elongation rates are only relevant if translation initiation is not rate-limiting.

In order to test whether Fmr1 has a similar function in *Drosophila* oocytes and the mouse cortex, we compared the results of ribosome profiling experiments performed using these respective tissues (Greenblatt et al. 2018, Das Sharma et al. 2019) but re-analyzed here (Table S1) from primary reads with identical bioinformatic pipelines (see Methods). Consistent with Greenblatt et al. 2018, we found that transcripts containing long CDS were concordantly translationally downregulated in Fmr1 RNAi oocytes as compared to controls (Fig. 2A). The median CDS length of genes significantly translationally decreased in Fmr1 KD oocytes was 3.1-fold longer than the median length of all oocyte expressed genes (4262 bps vs. 1372 bps respectively, Fig. 2B). We observed a strikingly similar pattern in the Fmr1 KO mouse cortex. Genes encoding large proteins were also concordantly translationally downregulated to a similar magnitude (Fig. 2C), and translationally downregulated genes had 3.4-fold longer CDS

lengths as compared to all cortex-expressed genes (4797 bps vs 1401 bps respectively, Fig. 2D).

We found that Fmr1 CLIP targets were concordantly translationally downregulated in the mouse cortex (Fig. 2E), consistent with previous findings (Das Sharma et al. 2019). This strongly suggests that the primary function of Fmr1 is to increase the translation of its bound targets, while the translational upregulation of genes in the absence of Fmr1 represents secondary effects resultant from Fmr1 loss. Indeed, as observed by (Das Sharma et al. 2019), many genes that are translationally activated in the absence of Fmr1 contain 5'TOP motifs that are regulated by mTOR signaling, which is over-activated in Fmr1 mutant neurons (Sharma et al. 2010). The long CDS lengths of genes translationally downregulated in Fmr1 neurons and oocytes, and the long CDS lengths of autism genes generally, suggests that a major reason for the previously established association of Fmr1 with ASD/ID genes is due to Fmr1's function to activate genes encoding large proteins. Indeed, we found that SFARI ASD genes were concordantly translationally downregulated in the Fmr1 KO cortex (Fig. 2F). These data suggest that the phenotypes associated with Fmr1 loss represent the combined effects of the reduced expression of dozens of individual ASD genes. Consistent with an ancient role for Fmr1 in the regulation of these targets, we found that a statistically significant subset of *Drosophila* orthologs of SFARI Class I and Class II ASD genes translationally downregulated in the Fmr1 KO mouse cortex (12/34) were also translationally downregulated in *Drosophila* oocytes ($p=2.3x10^{-9}$, Chi-squared test, Table S2).

While genes translationally downregulated in Fmr1-deficienct cells were highly enriched for those encoding large proteins, the magnitude of the reduction in translation did not appear to be length-dependent in either *Drosophila* oocytes or the mouse cortex (Fig. 2G,H). These data may be consistent with a "phase transition" phenomena, where genes surpassing a critical threshold related to mRNA length and perhaps other parameters are equivalently

translationally activated. Consistent with Fmr1 acting as a translation factor and not an mRNA stability factor, the reduced translation efficiency of large proteins in Fmr1-deficient cells was nearly entirely due to reduced translation, rather than changes to mRNA levels (Fig. 2I).

Fmr1 does not appear to affect translation rates through ribosome stalling

The reduction in ribosome footprints observed in Fmr1 targets in the absence of Fmr1 could in theory be due to either (1) reduced translation initiation, or (2) increased translation elongation. Prior analyses of *Drosophila* oocytes and neurons (Greenblatt and Spradling 2018; Sears et al. 2019) demonstrated that the steady state proteins levels of several large proteins were reduced in Fmr1-deficient cells, consistent with reduced translation initiation. Das Sharma et al. found that there are fewer ribosome pause/stall sites in Fmr1-deficient neurons as assessed by ribosome profiling (Das Sharma et al. 2019). These stall sites, however, were not specific to Fmr1 targets, suggesting that their appearance is a secondary rather than primary effect of Fmr1 loss (Das Sharma et al. 2019). The presence of a ribosome pause/stall site may or may not effect translation output - if and only if translation elongation is rate-limiting will changes to elongation rates impact overall translation rates. If translation elongation is rate-limiting due to ribosomal stalling, then this should lead to a 5' bias in the abundance ribosome footprints, with more reads aligning to the 5' ends of coding sequences than the 3' ends. Such stalls can be induced in bacteria lacking the factor EFP, which is required for the rapid translation of polyproline tracts (Doerfel et al. 2013; Ude et al. 2013). Elongation was found to be rate-limiting when polyproline-induced stalls were present in genes in which translation initiation rates were very high, and only when stalls were present near the 5' end of the coding sequence, which presumably led to the queuing of ribosomes to an extent that translation initiation was also inhibited (Woolstenhulme et al. 2015). Stalling associated with overall decreased translation led to asymmetric ribosome footprinting profiles, in which ribosome

footprints were more abundant prior to the stall vs. after the stall site (Woolstenhulme et al. 2015).

We developed a "stalling score" metric (see Methods) to determine whether the presence of a stall site led to decreased translation, as evidenced by asymmetry in the abundances of ribosome footprints aligned to either the 5' or 3' end of transcripts' CDS regions. Large positive stalling scores represent a build-up of reads at the 5' end of the CDS region as compared to the 3' end, consistent with a significant stalling effect leading to reduced translation; a negative stalling score represents a build-up of reads at the 3' end. For example, Xbp1, which has an internal ribosomal stall site essential for its membrane targeting, has a stalling score of 16.56 +/- 3.89, reflecting the greater abundance of ribosome footprints prior to the stall as compared to after the stall (Fig. 3A). By contrast, the validated *Drosophila* Fmr1 target Poe/Ubr4 (Greenblatt and Spradling 2018) does not appear to have a ribosomal stall peak, and accordingly Poe/Ubr4 has a low stalling score -2.05 +/- 1.95 (Fig 3B).

In order to determine whether the changes in ribosome footprint abundance of Fmr1 target genes are due to altered ribosome stalling, we compared stalling scores of genes in either control or Fmr1-deficient oocytes and neurons. Inconsistent with the model that the reduced levels of ribosome footprints in Fmr1 targets is due to alleviated ribosome stalling, we found that genes that were translationally downregulated or upregulated in the *Drosophila* oocytes (Fig. 3C, E) and in the mouse cortex (Fig. 3D, F) had statistically identical distributions of stalling scores when analyzed using multiple statistical tests. We found a high correlation between stalling scores genome-wide when comparing wild type vs. Fmr1 RNAi *Drosophila* oocytes, as well as the mouse cortex from control vs. Fmr1 KO animals (Pearson's r = 0.93 and r = 0.89 respectively, Fig. 3 G,H). These data argue that changes in the distribution of ribosomal stall sites in Fmr1 RNAi or KO animals do not impact the overall translation of Fmr1 targets and non-targets.

Transcripts from ASD/ID genes are enriched in stress granules upon acute cell stress

A unique feature of the cytoplasm of oocytes and neurons is the abundance of RNA granules. RNA granules such as P bodies form as a response to the accumulation of untranslated, stable mRNA molecules (Eulalio et al. 2007). P bodies and stress granules form through largely non-specific, multivalent interactions between and among RNA binding proteins and RNA molecules themselves (Molliex et al. 2015; Rao and Parker 2017; Van Treeck et al. 2018). The non-specific nature of recruitment of mRNAs into RNA granules would suggest that these structures and their associated proteins would bias their regulatory effects more strongly towards longer mRNA molecules, which contain more potential sites for non-specific protein-RNA and RNA-RNA interactions. Indeed, multiple groups have found that stress-induced RNA granules, as well as developmentally induced P granules, are enriched for long mRNA transcripts (Khong et al. 2017; Padrón et al. 2019; Lee et al. 2020). Given the requirement for Fmr1 targets such as Poe/Ubr4 in the maintenance of cell homeostasis, these results are consistent with the model that Fmr1 evolved to maintain uniform translation by counteracting the inherent tendency of long mRNAs to partition into repressive RNPs under stressful conditions. We tested this model by analyzing the enrichment for either Fmr1 target transcripts (defined as transcripts significantly translationally reduced in the absence of Fmr1) or ASD/ID transcripts in the transcriptomes of stress-induced granules following arsenite exposure (a stress granule-enhancing treatment) of a human cell line (Khong et al. 2017). We found that transcripts encoded by orthologs of mouse Fmr1 targets were enriched in abundance in stress granules compared to all mRNA transcripts, with a median enrichment of 2.4-fold for Fmr1 targets as compared to a 0.98-fold enrichment for all transcripts (Fig. 4A). Similarly, we observed an enrichment of ASD/ID genes in stress granule transcriptomes, with a median enrichment of 2.0-fold for transcripts from all ASD/ID SFARI Class I and Class II genes, while

those ASD/ID genes downregulated in Fmr1 KO cortex had a median enrichment of 3.3-fold. These data demonstrate that Fmr1 target mRNAs, and transcripts from ASD/ID genes generally, are preferentially recruited to stress granules that form under acute stress.

Mutations in Fmr1 ASD/ID targets frequently lead to haploinsufficiency

Our analyses suggest that Fmr1 acts as an activator of translation at the level of ~2-fold or less (Greenblatt and Spradling 2018). A 2-fold activation in protein production is potentially significant for "haploinsufficient" genes that cause defects when present in only one instead of two copies. Among genes translationally downregulated in the Fmr1 KO mouse cortex, we identified 39 as SFARI Class I and Class II ASD genes (Table 1). Fmr1 target ASD genes encoded proteins having an average length of >2,000 amino acids, and these genes were translationally downregulated to between 58-83% of control levels based on reduced ribosome footprints. Many of Fmr1 target ASD genes are likely activated translationally by Fmr1 through a direct mechanism, as ~80% (31/39) encode transcripts that were found to bind to FMRP protein in CLIP experiments (Darnell et al. 2011). Of Fmr1 target ASD genes, 20/39 were found to be dosage sensitive at the levels of “emerging evidence” and “sufficient evidence” (score of 2 and 3 respectively) based on clinical data curated by the Clinical Genome Resource (Rehm et al. 2015) (Table 1). Thus, at least 20 dose-sensitive ASD genes are translationally downregulated in the Fmr1 cortex, each potentially contributing to defects associated with Fmr1 deficiency even if the reduction of any one of them is less than 2-fold.

Discussion

FMRP plays an important and ancient role in the development and function of neurons, oocytes, and spermatocytes, where it is required in diverse animals ranging from *Drosophila* to humans (Hagerman et al. 2017; Drozd et al. 2018). Here we have addressed recent

controversies regarding the action of FMRP on target mRNAs (Greenblatt and Spradling, 2018; Das Sharma et al. 2019; Sawicki et al. 2019). We argue that the highly conserved FMRP protein evolved to activate protein synthesis to compensate for translational reductions when stressful or developmentally-induced conditions increase RNP granule content (Fig. 5). We find no evidence that FMRP directly influences ribosome stalling or translational elongation, but likely acts by increasing initiation. The large average size of FMRP target proteins reflects the size dependence of stress-related translational repression, driven by biochemical equilibria. We document here that proteins associated with intellectual disability and autism are also even larger on average than other neural-expressed gene products.

Large cells such as neurons and oocytes rely heavily on the transport and storage of stable, untranslated mRNAs. For example a large fraction of oocyte mRNAs are translationally repressed for later use during early embryonic development (Greenblatt et al. 2019; Kronja et al. 2014). Neurons transport translationally regulated synaptic mRNAs up to a meter or more down long axonal projections. Such transport requires cells to partition mRNAs into repressed or actively translated states through the use of P body and neuronal granule RNPs, that are structurally related to stress granules. Thus, even when tissue conditions are fully optimal, there are periods when mRNAs undergoing transport from the cell body along an axon, or while stored in the vicinity of a synapse, must remain potentially functional while in an inactive state. This storage process likely depend on RNPs and may itself have an mRNA length bias. Experiments in *Drosophila* oocytes where the duration of storage can be controlled show that Fmr1 function is particularly required under these conditions (Greenblatt and Spradling 2018).

Our analyses of ribosome profiling data lead us to propose that the primary function of Fmr1 in both *Drosophila* and mice, and in both neural tissue and oocytes, is to boost rather than to repress translation. Fmr1 has long been considered a translational repressor (Darnell et al. 2011; Laggerbauer et al. 2001; Li et al. 2001); however bulk translation in the brain may or

not be decreased in the absence of Fmr1 (Qin et al. 2013; Qin et al. 2015). We find that large protein production is similarly reduced in the Fmr1-null mouse cortex (Fig. 2A,C), consistent with observations that Fmr1 CLIP targets have concordantly reduced ribosome footprints (Fig. 2E and Das Sharma et al. 2019), and that Fmr1 preferentially binds to long transcripts (Sawicka et al. 2019; Li et al. 2020). The increased levels of overall translation in Fmr1 mutant animals are likely a secondary, indirect effect of Fmr1 loss. Genes with increased translation in Fmr1-null animals are not direct targets of Fmr1 as defined by ribosome profiling and binding experiments (Das Sharma et al. 2019; Darnell et al. 2011), and the indirect increase in the translation of some genes is in part due to the dysregulation of mTOR signaling (Das Sharma et al. 2019; Sharma et al. 2010). These data explain the lack of efficacy of the mTOR inhibitor rapamycin in reversing behavior deficits of Fmr1 null mice, as treating the increased translation levels of secondary Fmr1 targets would not rescue the primary translational defect in these animals - which is the underproduction of dozens of ASD/ID-associated proteins.

These new insights raise several questions and suggest new approaches to understanding and treating FMRP-caused syndromes and autism. The first is why large proteins are involved in neural development and are required for particularly complex tasks such as integrating sensory inputs and outputs in ways that are vital to survival. It may be advantageous to place many individual protein domains along a single polypeptide chain in order to ensure they are expressed at equimolar levels. This may help guarantee the proper stoichiometries of protein domains that function together in a complex process. Indeed, consistent with this idea many large autism-associated proteins are highly sensitive to changes in gene dosage (Table 1). Another possibility is that large proteins have many binding sites for additional binding partners, and thereby function as scaffolds for exceptionally large protein complexes that may be important in neurons. These complexes often associate binding partners together through phase separation, which is sensitive to protein concentration (Molliex

et al. 2015). If processes associated with neuronal-based decision making are especially dependent on such processes, this might explain the enrichment of such proteins in FMRP targets and in ASD/ID gene products.

Finally, it is important consider how the framework of FMRP function outlined here impacts ideas for mitigating the effects of genetic alterations in Fmr1 and in ASD/ID genes. The association of many transcripts encoded by ASD/ID genes with FMRP suggests that pharmacological interventions targeting this system could selectively promote the expression of dozens of genes individually associated with ASD/ID, and therefore may represent one means in which to rescue the expression of genes leading to syndromes caused by haploinsufficiency. Further efforts should address the exact mechanism by which FMRP promotes the translation of its targets, including a better understanding of the length-dependent recruitment of mRNAs into repressive RNPs, testing whether FMRP recruits its targets to specific sites of translation, and identifying suppressors that rescue the translation of FMRP targets in mutant cells.

Methods

Read alignments and quantification

Raw sequencing data from (Greenblatt and Spradling 2018) and (Das Sharma et al. 2019) were aligned to the Flybase Consortium/Berkeley Drosophila Genome Project (BDGP)/Celera Genomics *Drosophila* release dm 6.28 and Genome Reference Consortium mouse build 38 mm10 assemblies respectively. Alignments were performed with STAR (Dobin et al. 2013) and TPM values computed using RSEM (Li and Dewey 2011) for mRNA sequencing experiments. TPM values for ribosome profiling data were computed using featureCounts from the Subread (Liao et al. 2014) package for ribosome footprints aligning to CDS regions.

Length and expression analyses

SFARI Class I and Class II genes were obtained from the SFARI Gene database (Abrahams et al. 2013). For each gene, the highest expressed transcript isoform was used as the representative form. Transcripts with a TPM value of <2 were excluded from analyses. Intron, UTR, and CDS lengths for transcript isoforms were computed from BDGP and refGene gene models for *Drosophila* and mouse data respectively. For translation efficiency analyses, fold-change and p_{adj} values were computed using RiboDiff software (Zhong et al. 2017). Stress-granule enrichment data were obtained from (Khong et al. 2017).

Ribosome stalling analysis

Ribosome footprints were mapped to their P sites with a 3' offset of 14 bases using the make wiggle function from the Plastid software package (Dunn and Weissman 2016). Ribosome footprinting read densities mapping to the 5' and 3' ends (first and last quintile) of transcript CDS regions were quantified with Plastid. The stalling scores for representative transcripts were computed as the (5' CDS read density – 3' CDS read density) / (sum CDS read density) x

100, where 5' CDS and 3' CDS read densities represent the sum of reads coming from the 5'-most and 3'-most 20% of the CDS region respectively. Significantly upregulated and downregulated genes were defined as genes whose TE increased or decreased in Fmr1 mutant or RNAi / control with a p_{adj} value of $< .01$. For genome-wide stalling analysis, standard deviations of stalling scores from either control of Fmr1-deficient replicates were computed. The quality of the stalling scores were determined by taking the minimum stalling score to standard deviation ratio for control and Fmr1 KO/RNAi experiments for each transcript. Transcripts with the top 4,000 stalling quality scores (i.e. transcript stalling scores with the least noise) were used for genome-wide analyses.

References

Abrahams, B.S., Arking, D.E., Campbell, D.B., et al. 2013. SFARI Gene 2.0: a community-driven knowledgebase for the autism spectrum disorders (ASDs). *Molecular autism* 4(1), p. 36.

Anderson, B.R., Chopra, P., Suhl, J.A., Warren, S.T. and Bassell, G.J. 2016. Identification of consensus binding sites clarifies FMRP binding determinants. *Nucleic Acids Research* 44(14), pp. 6649–6659.

Antar, L.N., Dictenberg, J.B., Plociniak, M., Afroz, R. and Bassell, G.J. 2005. Localization of FMRP-associated mRNA granules and requirement of microtubules for activity-dependent trafficking in hippocampal neurons. *Genes, Brain, and Behavior* 4(6), pp. 350–359.

Ascano, M., Mukherjee, N., Bandaru, P., et al. 2012. FMRP targets distinct mRNA sequence elements to regulate protein expression. *Nature* 492(7429), pp. 382–386.

Barbee, S.A., Estes, P.S., Cziko, A.-M., et al. 2006. Staufen- and FMRP-containing neuronal RNPs are structurally and functionally related to somatic P bodies. *Neuron* 52(6), pp. 997–1009.

Bear, M.F. 2005. Therapeutic implications of the mGluR theory of fragile X mental retardation. *Genes, Brain, and Behavior* 4(6), pp. 393–398.

Bear, M.F., Huber, K.M. and Warren, S.T. 2004. The mGluR theory of fragile X mental retardation. *Trends in Neurosciences* 27(7), pp. 370–377.

Bülow, P., Murphy, T.J., Bassell, G.J. and Wenner, P. 2019. Homeostatic intrinsic plasticity is functionally altered in fmr1 KO cortical neurons. *Cell reports* 26(6), p. 1378–1388.e3.

Chen, L., Yun, S.W., Seto, J., Liu, W. and Toth, M. 2003. The fragile X mental retardation protein binds and regulates a novel class of mRNAs containing U rich target sequences. *Neuroscience* 120(4), pp. 1005–1017.

Comery, T.A., Harris, J.B., Willems, P.J., et al. 1997. Abnormal dendritic spines in fragile X knockout mice: maturation and pruning deficits. *Proceedings of the National Academy of Sciences of the United States of America* 94(10), pp. 5401–5404.

Darnell, J.C., Jensen, K.B., Jin, P., Brown, V., Warren, S.T. and Darnell, R.B. 2001. Fragile X mental retardation protein targets G quartet mRNAs important for neuronal function. *Cell* 107(4), pp. 489–499.

Darnell, J.C., Van Driesche, S.J., Zhang, C., et al. 2011. FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell* 146(2), pp. 247–261.

Das Sharma, S., Metz, J.B., Li, H., et al. 2019. Widespread alterations in translation elongation in the brain of juvenile fmr1 knockout mice. *Cell reports* 26(12), p. 3313–3322.e5.

Dictenberg, J.B., Swanger, S.A., Antar, L.N., Singer, R.H. and Bassell, G.J. 2008. A direct role

for FMRP in activity-dependent dendritic mRNA transport links filopodial-spine morphogenesis to fragile X syndrome. *Developmental Cell* 14(6), pp. 926–939.

Dobin, A., Davis, C.A., Schlesinger, F., et al. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29(1), pp. 15–21.

Doerfel, L.K., Wohlgemuth, I., Kothe, C., Peske, F., Urlaub, H. and Rodnina, M.V. 2013. EF-P is essential for rapid synthesis of proteins containing consecutive proline residues. *Science* 339(6115), pp. 85–88.

Dolzhanskaya, N., Sung, Y.J., Conti, J., Currie, J.R. and Denman, R.B. 2003. The fragile X mental retardation protein interacts with U-rich RNAs in a yeast three-hybrid system. *Biochemical and Biophysical Research Communications* 305(2), pp. 434–441.

Drozd, M., Bardoni, B. and Capovilla, M. 2018. Modeling fragile X syndrome in drosophila. *Frontiers in Molecular Neuroscience* 11, p. 124.

Dunn, J.G. and Weissman, J.S. 2016. Plastid: nucleotide-resolution analysis of next-generation sequencing and genomics data. *BMC Genomics* 17(1), p. 958.

Eulalio, A., Behm-Ansmant, I., Schweizer, D. and Izaurralde, E. 2007. P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. *Molecular and Cellular Biology* 27(11), pp. 3970–3981.

Feng, Y., Absher, D., Eberhart, D.E., Brown, V., Malter, H.E. and Warren, S.T. 1997. FMRP associates with polyribosomes as an mRNP, and the I304N mutation of severe fragile X syndrome abolishes this association. *Molecular Cell* 1(1), pp. 109–118.

Greenblatt, E.J., Obniski, R., Mical, C. and Spradling, A.C. 2019. Prolonged ovarian storage of mature Drosophila oocytes dramatically increases meiotic spindle instability. *eLife* 8.

Greenblatt, E.J. and Spradling, A.C. 2018. Fragile X mental retardation 1 gene enhances the translation of large autism-related proteins. *Science* 361(6403), pp. 709–712.

Hagerman, R.J., Berry-Kravis, E., Hazlett, H.C., et al. 2017. Fragile X syndrome. *Nature reviews. Disease primers* 3, p. 17065.

Hou, L., Antion, M.D., Hu, D., Spencer, C.M., Paylor, R. and Klann, E. 2006. Dynamic translational and proteasomal regulation of fragile X mental retardation protein controls mGluR-dependent long-term depression. *Neuron* 51(4), pp. 441–454.

Khong, A., Matheny, T., Jain, S., Mitchell, S.F., Wheeler, J.R. and Parker, R. 2017. The Stress Granule Transcriptome Reveals Principles of mRNA Accumulation in Stress Granules. *Molecular Cell* 68(4), p. 808–820.e5.

King, I.F., Yandava, C.N., Mabb, A.M., et al. 2013. Topoisomerases facilitate transcription of

long genes linked to autism. *Nature* 501(7465), pp. 58–62.

Kronja, I., Yuan, B., Eichhorn, S.W., et al. 2014. Widespread changes in the posttranscriptional landscape at the Drosophila oocyte-to-embryo transition. *Cell reports* 7(5), pp. 1495–1508.

Laggerbauer, B., Ostareck, D., Keidel, E.M., Ostareck-Lederer, A. and Fischer, U. 2001. Evidence that fragile X mental retardation protein is a negative regulator of translation. *Human Molecular Genetics* 10(4), pp. 329–338.

Lee, C.-Y.S., Putnam, A., Lu, T., He, S., Ouyang, J.P.T. and Seydoux, G. 2020. Recruitment of mRNAs to P granules by condensation with intrinsically-disordered proteins. *eLife* 9.

Li, B. and Dewey, C.N. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12, p. 323.

Li, M., Shin, J., Risgaard, R.D., et al. 2020. Identification of FMR1-regulated molecular networks in human neurodevelopment. *Genome Research* 30(3), pp. 361–374.

Li, Z., Zhang, Y., Ku, L., Wilkinson, K.D., Warren, S.T. and Feng, Y. 2001. The fragile X mental retardation protein inhibits translation via interacting with mRNA. *Nucleic Acids Research* 29(11), pp. 2276–2283.

Liao, Y., Smyth, G.K. and Shi, W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30(7), pp. 923–930.

Molliex, A., Temirov, J., Lee, J., et al. 2015. Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* 163(1), pp. 123–133.

Padrón, A., Iwasaki, S. and Ingolia, N.T. 2019. Proximity RNA Labeling by APEX-Seq Reveals the Organization of Translation Initiation Complexes and Repressive RNA Granules. *Molecular Cell* 75(4), p. 875–887.e5.

Park, S., Park, J.M., Kim, S., et al. 2008. Elongation factor 2 and fragile X mental retardation protein control the dynamic translation of Arc/Arg3.1 essential for mGluR-LTD. *Neuron* 59(1), pp. 70–83.

Pereanu, W., Larsen, E.C., Das, I., et al. 2017. AutDB: a platform to decode the genetic architecture of autism. *Nucleic Acids Research* 46.

Phan, A.T., Kuryavyi, V., Darnell, J.C., et al. 2011. Structure-function studies of FMRP RGG peptide recognition of an RNA duplex-quadruplex junction. *Nature Structural & Molecular Biology* 18(7), pp. 796–804.

Qin, M., Huang, T., Kader, M., et al. 2015. R-Baclofen Reverses a Social Behavior Deficit and Elevated Protein Synthesis in a Mouse Model of Fragile X Syndrome. *The International Journal*

of Neuropsychopharmacology 18(9).

Qin, M., Schmidt, K.C., Zametkin, A.J., et al. 2013. Altered cerebral protein synthesis in fragile X syndrome: studies in human subjects and knockout mice. *Journal of Cerebral Blood Flow and Metabolism* 33(4), pp. 499–507.

Rao, B.S. and Parker, R. 2017. Numerous interactions act redundantly to assemble a tunable size of P bodies inSaccharomyces cerevisiae. *Proceedings of the National Academy of Sciences of the United States of America* 114(45), pp. E9569–E9578.

Ray, D., Kazan, H., Cook, K.B., et al. 2013. A compendium of RNA-binding motifs for decoding gene regulation. *Nature* 499(7457), pp. 172–177.

Rehm, H.L., Berg, J.S., Brooks, L.D., et al. 2015. ClinGen--the Clinical Genome Resource. *The New England Journal of Medicine* 372(23), pp. 2235–2242.

Ronesi, J.A., Collins, K.A., Hays, S.A., et al. 2012. Disrupted Homer scaffolds mediate abnormal mGluR5 function in a mouse model of fragile X syndrome. *Nature Neuroscience* 15(3), pp. 431–40, S1.

Rosario, R., Filis, P., Tessyman, V., et al. 2016. FMRP Associates with Cytoplasmic Granules at the Onset of Meiosis in the Human Oocyte. *Plos One* 11(10), p. e0163987.

Rudelli, R.D., Brown, W.T., Wisniewski, K., et al. 1985. Adult fragile X syndrome. Clinico-neuropathologic findings. *Acta Neuropathologica* 67(3–4), pp. 289–295.

Sawicka, K., Hale, C.R., Park, C.Y., et al. 2019. FMRP has a cell-type-specific role in CA1 pyramidal neurons to regulate autism-related transcripts and circadian memory. *eLife* 8.

Sears, J.C., Choi, W.J. and Broadie, K. 2019. Fragile X mental retardation protein positively regulates PKA anchor Rugose and PKA activity to control actin assembly in learning/memory circuitry. *Neurobiology of Disease* 127, pp. 53–64.

Sharma, A., Hoeffer, C.A., Takayasu, Y., et al. 2010. Dysregulation of mTOR signaling in fragile X syndrome. *The Journal of Neuroscience* 30(2), pp. 694–702.

Sullivan, S.D., Welt, C. and Sherman, S. 2011. FMR1 and the continuum of primary ovarian insufficiency. *Seminars in Reproductive Medicine* 29(4), pp. 299–307.

Ude, S., Lassak, J., Starosta, A.L., Kraxenberger, T., Wilson, D.N. and Jung, K. 2013. Translation elongation factor EF-P alleviates ribosome stalling at polyproline stretches. *Science* 339(6115), pp. 82–85.

Van Treeck, B., Protter, D.S.W., Matheny, T., Khong, A., Link, C.D. and Parker, R. 2018. RNA self-assembly contributes to stress granule formation and defining the stress granule transcriptome. *Proceedings of the National Academy of Sciences of the United States of America*

115(11), pp. 2734–2739.

Vasilyev, N., Polonskaia, A., Darnell, J.C., Darnell, R.B., Patel, D.J. and Serganov, A. 2015. Crystal structure reveals specific recognition of a G-quadruplex RNA by a β-turn in the RGG motif of FMRP. *Proceedings of the National Academy of Sciences of the United States of America* 112(39), pp. E5391-400.

Woolstenhulme, C.J., Guydosh, N.R., Green, R. and Buskirk, A.R. 2015. High-precision analysis of translational pausing by ribosome profiling in bacteria lacking EFP. *Cell reports* 11(1), pp. 13–21.

Zalfa, F., Adinolfi, S., Napoli, I., et al. 2005. Fragile X mental retardation protein (FMRP) binds specifically to the brain cytoplasmic RNAs BC1/BC200 via a novel RNA-binding motif. *The Journal of Biological Chemistry* 280(39), pp. 33403–33410.

Zhang, Z., Marro, S.G., Zhang, Y., et al. 2018. The fragile X mutation impairs homeostatic plasticity in human neurons by blocking synaptic retinoic acid signaling. *Science Translational Medicine* 10(452).

Zhao, Y.-T., Kwon, D.Y., Johnson, B.S., et al. 2018. Long genes linked to autism spectrum disorders harbor broad enhancer-like chromatin domains. *Genome Research* 28(7), pp. 933–942.

Zhong, Y., Karaletsos, T., Drewe, P., et al. 2017. RiboDiff: detecting changes of mRNA translation efficiency from ribosome footprints. *Bioinformatics* 33(1), pp. 139–141.

Figure Legends

Figure 1 SFARI autism genes have longer open reading frames, UTRs and introns than average. Box plots of length distributions showing increased (A) coding sequences lengths, (B) 5' UTR lengths, (C) intron lengths (summed across all introns), and (D) 3' UTR lengths specifically for SFARI class I and class II autism spectrum disorder (ASD) genes (magenta) as compared to all genes expressed in the juvenile mouse cortex (blue). (E, F) No significant differences were detected between the distributions of transcript levels (E) or translation levels (F) specifically for ASD genes vs. all cortex-expressed genes as detected by mRNA sequencing and ribosome footprinting respectively.

Figure 2 Reduced translation of large proteins in Fmr1 RNAi *Drosophila* oocytes and Fmr1 KO mouse cortex. (A, C) Volcano plots showing concordantly diminished translation efficiency of large proteins (colored dark purple) in *Drosophila* oocytes (Greenblatt and Spradling 2018) (A) and in mouse cortex (Das Sharma et al. 2019) (C). (B, D) Box plots showing increased CDS lengths of genes translationally reduced in Fmr1 RNAi *Drosophila* oocytes (B) or in Fmr1 KO mouse cortex (D) as compared to all genes, or to genes that were increased in translation efficiency. (E,F) Volcano plots showing concordant reductions in the translation efficiency of FMRP clip targets (Darnell et al. 2011) (E) and SFARI Class I and Class II ASD genes (F) in the Fmr1 KO mouse cortex (Das Sharma et al. 2019). (G, H) Scatter plots showing reduced translation efficiency as a function of protein size for the 200 most statistically affected genes in Fmr1 RNAi *Drosophila* oocytes (G), or Fmr1 KO mouse cortex (H). Histograms showing size distributions for all *Drosophila* or mouse proteins respectively are shown to the right. (I) Heat map showing concordantly reduced translation levels but not mRNA levels for genes with reduced translation efficiency encoding large proteins >1,000 amino acids. (A) differs slightly from Greenblatt and Spradling, 2018, which plotted translation rather than

translation efficiency.

Figure 3. No apparent role for ribosome stalling in Fmr1-dependent translation in *Drosophila* oocytes or the mouse cortex. (A,B) Stalling scores (see Methods) were computed from the *Drosophila* oocyte ribosome profiles of the non-Fmr1 target Xbp1 (A) and the Fmr1 target Poe/Ubr4 (B). (C,D) Box plots showing no statistically significant changes in the distribution of stalling scores for genes that were translationally upregulated or downregulated in Fmr1 RNAi *Drosophila* oocytes (C) or the Fmr1 KO mouse cortex (D) (Student's t test, two tailed). (E, F). Cumulative distribution plots showing no statistically significant changes in the stalling scores for genes that were translationally upregulated or downregulated in Fmr1 RNAi *Drosophila* oocytes (E) or the Fmr1 KO mouse cortex (F) (Kolmogorov-Smirnov test). (G, H) Scatter plots showing high Pearson's correlation of stalling scores from genes expressed in controls vs. Fmr1 RNAi oocytes (G) or the Fmr1 KO mouse cortex (H).

Figure 4. Transcripts encoded by ASD/ID genes and Fmr1 targets become enriched in stress granules following acute stress. (A) Box plots showing human orthologs (Khong et al. 2017), specifically of genes translationally reduced in the Fmr1 KO cortex, become enriched in stress granules following arsenite treatment. (B) Box plots showing enrichment into stress granules of SFARI Class I and Class II ASD genes and ASD genes translationally reduced in the Fmr1 mouse cortex following acute arsenite treatment.

Figure 5. Model of translational activation through Fmr1-dependent mRNA partitioning. Long mRNAs (orange) preferentially condense into translationally repressive P bodies as compared to short mRNAs which remain outside of P bodies and are highly translated (blue).

Fmr1 (magenta circles) binds to specific long mRNAs and recruits them into an alternative RNP where their translation is maintained at a higher level.

Table 1. SFARI Class I and Class II ASD genes are translationally reduced in the Fmr1 KO mouse cortex. Table showing SFARI Class I and Class II ASD genes that are significantly altered in their translation efficiency in the Fmr1 KO mouse brain. 98% (39/40) of altered ASD genes are translationally reduced. ClinGen Dosage scores (Rehm et al. 2015) show that many of these ASD genes have strong (3), emerging (2), or preliminary evidence (1) of human syndromes associated with changes in gene dosage.

SUPPLEMENTAL TABLE LEGENDS

Supplementary Table 1. Compiled RiboDiff output for Fmr1 RNAi *Drosophila* oocyte and Fmr1 KO mouse cortex analyses.

Supplementary Table 2. ClinGen haploinsufficiency scores for *Drosophila* orthologs of SFARI Class I and Class II autism genes translationally downregulated in oocytes.

Figure 1

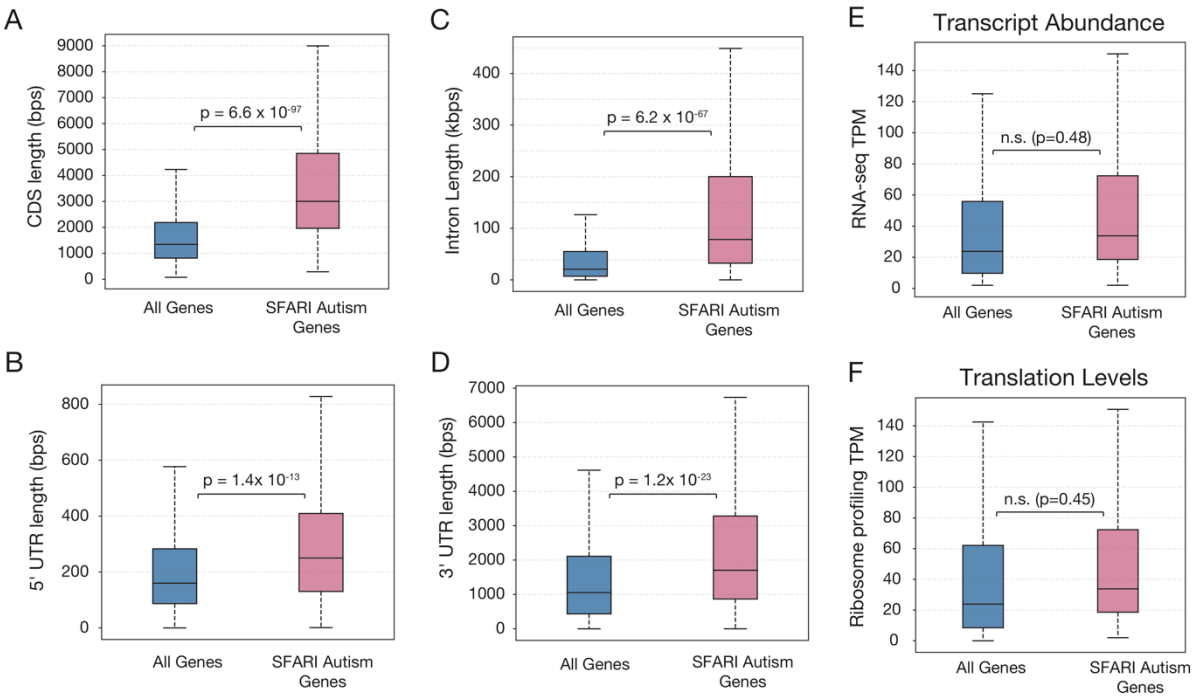

Figure 2

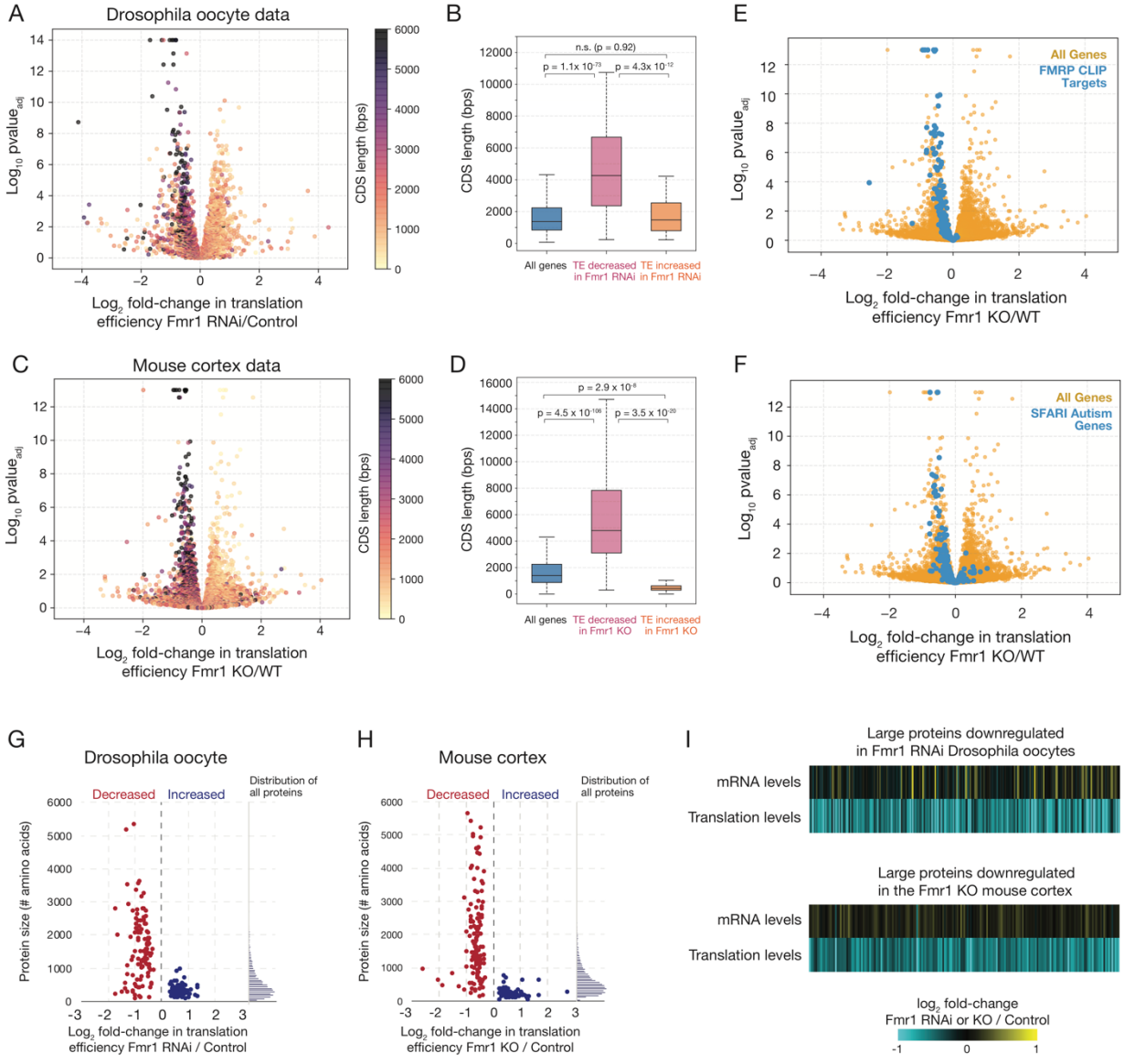

Figure 3

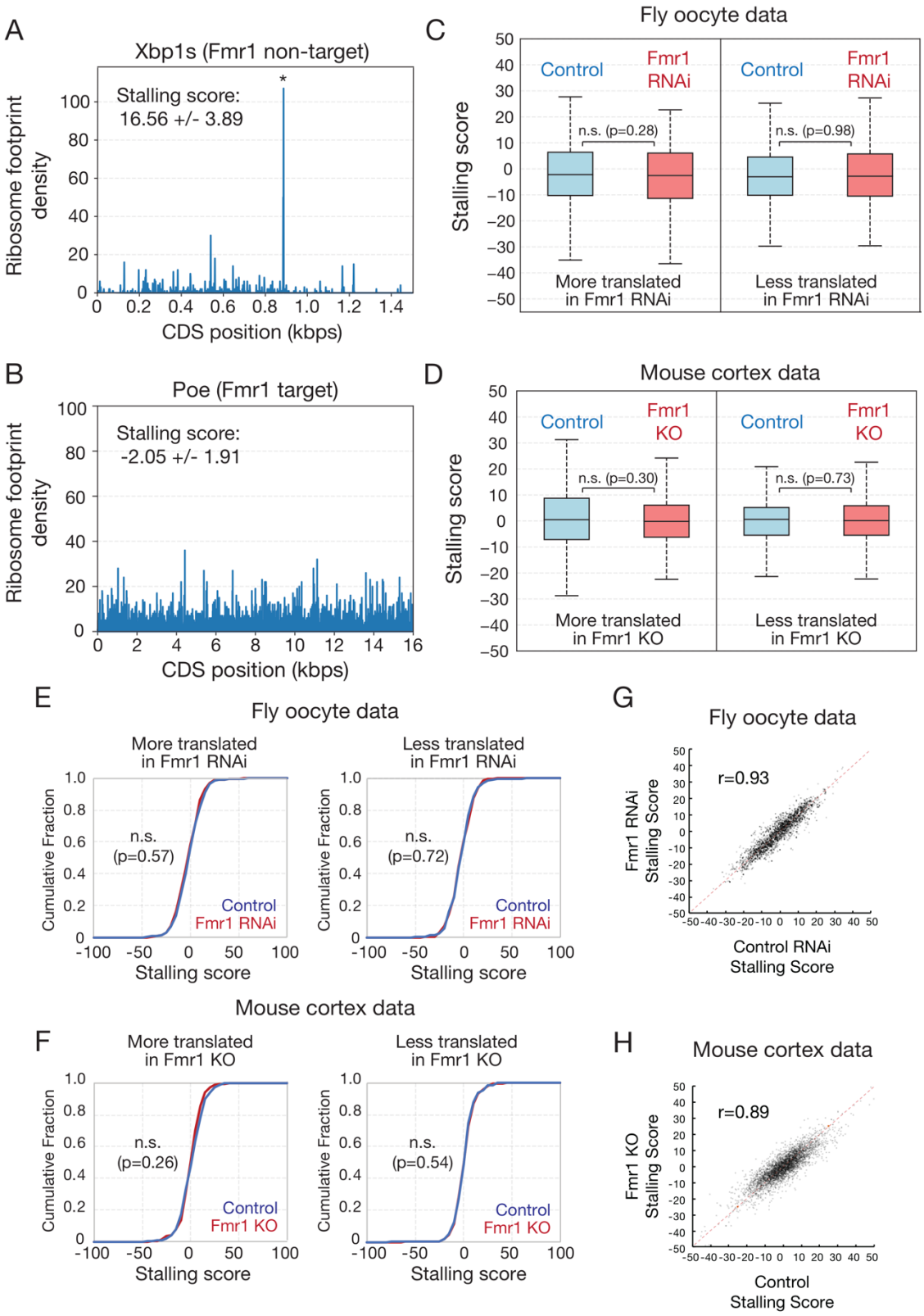

Figure 4

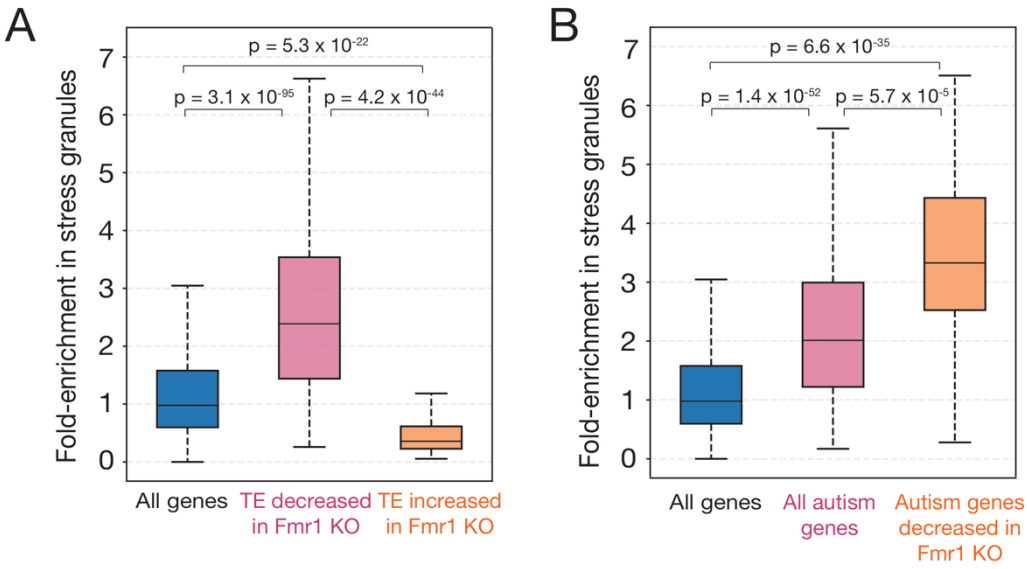

Figure 5

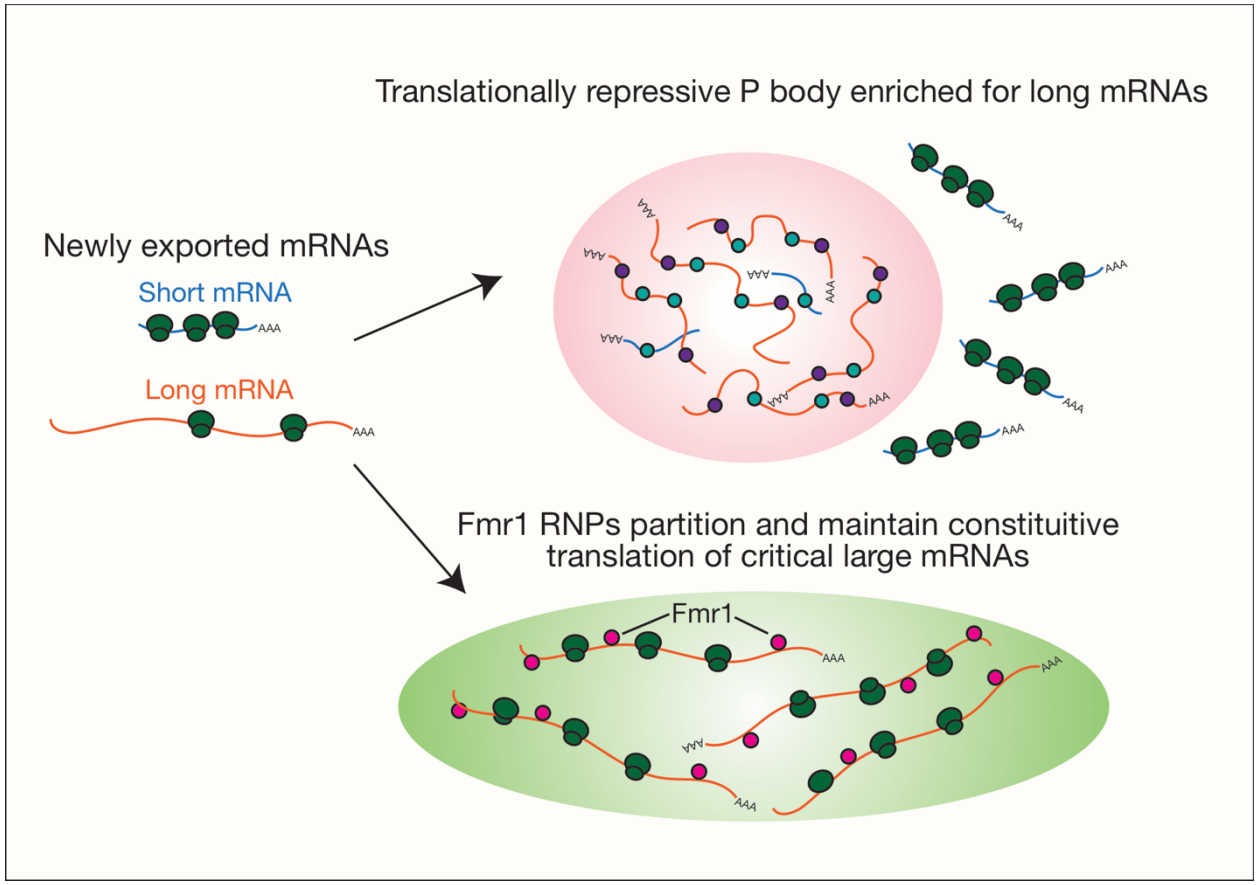

Table 1

SFARI Autism Gene	$Fmr1^{KO}$ / WT Translation Efficiency	$p\text{-value}_{adj}$	# Amino acids	ClinGen Dosage Score	Darnell et al. 2011 CLIP Target
KMT2A	0.58	9.0E-05	3966	3	✓
HIVEP3	0.58	4.7E-04	2348	1	✓
SHANK1	0.59	0.0E+00	2167	1	✓
KMT2C	0.61	4.0E-08	4904	3	✓
SPEN	0.63	2.7E-07	3620	1	✓
MED13L	0.63	4.8E-08	2207	3	✓
TCF20	0.64	7.1E-07	1987	3	✓
AHDC1	0.65	2.2E-07	1594	3	✓
SRCAP	0.65	6.2E-07	3271	1	
CIC	0.66	5.4E-07	1604	2	✓
KDM3B	0.66	1.3E-06	1762	n/a	
CNOT3	0.66	5.0E-03	751	0	
SHANK2	0.67	5.9E-08	1472	2	✓
RIMS1	0.67	9.7E-07	1190	1	
AUTS2	0.67	8.0E-06	1261	3	✓
RERE	0.68	1.0E-03	1558	n/a	✓
HECTD4	0.68	0.0E+00	4418	n/a	
LRP1	0.69	0.0E+00	4545	n/a	✓
CREBBP	0.69	1.1E-06	2441	3	✓
CHD8	0.70	2.7E-05	2582	3	✓
SHANK3	0.71	5.4E-03	1805	3	✓
ASH1L	0.71	2.9E-09	2958	3	✓
IQSEC2	0.71	1.3E-05	1479	3	✓
SON	0.72	5.2E-04	2444	3	✓
GRIK5	0.72	5.6E-03	979	n/a	✓
AKAP9	0.74	1.2E-03	3779	n/a	✓
NCOR1	0.74	4.1E-07	2454	n/a	✓
ARID1B	0.75	1.8E-04	2244	3	✓
MYT1L	0.77	6.6E-03	1185	3	✓
FOXG1	0.78	8.6E-03	481	3	
MECP2	0.78	9.5E-03	484	3	
NSD1	0.78	4.3E-03	2691	3	✓
CHD3	0.79	4.8E-04	2055	n/a	✓
EP400	0.80	8.2E-04	2999	n/a	✓
PHF3	0.80	2.7E-03	2025	1	
SLC12A5	0.81	5.0E-03	1138	n/a	✓
CTNND2	0.81	2.3E-04	1221	2	✓
TANC2	0.83	7.6E-03	1994	n/a	✓
ATP2B2	0.83	5.5E-04	1198	n/a	✓
ILF2	1.25	9.5E-03	390	n/a	