Ribosomes in RNA granules are stalled on mRNA sequences that are consensus sites for FMRP association.

Mina N. Anadolu[1], Senthilkumar Kailasam[2], Konstanze Simbriger[3], Jingyu Sun[4,5], Teodora Markova[1], Seyed Mehdi Jafarnejad[6], Francois Lefebvre[2], Joaquin Ortega[4,5], Christos G. Gkogkas[7], Wayne S. Sossin[1,4,*]

[1]Department of Neurology and Neurosurgery, Montreal Neurological Institute, McGill University, Montreal, Quebec, H3A 2B4, Canada.

[2]Canadian Centre for Computational Genomics, McGill University, Montreal, Quebec, H3A 0G1 Canada.

[3]Department of Pharmacology, Medical University Innsbruck, Innsbruck Austria Division of Biomedical Research, A-6020 Innsbruck, Austria

[4]Department of Anatomy and Cell Biology, McGill University, Montreal, Quebec H3A 0C7, Canada,

[5]Centre for Structural Biology, McGill University, Montreal, Quebec, H3G 0B1, Canada.

[6]Patrick G. Johnston Centre for Cancer Research, The Queen's University of Belfast, BT9 7AE, Belfast, Northern Ireland, UK

[7] Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology-Hellas, University Campus, 45110 Ioannina, Greece

*Corresponding Author. wayne.sossin@mcgill.ca (WSS)

Author contributions: Performed Research (MNA, KS, JS, SMJ), Analyzed Data (MNA, SK, TM, WSS), Designed research (FL, JO, CGG, WSS), Wrote paper (MNA, WS, JO, CGG)

Competing Interest Statement. There are no competing interests

Keywords: Ribosome Footprinting, FMRP, Stalled polysome, Neuronal RNA granule. Translational control

This PDF includes: Main Text. 6 figures, 2 tables, 5 supplemental figures.

Abstract

Local translation in neurons is mediated in part by the reactivation of stalled polysomes. However, the mechanism for stalling of the polysomes is not understood. Stalled polysomes may be enriched within neuronal RNA granules defined by dense collections of compacted ribosomes found in the pellet of sucrose gradients used to separate polysomes from monosomes. We find that this fraction, isolated from P5 rat brains of both sexes, is enriched in proteins implicated in stalled polysome function, such as the fragile X mental retardation protein (FMRP) and Up-frameshift mutation 1 homolog (UPF1). Ribosome profiling of this fraction showed an abundance of footprint reads derived from mRNAs of cytoskeletal proteins implicated in neuronal development and an enrichment of footprint reads on RNA binding proteins. Compared to those usually found in ribosome profiling studies, the footprint reads were more extended on their 3'end and were found in reproducible peaks in the mRNAs. These footprint reads were enriched in motifs previously associated with mRNAs cross-linked to FMRP in vivo, independently linking the ribosomes in the sedimented pellet to the ribosomes associated with FMRP in the cell. The data supports a model in which specific sequences in mRNAs act to stall translation elongation in neurons, attracting FMRP and beginning a process where stalled ribosomes are packaged and transported in RNA granules.

Significance Statement: This work finds that neuronal ribosomes in RNA granules are concentrated at consensus sites previously identified through cross-linking FMRP to mRNAs in the brain. This strongly links the compacted ribosomes found in the pellet of sucrose gradients from brain extracts to stalled ribosomes regulated by FMRP and provides important insights into how stalling is accomplished. Many mRNAs important for neurodevelopment are enriched in these ribosomes. These results suggest that many studies on translation in the brain may need to be revised. The larger size of the ribosomal footprints on stalled polysomes and their sedimentation in the pellet of sucrose gradients suggests mRNAs found in these structures have not been assessed in many studies of neuronal translation.

Introduction

In neurons, the local translation of mRNAs at distal synaptic sites is essential for maintaining the local proteome (1), developmental processes (2), homeostasis of excitability (3), and synaptic plasticity (4). Local translation requires the transport of mRNAs from the soma to distal sites in a repressed state, followed by their reactivation, either when the mRNA reaches its correct location or after an appropriate stimulus (5). Two major forms of mRNA transport have been defined in neurons: the transport of mRNAs that are repressed prior to the start of translational elongation, and the transport of mRNAs for which elongation has begun but then stalled (6). The mRNAs repressed before elongation lack large ribosomal subunits and are often transported in dedicated mRNA transport particles. In contrast, the mRNAs repressed at elongation are transported as stalled polysomes (6, 7). The local reactivation of translation from mRNAs transported in stalled polysomes can be distinguished from mRNAs blocked at initiation using drugs such as homoharringtonine (HHT) that specifically block the first step of translation elongation, thus blocking translation from mRNA transport particles but not RNA granules (5). Using this tool, several physiological processes have been shown to be supported by initiation-independent protein synthesis, such as the local production of microtubule associated protein 1B (Map1B) and the induction of a form of long-term depression stimulated by the activation of metabotropic glutamate receptors (mGluR-LTD) in vertebrates, as well as the induction of a type of intermediate term synaptic plasticity in the invertebrate model system *Aplysia californica* (8-10).

Stalled polysomes may be transported in large liquid-liquid phase separated structures termed neuronal RNA granules (5). These were first described in oligodendrocytes where they

transported myelin basic protein mRNA to myelin synthesis sites (11). The term neuronal RNA granule was first used to describe a sedimented fraction containing ribosomes and repressed mRNAs lacking initiation factors (7). These large collections of ribosomes can be separated from normal polysomes based on their high density (7, 12-15). The proteomic characterization of these structures is consistent with the possibility that they are stalled polysomes (13-15). Indeed, they are enriched in mRNAs such as Map1B that undergo initiation-independent protein synthesis (13).

Stalled polysomes may also be necessary for neuronal development and the regulation of developmentally expressed mRNAs through association with the fragile X mental retardation protein (FMRP), a protein that when lost, results in the neurodevelopmental disorder Fragile X syndrome (16). UV cross-linking of FMRP to mRNA in neurons showed that FMRP was mainly associated with the coding region of mRNAs (17), consistent with the possibility of FMRP associating with ribosomes. Indeed, FMRP has been shown to be associated with stalled ribosomes in several studies (12, 13, 17-19). A number of mRNA sequences enriched in the regions of mRNAs cross-linked with FMRP have been identified (20, 21) suggesting that specific mRNA sequences may be important for determining which mRNAs are recruited to FMRP-containing stalled polysomes. FMRP was also shown to be enriched in RNA granules isolated by sucrose granule sedimentation (13), consistent with the idea that this dense RNA granule fraction contains stalled polysomes.

The mechanism for stalling polysomes during elongation in neurons is not known. Recently, the nonsense-mediated decay factor UPF1 was implicated in this process. Decreasing levels of UPF1 reduced the levels of initiation-independent translation, the local production of

Map1B and disrupted the induction of mGluR-LTD (8). Since UPF1 is known to be attracted to ribosomes when they reach the stop codon through association with eukaryotic release factors (22), this suggests that stalled mRNAs within RNA granules are blocked at the release step of translation termination. To test this, we used ribosome profiling, also known as ribosome footprinting, to elucidate the position of ribosomes on mRNAs through sequencing the ribosome protected fragments (footprint reads) after nuclease treatment (23). We took advantage of the presumed enrichment of stalled polysomes in heavy fractions of sucrose gradients to create an enriched fraction and identify the sites occupied by ribosomes in these structures. The footprint reads identified by ribosome profiling of the sedimented RNA granules were larger than expected and produced reproducible peaks, which were highly enriched in motifs previously associated with FMRP target mRNAs (20, 21) and consensus sites for m6A modifications in the brain (24). Contrary to our prediction, the footprint reads were not enriched at the stop codon, but instead were slightly biased towards the first half of the open reading frame of transcripts. We propose a stochastic model in which stalling of ribosomes at specific motifs in mRNAs in neurons attracts FMRP, starting a process of ribosome compaction and RNA granule assembly that allows transport of stalled ribosomes to distal sites in neurons for local translation.

Results

Isolation of the pellet enriched in compacted ribosomes

To enrich for RNA granules and presumed stalled ribosomes, we adapted a previous protocol (13) for sucrose gradient sedimentation that involves a high-velocity 2-hour spin over a 60% sucrose pad to enrich for all polysomes, followed by an additional 45-minute short spin on a 15-60% sucrose gradient to further sediment heavier particles (Fig. 1A). From this second centrifugation, we collected fractions from the top and separated the proteins on an SDS gel followed by staining with Coomassie blue to visualize the contents of each fraction (Fig. 1B). We performed immunoblotting with an antibody against the ribosomal protein S6. We observed peaks of S6 both in fractions 5-6 and in the pellet (hereafter called granule fraction) consistent with the separation of polysomes from the denser compacted ribosomes of the RNA granule (Fig. 1B). Indeed, the granule fraction contained compacted ribosomes as observed under negative staining electron microscopy (Fig. 1C), consistent with previous descriptions of stalled polysomes (17) and RNA granules (7, 13, 14).

To further validate our enriched preparation, we performed western blot analysis on the starting material, fraction 5-6 (polysomes) and granule fraction, examining RNA binding proteins (RBPs) known to be markers for stalled polysomes and RNA granules. For quantification, we standardized the levels of the RBPs to the levels of the ribosomal protein S6 to normalize for the number of ribosomes in each fraction. FMRP has been linked to stalled polysomes in several studies (9, 13, 17, 18, 25) and was highly enriched in the granule fraction (Fig. 1D, E). Similarly Pur-alpha (PURA), a protein that marks RNA granules in neurons (13, 15) whose loss leads to

neurodevelopmental disorders (26), was significantly enriched in the granule fraction compared to polysomes (Fig. 1D, F). UPF1, the key component of the nonsense-mediated decay pathway (22), has been shown to play an independent role in the formation of stalled polysomes in neurons (8) and is enriched in the granule fraction, compared to polysomes (Fig. 1D, G). UPF1 interacts with Staufen 2 (STAU2), and this interaction is important for the formation of stalled polysomes (8). However, while STAU2 was present in the granule fraction, it was more enriched in polysomes (Fig. 1D, H). This may be due to an important role of STAU2 in regulating other forms of translation (27). Interestingly, we were unable to detect the presence of the eukaryotic elongation factor 2 (eEF2) in the polysome or the granule fraction (Fig. 1D, I). Given the proposed role for eEF2 phosphorylation in the reactivation of stalled polysomes (4, 10), the absence of eEF2 suggests that this may be due to a requirement for eEF2 release from normal translating polysome to restart stalled polysomes. Alternatively, the lack of cycloheximide in our purification may have allowed eEF2 to diffuse away, since cycloheximide blocks the elongation step of translation potentially trapping eEF2 on polysomes. Overall, the protein composition of the granule fraction is consistent with an enrichment of RNA granules containing stalled ribosomes.

Cleavage of compacted ribosomes into monosomes

Ribosome profiling is based on nuclease digestion of mRNA that is not protected by ribosomes, isolation of monosomes containing the protected mRNA, followed by library construction and sequencing. Compacted ribosomes are resistant to cleavage by nucleases (17). Based on the previous observation that high-salt conditions cause the unpacking of the compacted ribosomes (13), we treated the granule fraction with 400 mM of sodium chloride for

10 minutes before dilution to reduce the salt concentration back to physiological levels (150 mM) followed by incubation with RNase I. After RNase digestion we resedimented the fractions on a sucrose gradient to evaluate the digestion. We found that this treatment was able to cleave the compacted ribosome clusters to monosomes, measured based on the movement of S6 ribosomal protein immunoreactivity from the granule fraction to fraction 2 of the sucrose gradient (Fig. 2A-D). We observed that the high-salt treatment improved the digestion by nuclease at low nuclease concentrations, although some lots of nuclease led to complete cleavage even without the need for salt (Supp Fig. 1). At the higher concentration of nuclease (1 µl: 100 U), there was no difference in digestion in the presence or absence of salt (Fig 2A-D; Supp Fig. 1).

The effect of the high-salt and nuclease treatments was also visualized using negative staining electron microscopy (EM) (Fig 2E, F). Samples either untreated or treated with the high-salt buffer, nuclease (low concentration; 0.5 µl: 50 U) or both were deposited on the EM grids and negatively stained. Consistent with the sedimentation results, we observed that the high-salt and nuclease treatments, when applied separately, induced partial unpacking of the ribosome clusters. Only when both treatments were applied to the sample, the EM images showed that the ribosome clusters had dissociated into monosomes.

Ribosome Footprints of the mRNAs protected by ribosomes in the pellet fraction

After treatment with RNase I, the cleaved ribosomes from the granule fraction were loaded onto a second 15-60% sucrose gradient and centrifuged to separate the monosomes before RNA extraction, library preparation and sequencing of the footprint reads (Fig. 3A). As

we are interested in stalled polysomes, no attempt was made to prevent ribosome run-off, and cycloheximide was not present during the tissue homogenization. We tested whether the presence or absence of cycloheximide has an impact on ribosome footprints (Supp. Fig 2). There was little effect of the omission of cycloheximide on the mRNAs detected by footprint reads as seen by a high Pearson's correlation and clustering by principal components analysis (PCA) between relative numbers of footprint reads in the presence or absence of cycloheximide (Supp. Fig 2). Similarly, nuclease digestion was performed after a brief treatment with high salt; while this treatment slightly improved the digestion (Fig. 2), the mRNAs detected by footprint reads were very similar to the salt untreated sample, again determined by a high Pearson's correlation and clustering by PCA (Supp Fig 2). For consistency, only the samples prepared in the absence of cycloheximide and the presence of salt during digestion were used for the results reported below. Overall, we generated five biological replicates of libraries of footprint reads from the cleaved compacted ribosomes in the pellet fraction and the results below are from averages of the five replicates after normalizing for transcript length and library size (28).

Ribosome profiling reads are usually generated from canonical fragment sizes between 28 and 34 nucleotides (23). Surprisingly, the peak read size (36-37 nucleotides) from the ribosome protected reads was larger than this canonical ribosome protected fragment (Fig. 3B). It has been previously reported that classical ribosomes produce medium sized 27-29 nucleotide footprints and small 20-22 nucleotide footprints should the ribosome have an open A site (29). The larger reads we observed may be due to (i) an altered state of the stalled polysome, (ii) increased protection due to associated RNA binding proteins, or (iii) incomplete digestion. It is also possible that even larger reads are present since the libraries were generated after running

a gel to size select the ribosome protected RNAs (see Methods). The longer reads map better to the coding sequence (CDS) than shorter reads (Fig. 3C). Alignment at the start and stop codons showed that the excess length of the reads was due to extension at the 3'end of the footprint reads. While approximately 14 nt was protected at the 5' end, regardless of the length of the read, the extension at the 3'end increased with the read length giving a diagonal line on a plot of read length vs. distance from start or stop codon (Fig. 3D). Ribosome footprint reads should show periodicity due to the three nucleotide code in the mRNA and reads over 32 nt showed higher periodicity than the shorter reads (Fig. 3E). We had predicted that footprint reads from stalled polysomes would be most enriched at the stop codon and the 3' end of the message due to the requirement of UPF1 for the formation of stalled polysomes and the recruitment of UPF1 at the stop codon. However, we did not observe a bias for footprint reads near the stop codon. Instead, there was some bias in the large footprint reads for the first part of the message (Fig. 3F). Taken together, ribosome profiling of dissociated monosomes derived from compacted ribosomes in the RNA granule pellet from P5 rat brain reveals enrichment in large ribosomal footprints (>32 nt), which carry a prominent 3'end extension, display a preference for the first part of the mRNA CDS, and do not display a bias for the stop codon.

Analysis of mRNAs that are abundant and enriched in footprint reads from compacted ribosomes

We were interested in identifying what mRNAs represent the most abundant constituent of compacted ribosomes and what mRNAs are enriched in compacted ribosomes compared to

total mRNA levels. From the five biological replicates, we calculated the average for Abundance (reads per kb of mRNA; RPKM) of footprint reads to determine the most abundant constituents of the compacted ribosomes. To determine what mRNAs are enriched in these ribosomes, we calculated what is characteristically called the translation efficiency (abundance of ribosome footprints/abundance of total mRNA as determined by conventional RNA seq of the starting fraction) for each mRNA (Supp. Table 1). However, in the context of presumed stalled ribosomes, translation efficiency may be a misleading term, so we renamed this measure as enrichment for the present study. Gene Ontology (GO) analysis of the 100 mRNAs with the largest abundance showed significant over-representation of mRNAs encoding cytoskeletal proteins that are expressed developmentally in neuronal projection and synaptic compartments (Fig. 4A), including Map1b (Table 1), an mRNA we had previously shown to be translated in dendrites through reactivation of stalled polysomes (8, 9). While cytoskeletal mRNAs dominate the most abundant GO category and are also enriched (Table 1), they are also relatively abundant in the total mRNA population. The GO analysis for enriched mRNAs showed a significant over-representation of mRNAs encoding RNA binding proteins or proteins involved in RNA metabolism (Fig. 4B), including the gene mutated in ALS; FUS (Table 2). These results were similar regardless of whether the 50, 200 or 500 most abundant or enriched mRNAs were selected for the analysis (Supp. Table 2,3)

We next determined if mRNAs regulated by elongation or initiation in the nervous system were over-represented in our samples either through abundance or enrichment. Strikingly, mRNAs whose translation is regulated by elongation through eEF2 phosphorylation (30) were significantly abundant and enriched in the footprint reads (Fig. 5A, B). In contrast,

mRNAs regulated by initiation, through eIF4E phosphorylation (31), TOR activation (including TOP mRNAs) (32) or through upstream ORFs after eIF2 alpha phosphorylation was increased by the stimulation of metabotropic glutamate receptors (33), were not enriched in the preparation (Fig. 5A, B). While TOP mRNAs, mainly encoding ribosomal subunits are relatively abundant and known to be transported in neurons, we observed a significant de-enrichment of footprint reads from TOP mRNAs in footprint libraries of the compacted ribosomes (Fig. 5A), consistent with the notion that transported mRNAs blocked at initiation are not enriched in this preparation.

We next determined whether mRNAs that are predicted to be enriched in stalled polysomes are particularly enriched/abundant in the footprint reads of compacted ribosomes. A recent publication identified mRNAs protected by ribosomes resistant to ribosomal run-off in neuronal slices (19). While this publication focused on the proportion of each mRNA for which ribosomes ran-off, their data also identified the mRNAs with the most protected fragments remaining after a long period of run-off (60 min). The 200 runoff-resistant mRNAs with the most reads remaining at 60 minutes were significantly enriched and abundant in our preparation (Fig. 5C, D). FMRP, an RBP highly enriched in the sedimented pellet, has been associated with stalled polysomes (17, 18). The mRNAs associated with FMRP in neurons, identified through cross-linking immunoprecipitation (CLIP), are also more resistant to ribosome run-off (17). We examined the abundance/enrichment of two separate FMRP CLIP studies from brain tissue and both were significantly enriched/abundant in our preparation (17, 34) (Fig. 5C, D).

We also examined several other datasets to better evaluate the mRNAs with footprint reads in the pelleted compact ribosomes. Secreted mRNAs are stalled by their signal peptide and then co-translationally inserted into the endoplasmic reticulum (ER). The transport of

secreted mRNAs stalled at elongation would also involve the transport of ER in the RNA granules; while transportation of ER is possible, we suspected that secretory mRNAs would be de-enriched in the RNA granules and this was indeed the case (Fig. 5E). Recently, mRNAs that are preferentially translated from monosomes in neuronal processes were identified (35). While we predicted that these mRNAs would also be depleted from our preparation, they were significantly enriched (Fig. 5E), while mRNAs preferentially transported in polysomes were not significantly enriched (Fig. 5E). However, total mRNA levels for the preferentially polysomal transported mRNAs were significantly higher than the total mRNA levels for the preferentially monosomal translated mRNAs at this developmental timepoint (158 ± 17 polysome RPKM (n=327) vs. 72 ± 4 monosome RPKM (n=458), S.E.M, $p<0.001$ Student's t-test).

Since translation from stalled polysomes is implicated in neurodevelopmental disorders, we examined if protected reads from Autism-related genes from the Simons Foundation Autism Research Initiative (SFARI) database were enriched and abundant in the protected reads. These mRNAs were significantly enriched in our data compared to all mRNAs (Fig. 5G, H). There is a significant overlap with FMRP CLIPped mRNAs in this dataset, but both FMRP CLIPped SFARI mRNAs and non-CLIPped SFARI mRNAs were equally enriched in the protected reads (Fig. 5 G, H).

Finally, some of the enrichments we observed may be due to our preparation being specifically from the nervous system. To account for this bias, we repeated this analysis using only mRNAs known to be transported in neuronal processes (35). While the mRNAs known to be transported in neuronal processes were significantly enriched and abundant in our preparation (Supp Fig 3) when we restricted the total set of mRNAs to only include this set, all of the results

above were replicated (Supp Fig 4) suggesting our results are not biased due to enrichment of neuronally transported mRNAs.

The ribosome protected reads are enriched in sequences matching FMRP CLIPs

If stalled polysomes are indeed enriched in the compacted ribosome-containing pellet, the footprint reads should help identify where on the mRNA ribosomes are stalled. Examination of the distribution of the large reads (>32 nt) on individual messages revealed highly non-uniform distribution of reads on mRNAs (Fig. 6A). As peaks in ribosome protected fragments are often non-reproducible (36), we used stringent criteria to identify peaks. First, peaks were defined for each mRNA in each library based on a maximum value higher than the average RPKM for the mRNA and a minimum width of 18 nucleotides at the half maximum height. Second, the peak had to be present in at least 3 of the 5 biological replicates (Fig. 6B; Supp Fig. 5). Using these stringent criteria, we identified 766 peaks in 524 mRNAs (Supp Table 4). While 90% of the mRNAs had only one or two peaks, Map1B, the mRNA most associated with stalled polysomes, had the largest number of identified peaks (10; Fig. 6A).

Contrary to our initial hypothesis that peaks representing stalled footprint reads would be clustered around the stop codon, only 6 of the 766 total peaks were at the stop codon, and, similar to the overall coverage of footprint reads, the peaks were biased to the first half of the message with an average position of 0.37 ± 0.26 (SD) where 0 is the start codon, and 1 is the stop codon. The average length of the sequence within a peak was 36 nt ± 6 nt, similar to the most common footprint read size (Fig. 3B), and there was no clear evidence for the presence of

collided ribosomes (peaks of twice the size of a ribosome protected fragments). We also identified peaks in the smaller reads, but only 65% of these peaks were in the CDS, and there was little overlap with the peaks in the long reads. In contrast, 94% of the peaks from large reads were in the CDS.

We next examined whether there were any consensus sequences in these peaks of footprint reads. We used an unbiased sequence motif search approach with the HOMER program (37). HOMER uses relative enrichment and requires a background sequence. To remove usage bias for mRNA sequences, we used similarly sized fragments from mRNAs with no peaks as our background selection. The Homer program identified three highly significant consensus sequences in the peaks above the false discovery rate determined by the program (Fig. 6C). Notably, the most significant consensus sequence (p=1e-67) matched a consensus sequence (WGGA) previously derived from analyzing FMRP CLIP sequences (20, 21), which also overlapped with the consensus sites for m6A methylation (RGACH or RRACT) in the nervous system (24) (Fig. 6D-F). The motif was not biased to the start or end of the protected reads (average position of motif in reads was 19 ± 9 bases) and thus does not represent sequences selected because they were difficult for nucleases to digest, since sequences selected due to resistance to digestion would be at the end or beginning of the read. Analysis of 36 nt in front of the peaks or 36 nt behind the peaks did not result in a HOMER consensus sequence above the false discovery rate determined by the program. Similarly, peaks identified using small or medium-sized reads also did not result in a HOMER consensus sequence above this cut-off.

We also performed a directed search of the peaks for sequences matching consensus binding sites of RBPs (38) using FIMO (39). Consistent with the non-biased search, FMRP CLIP

consensus sites had the most matches (Fig. 6D, E; Supp Table 5). Strikingly, these motifs mainly recognized purine-rich sequences (Fig. 6E; 76% of nucleotides in the top 10 RBP motifs are purines). However, the overall peak nucleotides are not enriched in purines compared to the background sequences (both at 57% purines). The abundance of purines in the consensus peaks suggests a possible role for the RBP PURA, a protein highly enriched in RNA Granules, but without a known consensus binding site. Of the 766 peaks, there were 415 peaks with a WGGA site, and also many peaks with multiple WGGA sites in the peak making a total of 615 WGGA sites in consensus peaks (Fig. 6E; Supp Table 4). The number of multiple matches (39% of peaks with a WGGA have multiple WGGA sites) was far in excess of those found in the corresponding background sequences (11 ± 1 %, from 10 separate selections of 750 background sequences with a WGGA peak). Of the 615 WGGA sites found in the peaks, 182 WGGA sites overlapped with an m6A site (examples shown in Fig. 6F). This percentage of WGGA sites overlapping with an RGACH site (30%) was significantly more than the number of overlaps seen in WGGA sites in our background samples (19 + 1.3%; 10 separate selections of 1300 WGGA sites from background sequences). Moreover, there were also many examples where m6A sites and WGGA sites were in the same peak but did not overlap. Of the 415 peaks with a WGGA site, 315 (76%) had an m6A consensus site in the same peak significantly more than were found in the background sample (39 ± 4%, 10 separate selections of 750 background sequences with a WGGA peak).

There were also many matches to other FMRP CLIP consensus sites (Fig. 6D). After including GVAGAW and GACAAG over 80% of the 766 consensus peaks contained an FMRP CLIP

sequences or m6A consensus sites suggesting a strong sequence bias for these sites in the regions of the mRNA enriched in footprint reads (Supp Table 4).

Discussion

The granule fraction containing compacted ribosomes is an enriched preparation for stalled ribosomes

Previous reports showed that increased density could be used to segregate compacted ribosomes from normal polysomes (7, 12-14). However, these studies did not show that these compacted ribosomes are stalled. An association between stalled polysomes and compacted ribosomes was first made based on FMRP association (17), however in that study the RNA granules were not isolated by sedimentation. Here, we show that the sedimented pellet (granule fraction) containing compacted ribosomes is not only enriched for FMRP (as had been previously shown (13)), but also for UPF1, an additional protein functionally implicated in stalled polysomes (8). Moreover, mRNAs that were previously shown to resist ribosomal run-off by initiation inhibitors (19), or to be regulated at the elongation stage (30) were over-represented in the footprint reads generated from this fraction. All of these findings are consistent with the granule fraction enriching for ribosomes that were stalled in neurons.

It should be noted that this preparation represents a snapshot of P5 brains. This sedimentation protocol is complicated by the presence of myelin at later stages of development (13). Still, if feasible, it is likely that the mRNAs found in the granule fraction would be different at the adult stage since many of mRNAs with the most abundant footprint reads that we found in this fraction are developmentally implicated in neuronal outgrowth. A related protocol developed for E18 rat brains showed fewer ribosomes when performed in adults (14). However, stalled polysomes and some of the mRNAs isolated here (notably Map1b) are implicated in

mGluR-LTD, a plasticity that is present mainly in mature brains (40). Thus, it is plausible that the role of stalled polysomes may change as the brain develops.

Cytoskeletal and RNA binding proteins represent abundant and enriched mRNAs, respectively

Axon and dendrite outgrowth, coupled with synapse formation, occur at high rates in P5 rat brains (41). Indeed, the most abundant footprint reads are on cytoskeletal mRNAs and mRNAs encoding proteins that are highly enriched in growth cones (such as 14-3-3 proteins)(42). Indeed, the most abundant number of footprint reads is found on beta-actin, whose local translation in both axons and dendrites is important for neuronal outgrowth (43, 44). Somewhat more surprising was the high enrichment for RNA binding proteins in the footprint reads (Fig. 4). This suggests an important homeostatic aspect for translation from stalled polysomes, where the increased translation of RBPs will have critical effects on the translation of other messages not necessarily present in stalled polysomes.

Identification of conserved motifs enriched in footprint read peaks from compacted ribosomes.

The finding that the footprint reads derived from the compacted ribosome granule fraction are distributed mainly in large peaks is consistent with the enrichment of stalled ribosomes in this fraction. These peaks are strikingly enhanced in sequences previously defined as enriched in FMRP CLIPs (Fig. 6). This is consistent with the strong enrichment of FMRP in the granule fraction and the abundance and enrichment of mRNAs previously identified as associated with FMRP using CLIP experiments (Fig. 5). While most of the FMRP CLIP consensus sequences have not been directly shown to bind to FMRP, the WGGA sequence can be directly

bound by FMRP (21). However, the major FMRP binding sites (G quartets and Kissing sequence) are not enriched in FMRP CLIP consensus sites (20). Moreover, if ribosomes protect this sequence, it is unclear how FMRP would gain access to these sequences. However, it is possible that FMRP initially binds this sequence and this is followed by ribosome occupation and stalling. It is also possible that these sequences are enriched in FMRP CLIPs because FMRP is specifically associated with stalled ribosomes and these sequences specify where ribosomes would be stalled independently of FMRP. In this scenario, the sequences would not be directly bound by FMRP. Instead, FMRP would be crosslinked to sequences near to, but not protected by the ribosome. Since the CLIP sequences are approximately 100 bp, this is entirely consistent with both the CLIP and ribosome footprint read data.

The sites are also enriched in a consensus site for m6A modification. Interestingly, mRNAs with m6A sites are selectively transported in neurons (45) and this methylation plays an important role in neurodevelopment (46). Moreover, mRNAs that are associated with FMRP by CLIP experiments have previously been shown to be highly enriched for m6A modifications in neurons (24) and this has been proposed to play a role in FMRP-mediated nuclear export (47, 48). However, whether FMRP directly binds to m6A, interacts with m6A readers, or is associated with m6A through some other indirect interaction is not clear. There has been some indication that m6A directly leads to the stalling of ribosomes (49) or it may be that some m6A reader in neurons is important for the stall. Again, it is unclear how the reader would get access to the mRNA sequences protected by the ribosome, but similar to FMRP the sequence may be recognized first and then later occupied by the ribosome. It will be interesting in the future to determine the specific relationship between m6A methylation and stalled polysomes and

whether initial findings of specific roles for m6A methylation in the developing brain are linked to their possible role in stalling translation in RNA granules.

Stalled Polysomes or Stalled ribosomes?

Stalled polysomes would be predicted to show a series of protected reads following the initial stalled peak. These peaks downstream of the stall site, corresponding to ribosomes stalled behind the leading ribosome, would not be predicted to show any specificity (consensus sequences), as the leading ribosome dictates the stall. If, for example, four ribosomes were stalled behind the leading ribosome, this would mean only 20% of the ribosomes would be stalled on a stalling sequence. This model does not fit the data we observe. The majority of consensus peaks contain sites previously observed for FMRP Clips or m6A consensus sites, and there is little periodicity in the placement of peaks.

We did not search for disome or higher periodicity ribosome footprints (50), as we performed ribosome footprints on purified monosomes. However, we did not observe other signs of collided ribosomes, such as periodic peaks or a size of peaks corresponding to two or more ribosomes. Indeed, ribosomes that are stalled due to collision with a downstream ribosome recruit a collection of specific factors to resolve the stall and rescue the ribosomes (51). These factors are not present in proteomics of RNA granules (13-15) and indeed, if these stalled ribosomes are meant to be transported and reactivated, they need to be protected from the surveillance mechanisms normally used for prematurely stopped or collied ribosomes, which lead to unstalling and/or degradation of the proteins (51). Our data is consistent with a model in which a controlled form of stalling attracts specific factors in neurons, such as FMRP,

that likely play a role in the compaction the stalled ribosome/polysome and package the structure into a granule for transport before collisions occur. Thus, the compacted ribosomes may not represent collided ribosomes on a polysome, but collections of ribosomes from distinct mRNAs that are packaged together. Indeed, the mRNAs that are preferentially translated by monosomes in dendrites are enriched in this compacted ribosome pellet. The packaging of many distinct mRNAs in the same mRNA granule is consistent with the finding that *in situ* RNA granules containing stalled polysomes probably contain many distinct mRNAs (52).

Conclusions

While there have been assumed links between the compacted ribosomes that sediment in sucrose gradients, stalled polysomes identified by resistance to ribosomal run-off in neuronal dendrites and FMRP's association with stalled polysomes, there were previously no direct connections between these various lines of research. Identifying an enrichment for FMRP CLIP consensus sequences in protected reads in compacted ribosomes from the pellet of sucrose gradients establishes these links. Notably, most investigations of translation regulation in neuronal tissues do not consider the pellet fraction after separation of polysomes using sucrose gradients. Thus, they are not accounting for this pool of translationally repressed mRNAs. For ribosome profiling from the initial polysome pellet, the use of translation efficiency needs to be re-evaluated, because many of the ribosomes in this initial pellet are presumably stalled. Moreover, some studies only include footprints of the canonical size and may exclude the larger

size footprints that we observed in this preparation. Thus, these results have large implications for the interpretation of many studies on neuronal translation.

Materials and Methods

Purification of RNA Granules.

RNA Granules were purified, from whole brain homogenate harvested from five-day-old (P5) Sprague Dawley rats, using a protocol adapted from a previous purification (13). Five P5 rat brains were homogenized in RNA Granule Buffer (20 mM TRIS-HCl pH 7.4 [Fisher; BP152-1], 150 mM NaCl [Fisher; BP358-212], 2.5 mM $MgCl_2$ [Fisher; M33-500]) supplemented with 1 mM DTT [Sigma; D9163], 1 mM EGTA [Sigma; E8145], EDTA-free protease inhibitor [Roche; 04693132001]. Note that cycloheximide [Abcam; ab120093] was not added to the homogenization buffer unless explicitly stated. Homogenate was centrifuged 15 minutes in a Thermo Scientific T865 fixed angle rotor at 9200 rpm (6117 RCF) at 4˚C to spin down cellular debris. The supernatant was treated with 1% IGEPAL CA-630 [Sigma; I8896] for 5 minutes at 4˚C on a rocker. The sample was then loaded onto a 2 ml 60% sucrose [Calbiochem; 8550] cushion (dissolved in supplemented RNA Granule Buffer) in a Sorvall 36 ml tube [Kendro; 3141, Thermo Scientific], filled to top with additional RNA Granule Buffer and centrifuged for 2 hours in a Thermo Scientific AH-629 swing-bucket rotor at 28500 rpm (56660 RCF) at 4˚C to achieve the polysome pellet. The pellet was re-suspended in RNA Granule Buffer, gently dounced and loaded over a 15-60% linear sucrose gradient (gradient was made with RNA Granule Buffer) that was prepared in advance using a gradient maker [Biocomp Gradient Master], and centrifuged for 45 minutes at 28500 rpm (56660 RCF) at 4˚C in an AH-629 swing bucket rotor. Fractions of 3.5 ml were then collected from the top, and the remaining pellet was rinsed once and then resuspended using RNA Granule Buffer. For some experiments, such as Electron Microscopy, the

resuspended pellet was used directly, or treated with salt and nuclease to break up the ribosome clusters into monosomes (see below). For some experiments, such as Ribosome Footprinting, the fractions were precipitated overnight at -20 ˚C by adding 7 ml of chilled 100% ethanol. The precipitated samples were then centrifuged for 45 minutes at 3900 RPM (2177 RCF) at 4˚C in an Eppendorf 5810/5810 swing bucket rotor before collection using RNA Granule Buffer.

Nuclease and Salt Treatments.

To break up the stalled polysomes compacted in the RNA Granule into monosomes for ribosome footprint analysis, we treated the pellet fraction with salt and nuclease. The pellet was incubated with RNA granule buffer containing 400 mM NaCl for 10 minutes at 4˚C on a rocker(13). Before nuclease treatment, the NaCl concentration was reduced back to 150 mM by diluting the sample with a NaCl-free RNA Granule Buffer. The sample was then treated with 100 U of RNase I [100 U/µl; Ambion AM2294, Thermo Fisher] for 30 minutes at 4˚C on a rocker. The nuclease was quenched with 100 U of SuperaseIN [20 U/µl; Invitrogen #AM2696, Thermo Fisher], and the samples were re-run on a fresh 15-60% sucrose gradient to separate mononsomes. Fraction 2 was precipitated overnight at -20 ˚C by adding 7 ml of chilled 100% ethanol. The precipitated samples were then centrifuged for 45 minutes at 3900 RPM (2177 RCF) at 4˚C in an Eppendorf 5810/5810 swing bucket rotor before collection using RNA Granule Buffer.

Electron microscopy

The untreated pellet fraction, or after treatment with either high-salt or nuclease or both treatments were deposited on the EM grids. The ribosome concentration of each sample was adjusted to ~80 ng/μl (~25 nM) using RNA Granule Buffer before applying them to the grids. In the case of the sample treated with both nuclease and high salt the concentration of sample applied to the grid was 9.2ng/μl (2.9 nM). We used 400-mesh copper grids freshly coated with a continuous layer of thin carbon for these experiments. Grids were glow-discharged at 5 mA for 15 seconds and then floated on a 5-μl drop of the diluted sample for 2 min. Excess of sample was blotted away with filter paper (Whatman #1), and to stain them, they were subsequently floated in a 5-μl drop of a 1% uranyl acetate solution for 1 min. Excess of stain was blotted away, and the grids were dried on air and store in regular grid boxes. The EM images were acquired on a Tecnai F20 electron microscope operated at 200 kV using a room temperature side entry holder. Images were collected in a Gatan Ultrascan 4000 4 k × 4 k CCD Camera System Model 895 at a nominal magnification of 60,000x. Images produced by this camera had a calibrated pixel size of 1.8 Å/pixel. The total electron dose per image was ~50 e-/Å2. Images were collected using a defocus of approximately −2.7 μm. Images were prepared for figures using the Adobe Photoshop program.

Immunoblotting and Quantification of Enrichment.

For immunoblotting, SDS sample buffer was added to each sample before loading onto a 10% acrylamide gel. The resolved proteins were either stained with Coomassie brilliant blue [Fisher; 821616] or transferred onto a 0.45 μm nitrocellulose membrane (Bio Rad; 1620115) for immunoblotting. The membranes were blocked with 5% BSA (Sigma; 9647) in Tris Buffered Saline with Tween (TBS-T [Tris: Fisher, BP152-1; NaCl: Fisher, BP358-212; Tween: Fisher, BP337])

before incubation with the following primary antibodies in 1:1000 dilution: rabbit anti-S6 [Cell Signaling #2217], rabbit anti-FMRP [Cell Signaling; #4317], rabbit anti-eEF2 [Cell Signaling; #2332S], rabbit anti-Upf1 [Abcam; ab133564], mouse anti-Stau2 [Medimabs; MM0037-P], rabbit anti-PurA [Abcam; ab79936]. Membranes were washed with TBS-T after incubation. Detection was done using HRP-conjugated secondary antibodies [Invitrogen] before ECL [Perkin Elmer; NEL105001EA] reaction and imaging using the Bio Rad ChemiDoc digital imager. For quantification of RBP enrichment, membranes were stripped with a Western Blot Stripping Buffer [ZmTech Scientific; S208070] and re-probed with rabbit anti-S6 [Cell Signaling #2217] antibody, followed by detection with HRP-conjugated secondary antibodies [Invitrogen].

Quantification of signal intensity was done using ImageJ software. We selected full lane ROIs and quantified single bands corresponding to the observed kDa size of each protein. We then used the corresponding anti-S6 signal intensity to quantify the amount of examined protein per S6 ribosomal protein for each fraction. For each experiment the protein/S6 value was normalized to the protein/S6 value from the starting material. Following this normalization, biological replicates were averaged. For salt and nuclease experiments, we calculated the proportion of Granule Fraction Polysomes (Pellet) digested into monosomes (Fraction 2) by doing the following calculation: Fraction 2/(Fraction 2 + Pellet). In one experiment we added Fraction 1 and 2 together because we detected some residual monosomes in Fraction 1, likely caused by spillover of Fraction 2 from pipetting.

Footprint sequencing library construction and sequencing.

Fraction 2 from the monosome purification was centrifuged at 4,444 x g at 4 °C for 60 min, the isopropanol taken off, the pellet air-dried, and re-suspended in 10 mM Tris-HCl pH 7. We then proceeded with a standard hot phenol-acid extraction (Acid Phenol:Chloroform mix 125:24:1, Invitrogen, AM9722). The samples were resuspended in nuclease-free water and quantified on a Spectrophotometer.

All samples were depleted of ribosomal RNA using the Ribo-Zero Gold (Human/Mouse/Rat) Kit (Illumina) (replicates 1-3 and cycloheximide treated) or the NEBNext® rRNA Depletion Kit (Human/Mouse/Rat)(E6350, NEB) (replicates 4-5 and no salt treatment) The footprint samples were size-selected using the 17 and 34 nt markers as a guide on a 15% TBE-Urea polyacrylamide gel (Thermo Scientific), while total RNA samples were randomly heat fragmented by incubating at 95 °C in an alkaline fragmentation solution (50 mM $NaCO_3$ pH 9.2, 2 mM EDTA) for 40 min, yielding comparably sized fragments to footprints. All samples were de-phosphorylated using PNK (T4, NEB). The quality and concentration of samples was assessed by running an Agilent Small RNA chip (Agilent Technologies), and sequencing libraries were generated using the NEXTflex™ Small RNA Sequencing Kit v3 (PerkinElmer, NOVA-5132-06), according to the manufacturer's instructions. Samples were balanced and pooled for sequencing with Edinburgh Genomics on NovaSeq S1/2 flow cells yielding 50 bp paired-end reads.

Riboseq and RNAseq data analysis

Adaptor sequences and low-quality score containing bases (*Phred score < 30*) were trimmed from reads using Cutadapt v2.8 (*-j 8 -u 4 -u -4 -Z -o*) (53). Noncoding RNAs were removed by custom scripts following mapping these contaminant reads using Bowtie2 v2.3.5 (*--phred33 --*

very-sensitive) (54). The unmapped reads were aligned to reference rat genome (Rnor_6.0.94) using *STAR v2.7.3a (55)*with options previously described (35) (*--twopassMode Basic --twopass1readsN -1 -- seedSearchStartLmax 15 --outSJfilterOverhangMin 15 8 8 8 --outFilterMismatchNoverReadLmax 0.1*). QuantMode with STAR was used to obtain genomic and transcript coordinates (55). Assigning RPF reads to genomic features (CDS, UTRs) was based on Genome annotation (Rnor_6.0.94). Only a single transcript isoform, with the highest APPRIS score (56), was considered per gene. All raw reads from these experiments have been deposited at the GEO dataset at NCBI (number to be confirmed).

Raw counts obtained using *featurecounts v2.2.0* (57) were analyzed using the R package *limma* (28). From this package, transcript abundance was obtained for riboseq and RNAseq data in reads per kb of mRNA (RPKM). Enrichment was determined for each transcript by dividing riboseq RPKM by RNAseq RPKM. False discovery rates (FDR) and p values were determined for enrichment in the package, and nominal p-values were corrected for multiple testing using the Benjamini-Hochberg method.

Gene Ontology (GO) enrichment analysis was performed with *gProfileR* (58). Enrichment p-values were based on a hypergeometric test using the set of known Rat genes as background. Sample correlation based on normalised read count was obtained using the R package *limma* (28). We used Ribowaltz (59) to assign length dependant P-site correction and periodicity assignment to each reads.

Peak identification

Sites enriched with RPF were identified using normalized riboseq profiles using peak identification function within *IDPmisc,* a R package (60). A peak width of minimum 18 nt and peak height above the mean peak height within the transcript was used as the criteria to define a region with significant enrichment of RPFs. Only peaks region that overlap (at least 90% nt overlap) in at least 3 samples were considered using BEDtools v2.29.2 (61)

Motif analysis:

The peak regions were scanned for known human RNA-binding protein motifs using the *FIMO* program (39) which is part of MEME Suite (62). Only search results with a p-value less than the threshold of 0.5 were considered. De novo motif finding was performed using peak from RPFs. *HOMER* tool (37) was used for this analysis (*findMotifs.pl -rna)*. Background sequences were randomly selected from transcripts with no peaks.

Acknowledgements

This work was supported by Canadian Institute of Health Research (CIHR) project grant 374967 to W.S.S, an award from the Azrieili foundation to WSS and JO, and by a Fondation Santé grant and a FORTH-IMBB-BR start-up grant to CGG. W.S.S. is a James McGill Professor and SMJ is supported by a Queen's University Belfast Patrick Johnston Research Fellowship. We thank staff at the Facility for Electron Microscopy Research (FEMR) at McGill University. FEMR is supported by the Canadian Foundation for Innovation, the Quebec government and McGill University. Figures were created using Biorender.com

References

1. C. Glock, M. Heumüller, E. M. Schuman, mRNA transport & local translation in neurons. *Current Opinion in Neurobiology* **45**, 169-177 (2017).
2. J. M. Cioni, M. Koppers, C. E. Holt, Molecular control of local translation in axon development and maintenance. *Curr Opin Neurobiol* **51**, 86-94 (2018).
3. M. Mori, J. Penney, P. Haghighi, "Regulation of synaptic homeostasis by translational mechanisms" in The Oxford Handbook of Neuronal Protein Synthesis*,* W. Sossin, Ed. (Oxford Press, 2019), 10.1093/oxfordhb/9780190686307.013.14.
4. W. S. Sossin, M. Costa-Mattioli, Translational Control in the Brain in Health and Disease. *Cold Spring Harb Perspect Biol* **11** (2019).
5. M. N. Anadolu, W. S. Sossin, "Focusing on mRNA granules and stalled polysomes amidst diverse mechanisms underlying mRNA transport, mRNA storage and local translation." in Oxford Hanbook of Neuronal Protein Synthesis*,* W. Sossin, Ed. (Oxford Press, 2020).
6. W. S. Sossin, L. DesGroseillers, Intracellular trafficking of RNA in neurons. *Traffic* **7**, 1581-1589 (2006).
7. A. M. Krichevsky, K. S. Kosik, Neuronal RNA granules: a link between RNA localization and stimulation-dependent translation. *Neuron* **32**, 683-696 (2001).
8. T. E. Graber *et al.*, UPF1 Governs Synaptic Plasticity through Association with a STAU2 RNA Granule. *J Neurosci* **37**, 9116-9131 (2017).
9. T. E. Graber *et al.*, Reactivation of stalled polyribosomes in synaptic plasticity. *Proc Natl Acad Sci U S A* **110**, 16205-16210 (2013).
10. P. K. McCamphill, C. A. Farah, M. N. Anadolu, S. Hoque, W. S. Sossin, Bidirectional regulation of eEF2 phosphorylation controls synaptic plasticity by decoding neuronal activity patterns. *J Neurosci* **35**, 4403-4417 (2015).

11. E. Barbarese *et al.*, Protein translation components are colocalized in granules in oligodendrocytes. *Journal of Cell Science* **108**, 2781-2790 (1995).

12. A. Aschrafi, B. A. Cunningham, G. M. Edelman, P. W. Vanderklish, The fragile X mental retardation protein and group I metabotropic glutamate receptors regulate levels of mRNA granules in brain. *Proc Natl Acad Sci U S A* **102**, 2180-2185 (2005).

13. R. El Fatimy *et al.*, Tracking the Fragile X Mental Retardation Protein in a Highly Ordered Neuronal RiboNucleoParticles Population: A Link between Stalled Polyribosomes and RNA Granules. *PLoS Genet* **12**, e1006192 (2016).

14. G. Elvira *et al.*, Characterization of an RNA granule from developing brain. *Molecular & cellular proteomics : MCP* **5**, 635-651 (2006).

15. Y. Kanai, N. Dohmae, N. Hirokawa, Kinesin transports RNA: isolation and characterization of an RNA-transporting granule. *Neuron* **43**, 513-525 (2004).

16. K. B. Garber, J. Visootsak, S. T. Warren, Fragile X syndrome. *Eur J Hum Genet* **16**, 666-672 (2008).

17. J. C. Darnell *et al.*, FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell* **146**, 247-261 (2011).

18. S. Ceman *et al.*, Phosphorylation influences the translation state of FMRP-associated polyribosomes. *Hum Mol Genet* **12**, 3295-3305 (2003).

19. S. Shah *et al.*, FMRP Control of Ribosome Translocation Promotes Chromatin Modifications and Alternative Splicing of Neuronal Genes Linked to Autism. *Cell Rep* **30**, 4459-4472 e4456 (2020).

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

21. M. Ascano, Jr. *et al.*, FMRP targets distinct mRNA sequence elements to regulate protein expression. *Nature* **492**, 382-386 (2012).

22. Y. K. Kim, L. E. Maquat, UPFront and center in RNA decay: UPF1 in nonsense-mediated mRNA decay and beyond. *RNA* **25**, 407-422 (2019).

23. N. T. Ingolia, Ribosome profiling: new views of translation, from single codons to genome scale. *Nat Rev Genet* **15**, 205-213 (2014).

24. F. Zhang *et al.*, Fragile X mental retardation protein modulates the stability of its m6A-marked messenger RNA targets. *Hum Mol Genet* **27**, 3936-3950 (2018).

25. E. Chen, M. R. Sharma, X. Shi, R. K. Agrawal, S. Joseph, Fragile X mental retardation protein regulates translation by binding directly to the ribosome. *Mol Cell* **54**, 407-417 (2014).

26. M. R. F. Reijnders *et al.*, PURA syndrome: clinical delineation and genotype-phenotype study in 32 individuals with review of published literature. *J Med Genet* **55**, 104-113 (2018).

27. J. E. Heraud-Farlow, M. A. Kiebler, The multifunctional Staufen proteins: conserved roles from neurogenesis to synaptic plasticity. *Trends Neurosci* **37**, 470-479 (2014).

28. M. E. Ritchie *et al.*, limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47 (2015).

29. C. C. Wu, B. Zinshteyn, K. A. Wehner, R. Green, High-Resolution Ribosome Profiling Defines Discrete Ribosome Elongation States and Translational Regulation during Cellular Stress. *Mol Cell* **73**, 959-970 e955 (2019).

30. J. W. Kenney *et al.*, Eukaryotic elongation factor 2 kinase regulates the synthesis of microtubule-related proteins in neurons. *J Neurochem* **136**, 276-284 (2016).

31. I. S. Amorim *et al.*, Loss of eIF4E Phosphorylation Engenders Depression-like Behaviors via Selective mRNA Translation. *J Neurosci* **38**, 2118-2133 (2018).

32. C. C. Thoreen *et al.*, A unifying model for mTORC1-mediated regulation of mRNA translation. *Nature* **485**, 109-113 (2012).

33. G. V. Di Prisco *et al.*, Translational control of mGluR-dependent long-term depression and object-place learning by eIF2alpha. *Nat Neurosci* **17**, 1073-1082 (2014).

34. T. Maurin *et al.*, HITS-CLIP in various brain areas reveals new targets and new modalities of RNA binding by fragile X mental retardation protein. *Nucleic Acids Res* **46**, 6344-6355 (2018).

35. A. Biever *et al.*, Monosomes actively translate synaptic mRNAs in neuronal processes. *Science (New York, N.Y.)* **367** (2020).

36. B. Liu *et al.*, Optimization of ribosome profiling using low-input brain tissue from fragile X syndrome model mice. *Nucleic Acids Res* **47**, e25 (2019).

37. S. Heinz *et al.*, Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**, 576-589 (2010).

38. E. L. Van Nostrand *et al.*, A large-scale binding and functional map of human RNA-binding proteins. *Nature* **583**, 711-719 (2020).

39. C. E. Grant, T. L. Bailey, W. S. Noble, FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**, 1017-1018 (2011).

40. E. D. Nosyreva, K. M. Huber, Developmental switch in synaptic mechanisms of hippocampal metabotropic glutamate receptor-dependent long-term depression. *J Neurosci* **25**, 2992-3001 (2005).

41. B. D. Semple, K. Blomgren, K. Gimlin, D. M. Ferriero, L. J. Noble-Haeusslein, Brain development in rodents and humans: Identifying benchmarks of maturation and vulnerability to injury across species. *Prog Neurobiol* **106-107**, 1-16 (2013).

42. C. B. Kent *et al.*, 14-3-3 proteins regulate protein kinase a activity to modulate growth cone turning responses. *J Neurosci* **30**, 14059-14067 (2010).

43. K. M. Leung *et al.*, Asymmetrical beta-actin mRNA translation in growth cones mediates attractive turning to netrin-1. *Nat Neurosci* **9**, 1247-1256 (2006).

44. T. Eom, L. N. Antar, R. H. Singer, G. J. Bassell, Localization of a beta-actin messenger ribonucleoprotein complex with zipcode-binding protein modulates the density of dendritic filopodia and filopodial synapses. *J Neurosci* **23**, 10433-10444 (2003).

45. D. Merkurjev *et al.*, Synaptic N(6)-methyladenosine (m(6)A) epitranscriptome reveals functional partitioning of localized transcripts. *Nat Neurosci* **21**, 1004-1014 (2018).

46. J. Widagdo, V. Anggono, The m6A-epitranscriptomic signature in neurobiology: from neurodevelopment to brain plasticity. *J Neurochem* **147**, 137-152 (2018).

47. P. J. Hsu *et al.*, The RNA-binding protein FMRP facilitates the nuclear export of N (6)-methyladenosine-containing mRNAs. *J Biol Chem* **294**, 19889-19895 (2019).

48. C. J. Westmark, B. Maloney, R. S. Alisch, D. K. Sokol, D. K. Lahiri, FMRP Regulates the Nuclear Export of Adam9 and Psen1 mRNAs: Secondary Analysis of an N(6)-Methyladenosine Dataset. *Sci Rep* **10**, 10781 (2020).

49. J. Choi *et al.*, N(6)-methyladenosine in mRNA disrupts tRNA selection and translation-elongation dynamics. *Nat Struct Mol Biol* **23**, 110-115 (2016).

50. P. Han *et al.*, Genome-wide Survey of Ribosome Collision. *Cell Rep* **31**, 107610 (2020).

51. A. R. Buskirk, R. Green, Ribosome pausing, arrest and rescue in bacteria and eukaryotes. *Philos Trans R Soc Lond B Biol Sci* **372** (2017).

52. J. J. Langille, K. Ginzberg, W. S. Sossin, Polysomes identified by live imaging of nascent peptides are stalled in hippocampal and cortical neurites. *Learn Mem* **26**, 351-362 (2019).

53. M. Martin, Cutadapt removes adapter sequences from high-trhoughput sequencing reads. *EMBnet.journal* **17**, 10-12 (2011).

54. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-359 (2012).

55. A. Dobin *et al.*, STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).

56. J. M. Rodriguez *et al.*, APPRIS: annotation of principal and alternative splice isoforms. *Nucleic Acids Res* **41**, D110-117 (2013).

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

58. J. Reimand *et al.*, g:Profiler-a web server for functional interpretation of gene lists (2016 update). *Nucleic Acids Res* **44**, W83-89 (2016).

59. F. Lauria *et al.*, riboWaltz: Optimization of ribosome P-site positioning in ribosome profiling data. *PLoS Comput Biol* **14**, e1006169 (2018).

60. F. E. Harrell Jr, M. C. Dupont (2019) Package Hmisc. in *R Foundation for Statistical Computing* (Vienna, Austria).

61. A. R. Quinlan, I. M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842 (2010).

62. T. L. Bailey *et al.*, MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res* **37**, W202-208 (2009).

Table 1

Top 20 most enriched mRNAs

Gene Name	logFC	P-value	FDR	Total RNA RPKM	Footprints RPKM
FUS	3.16434772	6.49E-10	9.56E-08	69.7831423	337.199324
FAM168A	2.94304308	7.81E-07	1.60E-05	52.1322914	130.2335745
PRRT1	2.89433174	1.85E-06	3.16E-05	26.2478904	55.2638004
NCL	2.87065155	4.50E-13	4.21E-10	63.0689626	366.6690471
RBM25L1	2.67673894	5.59E-08	2.27E-06	4.69685407	22.57140369
HNRNPD	2.63687085	5.04E-10	8.14E-08	135.136829	590.9433885
GATAD2B	2.55905455	1.39E-08	8.75E-07	34.7720695	118.9687961
SOX9	2.45296112	1.01E-05	0.00012227	25.4556107	59.89774259
UBQLN2	2.43529701	4.11E-11	1.25E-08	101.635634	385.068685
HNRNPA2B1	2.43013679	5.74E-10	8.95E-08	158.368037	668.3078643
PRR36	2.41730546	1.44E-06	2.58E-05	23.624002	46.51721335
ZMIZ1	2.37502673	7.30E-07	1.53E-05	38.2509104	89.47685878
LOC103694210	2.3707282	2.29E-07	6.30E-06	16.8011259	76.08347195
TAF15	2.36234981	2.89E-07	7.61E-06	10.1106776	38.79032835
MNT	2.36190606	3.14E-08	1.55E-06	10.7221613	28.72871691
AHCYL1	2.36154326	3.68E-08	1.70E-06	59.2481984	150.2720528
HNRNPA3	2.35579919	1.63E-10	3.46E-08	53.2766435	197.1313486
LOC108348151	2.35324232	1.93E-05	0.00020574	11.6881118	46.27995067
VAMP2	2.30892138	4.34E-07	1.00E-05	473.446629	912.2714551
SF3A2	2.29351137	3.99E-07	9.54E-06	20.2483045	41.1681424

Table 1. Top 20 most enriched mRNAs
Top 20 enriched mRNAs sorted on the fold change (FC) in global footprint reads per kilobase pre million reads (Footprint RKPM) compared to Total RNA reads per kilobase per million reads (Total RNA RPKM). mRNAs were excluded if the RPKM was predominantly due to reads less than 32 bp (>2 fold) or if the total RPKM was less than 5. Also shown are two evaluations of the significance of the fold change, the P value, and the false discovery rate (FDR). mRNAs highlighted in red match the GO term mRNA or RNA binding.

Table 2

Top 20 most abundant mRNAs

Gene Name	logFC	P-value	FDR	Total RNA RPKM	Footprints RPKM
ACTB	1.940985547	1.53E-08	9.23E-07	3311.57669	9408.327001
TUBB5	1.033645345	1.37E-05	1.55E-04	3698.919325	7404.359382
TUBB2B	1.63257495	4.98E-07	1.13E-05	1913.684313	5570.961651
YWHAE	1.711033554	7.60E-09	5.65E-07	1688.064086	4751.885604
TUBB3	1.244833653	2.40E-07	6.48E-06	2052.960915	4594.904687
YWHAZ	1.05722464	1.71E-06	2.96E-05	2004.358037	4216.85062
STMN2	1.400937863	2.81E-06	4.44E-05	1692.26231	3719.643634
CALM1	0.154694995	4.87E-01	6.19E-01	3093.888941	3625.962055
DPYSL2	1.24010485	2.46E-09	2.57E-07	1568.286012	3524.363783
YWHAG	1.82197446	1.01E-09	1.32E-07	1166.611045	3340.758587
CALM3	-0.309008441	7.89E-02	1.58E-01	3403.092834	2994.609636
CALM2	0.862970381	7.06E-04	3.78E-03	1549.807564	2953.038226
UCHL1	1.003234991	4.34E-05	3.93E-04	1439.792879	2647.37674
DPYSL3	2.223810027	1.16E-10	2.71E-08	698.2863694	2518.842074
ZWINT	0.887865185	8.91E-04	4.56E-03	1426.858532	2492.284205
MAP1B	1.865404024	1.12E-10	2.71E-08	755.4364087	2259.892443
STMN1	0.273922938	2.08E-01	3.31E-01	1722.323579	2160.340529
GAP43	0.569199253	1.27E-03	6.09E-03	1409.864902	2067.364803
CRMP1	1.496025361	3.20E-09	3.00E-07	882.8247474	1959.479377
YWHAB	1.791487719	8.44E-09	6.04E-07	578.6707707	1808.787351

Table 2. Top 20 most abundant mRNAs
Top 20 most abundant mRNAs sorted on the Footprint reads per kilobase per million reads (RKPM) compared to Total RNA reads per kilobase per million reads (Total RNA RPKM). Also shown are the log fold change (Footprint RPKM compared to total RNA RPKM). Also shown are evaluations of the significance of the fold change (the P value and the False Discovery Rate (FDR). mRNAs highlighted in blue match the GO term for cytoskeleton constituent of cytoskeletal binding.

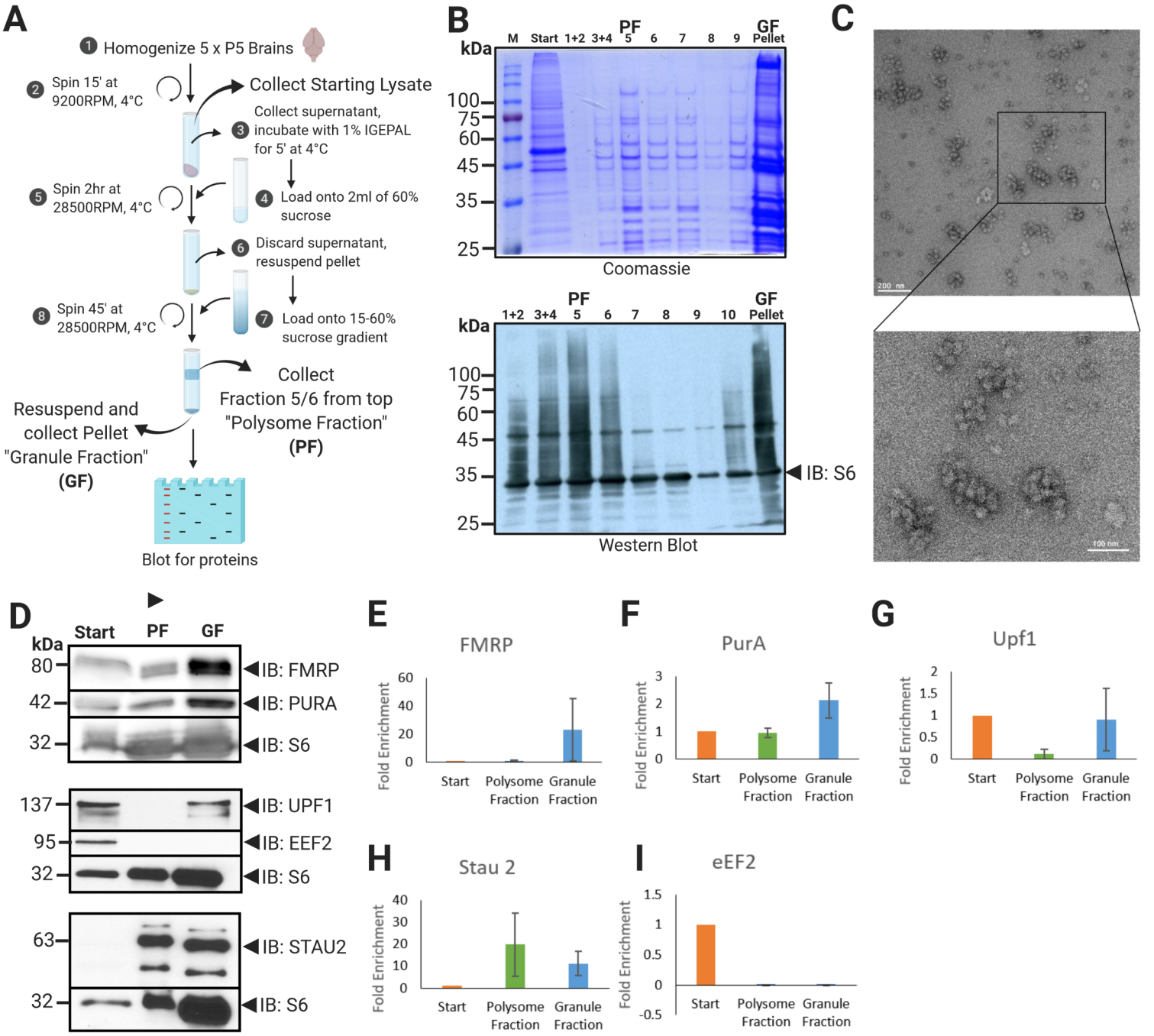


Figure 1. Isolation of a pellet enriched in compacted ribosomes.
(A) Summary of protocol for isolating the pellet from P5 rat whole brain homogenate using sucrose gradient fractionation. (B) Top: SDS-page stained with Coomassie brilliant blue showing enrichment of the characteristic distribution of ribosomal proteins in Fraction 5 (Polysome Fraction; PF) and Pellet (Granule Fraction; GF); Bottom: Immunoblot analysis demonstrating strong enrichment of S6 ribosomal protein in Fraction 5 (PF) and Pellet (GF); Lanes are described on top (M for molecular weight marker). P, pellet. (C) Negative stained electron micrograph of the Pellet fraction shows compacted ribosomes; inset showing close up of compacted ribosome structure (D) Immunoblot analysis of starting material, Polysome Fraction (PF; Fraction 5/6) and Granule Fraction (GF; Pellet) stained for RBPs and other factors implicated in RNA granules and stalled polysomes: FMRP, PURA, Upf1, Stau2 and eEF2. (E-I) Quantification of Western Blots. The fold enrichment compared to starting material normalized to levels of S6 (see methods) is shown for each RBP: FMRP (N=3), PURA (N=3), Upf1 (N=2), Stau2 (N=2), eEF2 (N=3). Error bars represent SEM.

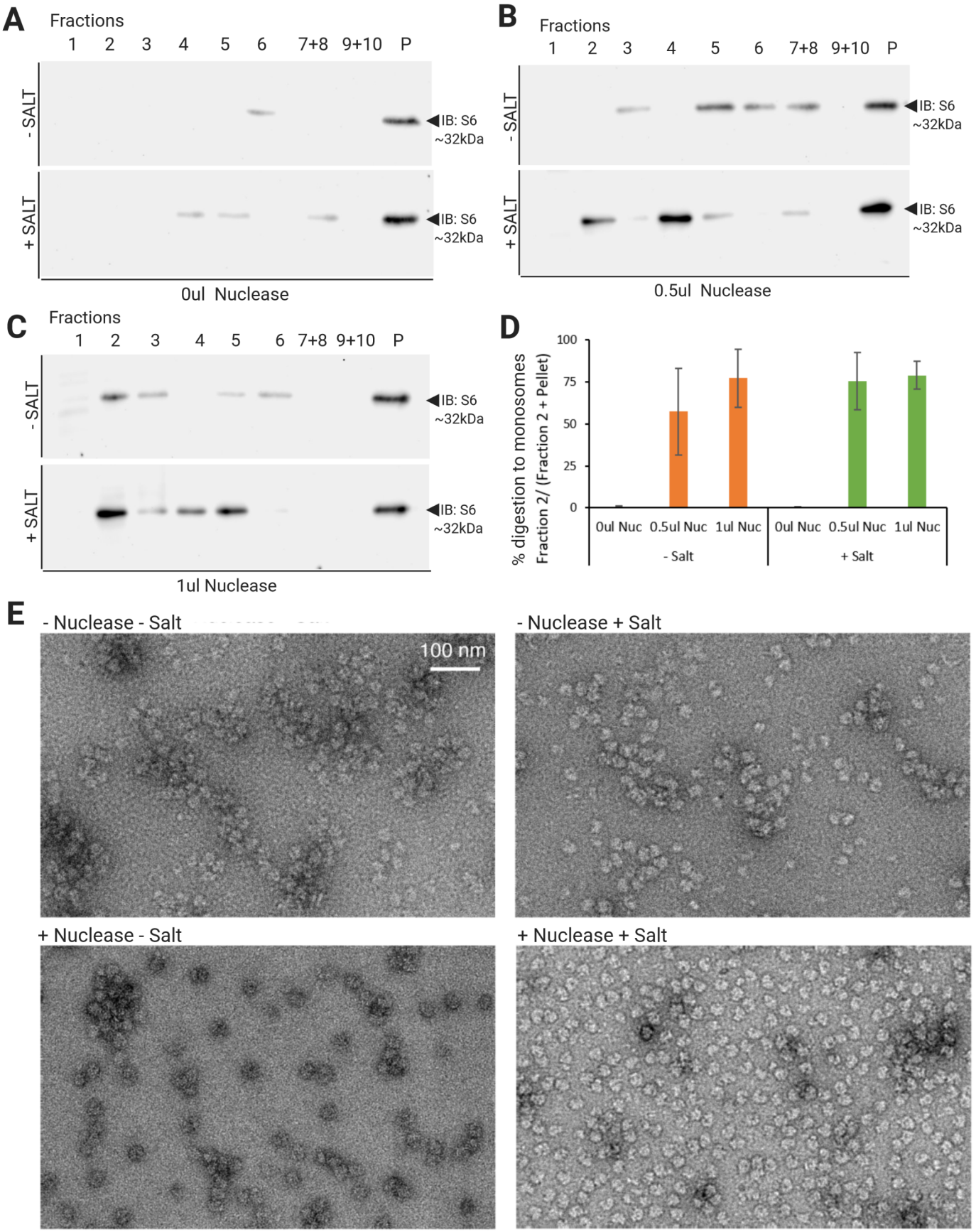


Figure 2. Cleavage of compacted stalled polysomes into monosomes
(A-C) Western Blot analysis for S6 ribosomal protein of fractions collected after first treating the pellet from the first sucrose gradient with 0ul (A), 0.5ul (B) and 1ul (C) RNase I, with or without pre-treatment of 400 mM NaCl (-SALT: TOP; +SALT: BOTTOM), followed by a second sucrose gradient (see Methods). (D) Generation of monosomes by digestion of pelleted granule fraction with RNase I. Quantification of digestion represented as the ratio of Fraction 2/(Fraction 2 + Pellet), N=4 biological replicates, error bars represent SEM. (E) Electron micrographs of negatively stained pellet fraction following treatment with RNase I, with and without pre-treatment with Salt (-Nuclease -Salt: TOP LEFT; -Nuclease +Salt: TOP RIGHT; +Nuclease – Salt: BOTTOM LEFT; +Nuclease +Salt: BOTTOM LEFT). Scale bar represents 100 nm. Note that EM images represent the pellet fraction treated with Nuclease and Salt prior to the second sucrose gradient, and not the purified monosomes in Fraction 2 that are quantified in the above Western Blots.

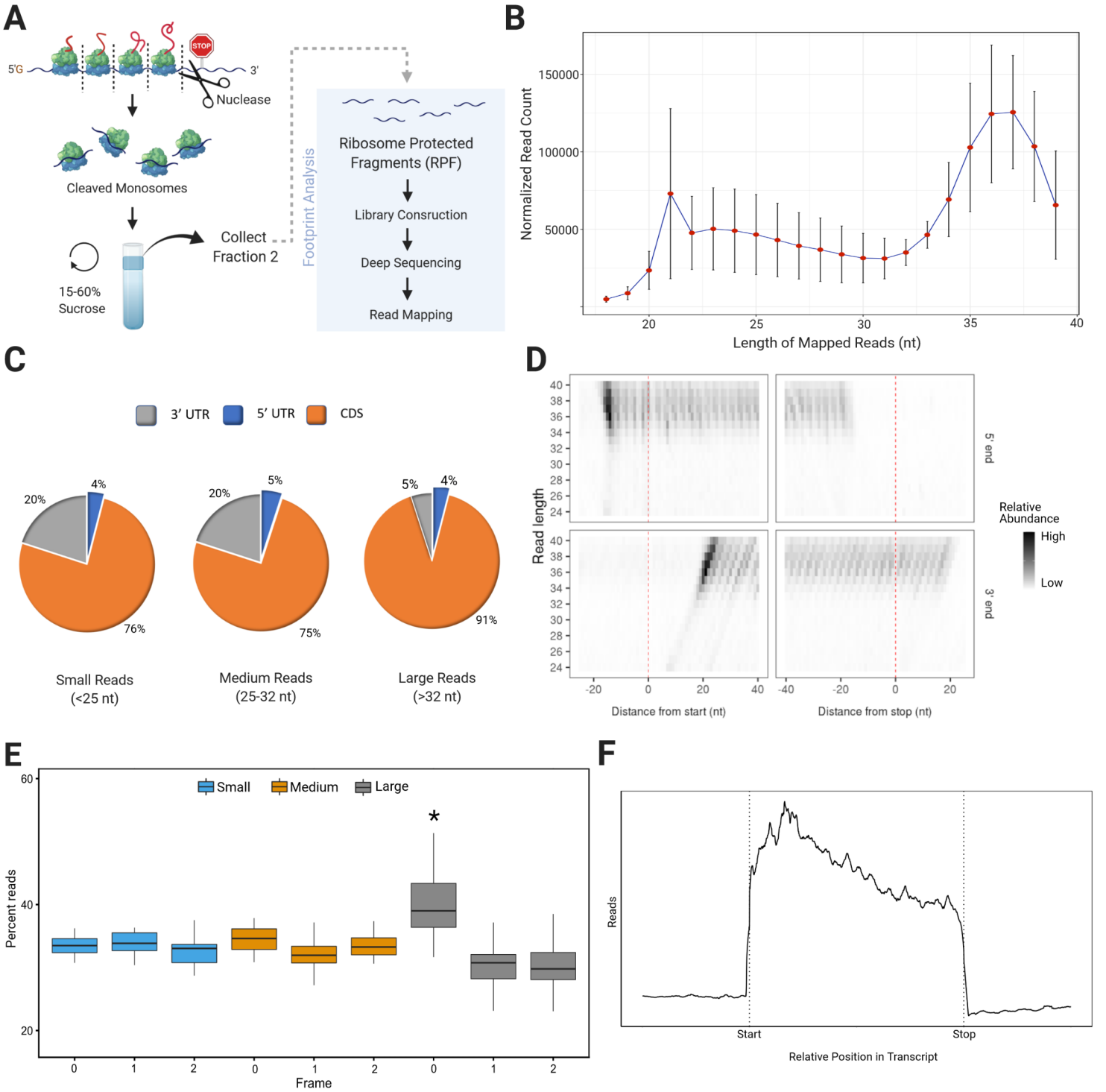


Figure 3. Ribosome footprinting of the pellet fraction
(A) Diagram summarizing Footprinting procedure. (B) Size distribution of normalized footprint reads, n=5 biological replicates, SEM. (C) Read coverage of different sized footprints (small <25 nt, medium 25-32 nt, large >32 nt) to 3'UTR, 5'UTR and CDS. (D) The number of read extremities (shading) for each read length (Y-axis) based on distance from start (left) and stop (right) with the beginning of the read (5') on top and the end of the read (3') on bottom. Data are shown for one biological replicate, but results are similar for all replicates. (E) Periodicity statistics indicate that long reads (33-40) in frame 0 have significantly more periodicity than frame 1 and frame 2 for long reads, or any frame for short (21-24) and medium reads (25-32) (ANOVA $F(261,8) = 13.9$, $p<0.001$; Post-Hoc Tukey HSD test, Long reads in 0 frame*, $p<0.001$ against long reads in frame 1 and frame 2 and other reads in frame 0) SD, n=39 long, 40 medium, 11 short (N is based on each read length in each biological replicate; not all read lengths are present in all biological replicates); Errors are SD. (F). Distribution of large reads with the CDS of all transcripts normalized to the same length shows that reads are biased to the first half of the transcript. Data are shown for one biological replicate, but results are similar for all replicates.

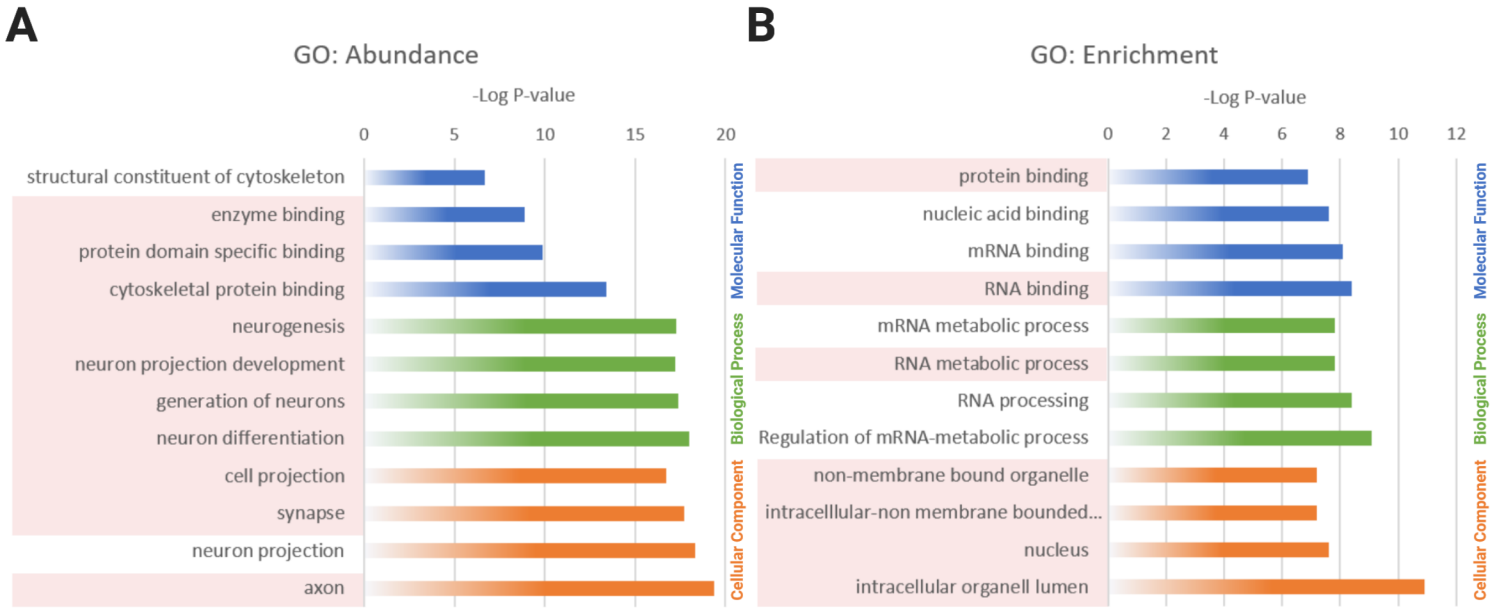


Figure 4. GO Analysis of footprint mRNAs
(A-B) Gene Ontology (GO) terms of selected comparisons for most enriched (A) and most abundant (B) mRNAs found in reads based on -Log P-values. Terms highlighted in red represent terms involving RNA binding for enrichment and cytoskeletal related terms for abundance. Complete table with all significant findings in Supp Table 2.

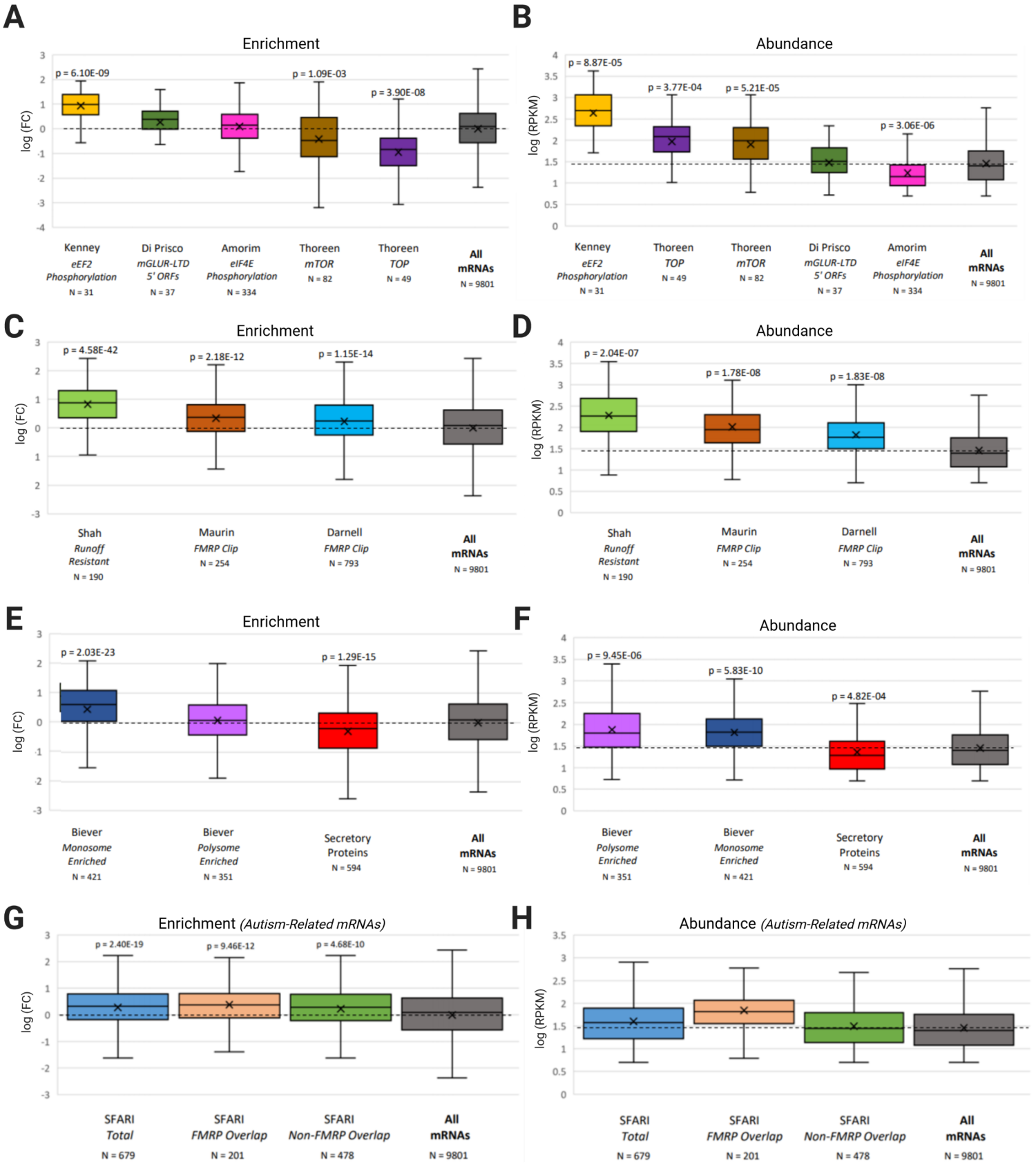
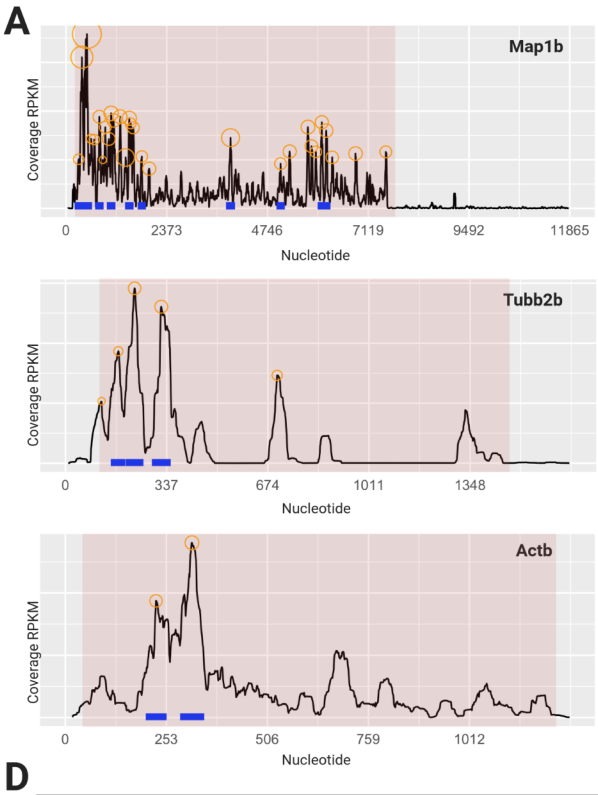
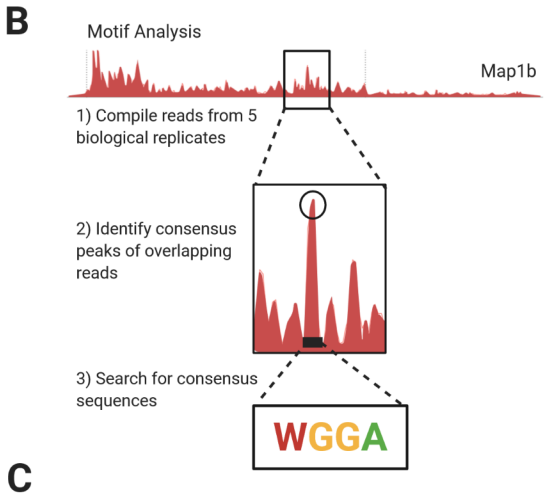


Figure 5. Correlation analysis of mRNAs abundant and enriched in ribosome footprints
(A) Enrichment and (B) Abundance comparison of footprint reads to mRNAs regulated by translation elongation (30) eIF4E phosphorylation (31) mTOR (32) TOP mRNAs (32) and mRNAs upregulated by mGluR with upstream open reading frames (33). (C) Enrichment and (D) Abundance comparison of runoff-resistant mRNAs (19) and mRNAs that are CLIPped by FMRP (17, 34). (E) Enrichment and (F) Abundance comparison of mRNAs translated preferentially by monosomal and polysomal mRNAs in the neuropil (35) and secretory mRNAs (secretory proteins with reviewed annotation from UNIPROT), compared to all mRNAs. (G-H) Enrichment and Abundance comparison of autism-related mRNAs from the SFARI database (syndromic and levels 1-3). The total SFARI group was also divided into ones that are also in the FMRP CLIP group (17,34). and ones that are not. For all groups there was a cut-off of 5 RPKM to avoid mRNAs not expressed in the nervous system. P values from comparison to all mRNAs (Students t test with Bonferroni correction for multiple tests (n=14 for all comparisons in figure). Only Significant P values (p<0.01 after correction are shown). The N for each comparison group is shown under the group.





C

Rank	Motif	P-value	% Targets	% Background
1	CAUGGACU	1e-67	52.60	21.74
2	CAUCGAGAACAU	1e-49	77.75	49.97
3	UAUGUAGA	1e-34	29.48	11.63

D

Rank	1	2	3
Motif	CAUGGACU	CAUCGAGAACAU	UAUGUAGA
Matching RBP Consensus Sequence	WGGA (FMR1-1) RGACH (m6A) BV TGGA (FMR1-2) A NGGACA (FMR1-3)	G VAGAW (FMRI)	G v AGAw (FMRI)

Code	Residue
R	Purine (A & G)
W	Weak (A & T)
V	Not T
H	Not G
B	Not A
N	Unknown

E

Top 10 hits from RBP list

RNA Binding Protein	Interaction Motif	Number of Consensus Peaks with Motif	Total Occurrence of Motif	Frequency of Motifs per Consensus Peak			
				n=1	n=2	n=3	n=4+
FMR1-1	WGGA	415	615	252	133	23	7
FMR1-2	BVTGGA	408	585	269	109	25	2
SRSF1	AGGAY	371	500	264	87	18	2
m6A	RGACH	366	472	272	83	10	1
FMRI	GVAGAW	349	489	238	89	16	5
FMR1-3	ANGGACA	334	416	255	76	3	0
LIN28A	GGAGWA	322	419	243	64	13	1
RBM5	GAAGGAG	303	465	190	80	19	12
FMR1-4	GACAAG	299	379	228	63	7	1
FXR2	GACYYY	296	372	229	59	7	1

F

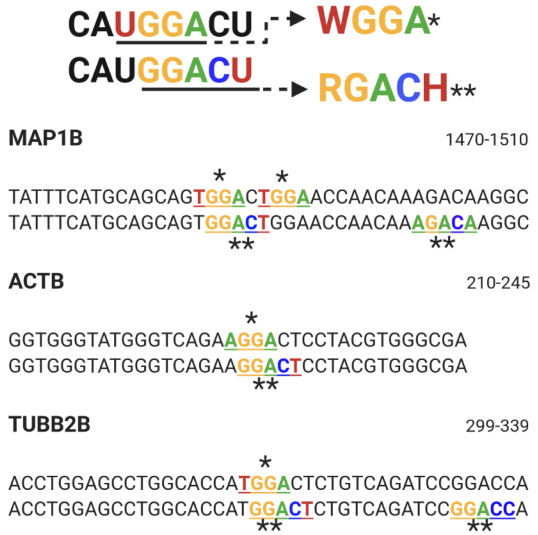
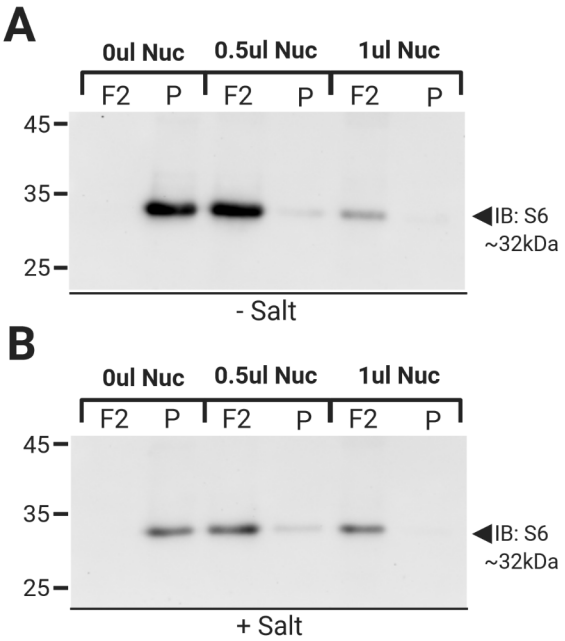
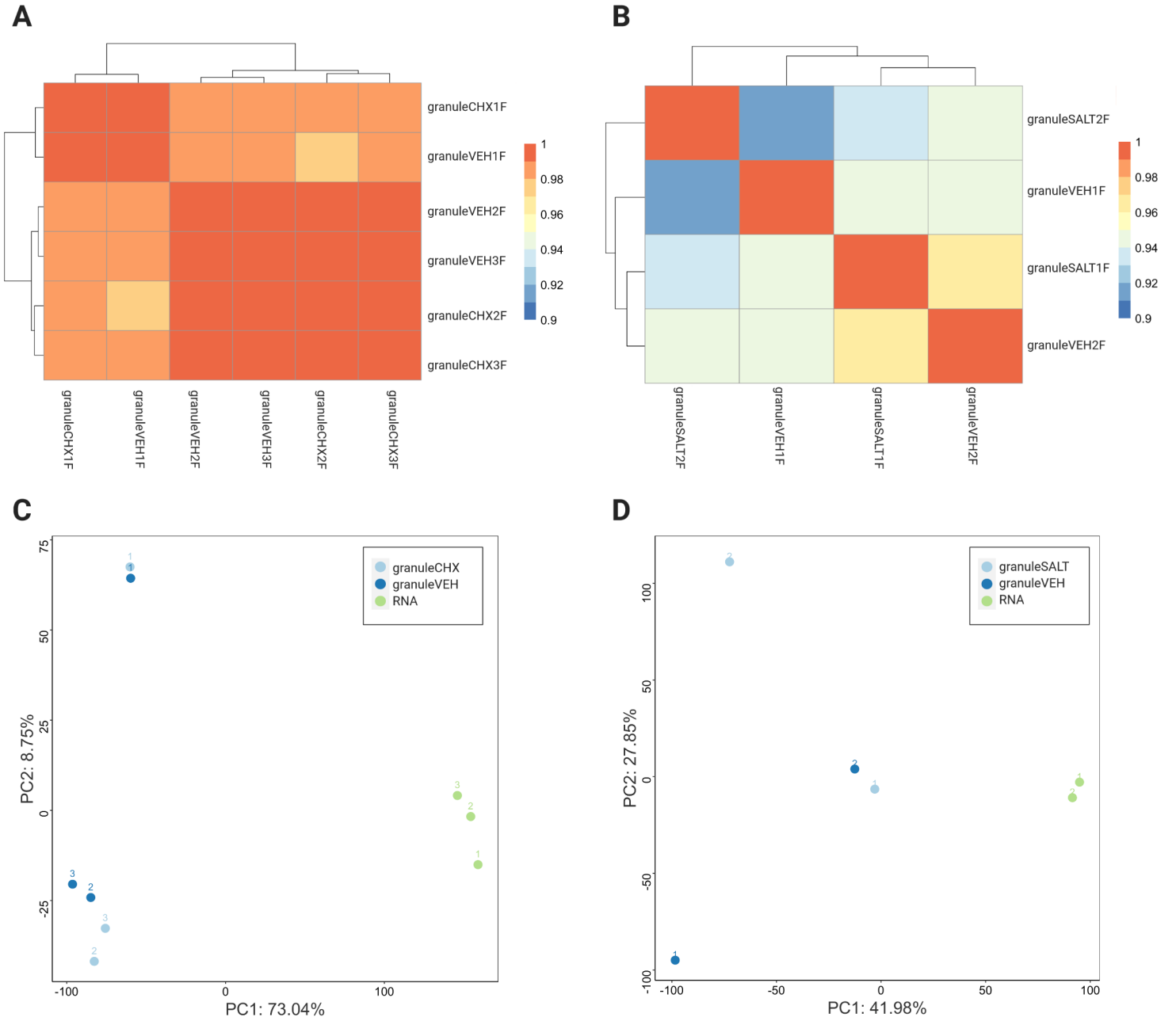


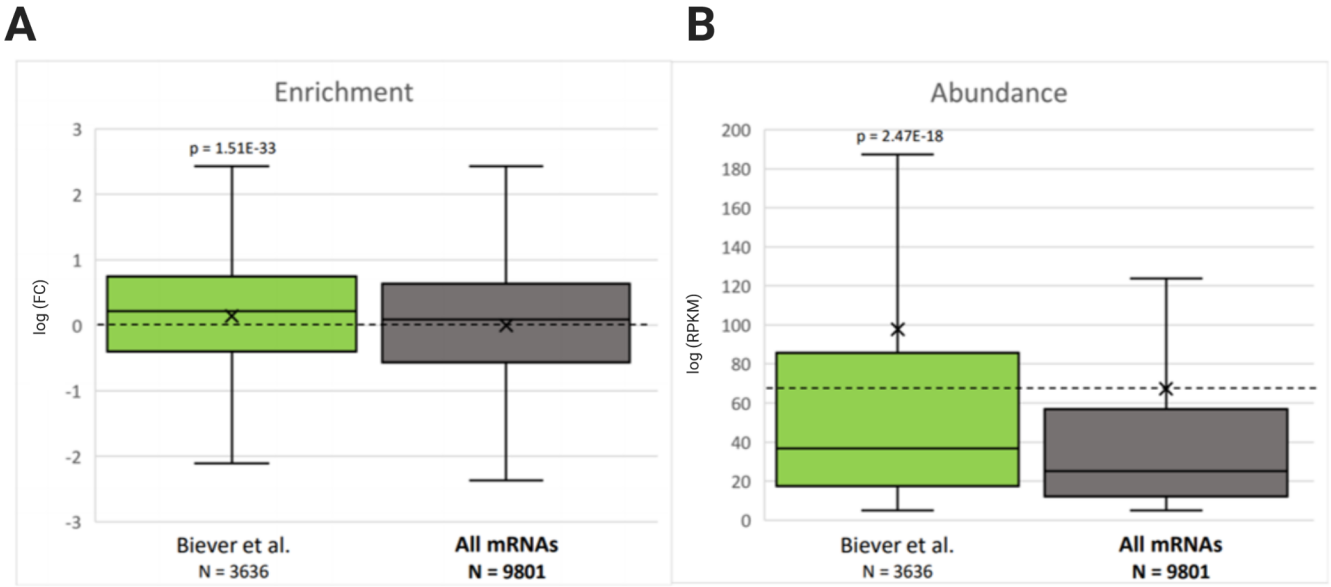
Figure 6. Sequences underlying ribosome protected fragments are enriched in sequences matching FMRP CliPs. (A) mRNA profiles of Map1b, Beta-actin and Tubulin 2b showing reproducible consensus peaks in the CDS; circles represent consensus peaks of footprint reads mapping to the same sequence, blue lines represent reproducible consensus peaks across biological replicates, red shading represents CDS. All replicates are shown in Supp. Fig 5. (B) Diagram summarizing how motif analysis is done. (C) Results from the HOMER program showing the only 3 consensus sequences above the cut-off provided by HOMER. (D) Homer identified motifs overlapped with matching interaction motifs for RBPs listed in brackets. Residues that do not match are given in smaller font. RIGHT: Code key for residue annotation. (E) Table of top 10 RBPs with RNA interaction motifs present in consensus peaks. The number of peaks with multiple hits are also shown as Frequency of Motifs per Consensus Peak; n=x represents the number of motifs per consensus peak. All RBPs motifs examined are shown in in Supp. Table 4. (F) Top ranked consensus sequence from HOMER showing overlapping sites for interaction motifs WGGA and RGACH, and their corresponding residues on a single peak from Map1b, Beta-actin and Tubulin 2b. Both sequences given for each protein are identical, but they have been annotated to show clusters of motifs that map to the same consensus sequence. Numbers on the right indicate the location of the sequence in each mRNA that correspond to a consensus peak (blue) on the mRNA profiles shown in (A).



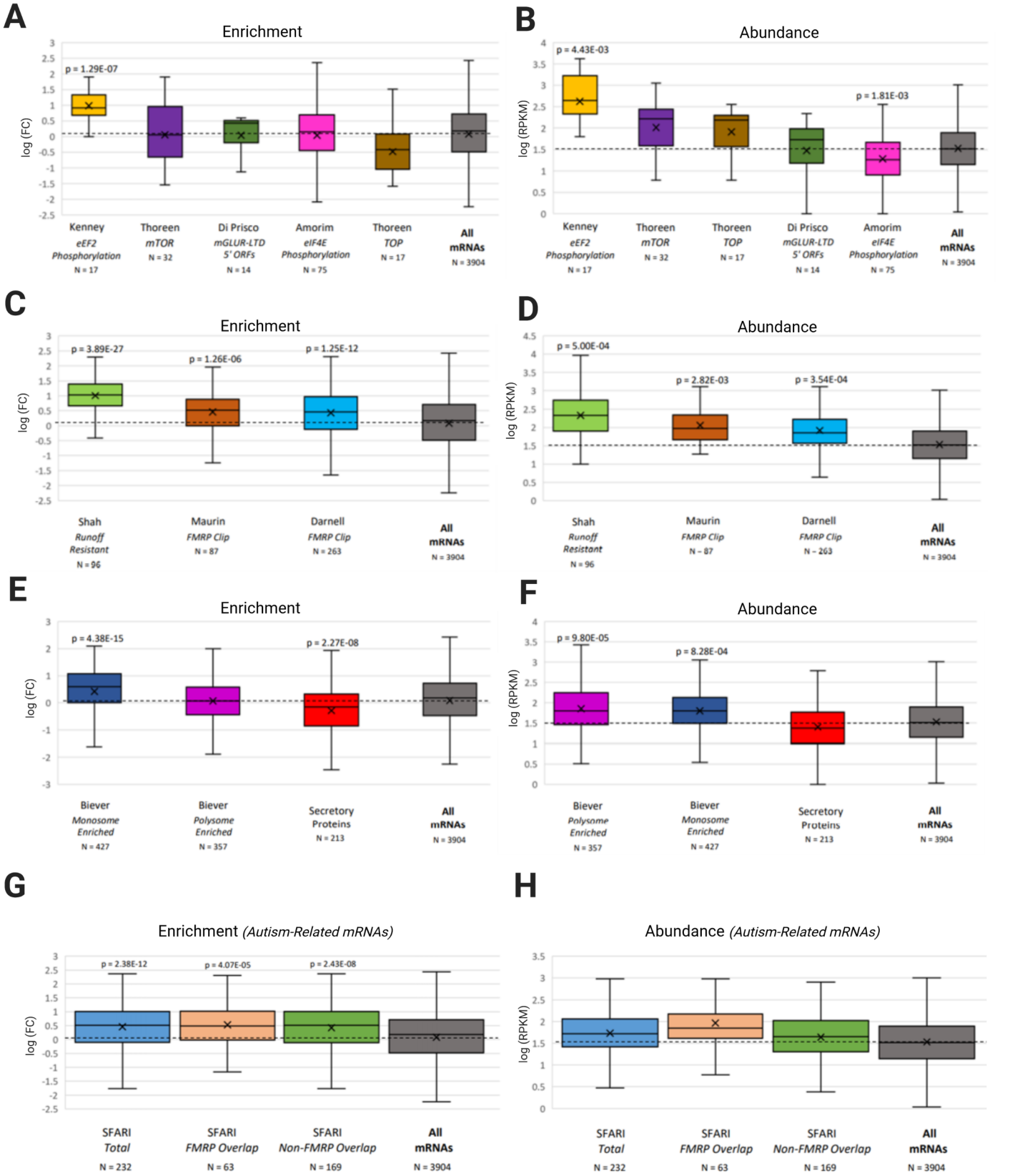
Supplemental Figure 1. Complete digestion of pellet fraction.
Western Blot analysis for S6 ribosomal protein of purified fractions collected after first treating the pellet from the first sucrose gradient with 0ul (A), 0.5ul (B) and 1ul (C) RNase I, with or without pre-treatment of 400 mM NaCl (-SALT: TOP; +SALT: BOTTOM), followed by a second sucrose gradient (see Methods). Only fraction 2 was loaded in this experiment. In this example, almost complete digestion even with lower concentrations of nuclease and no salt was observed. Increased nuclease digestion appeared to be related to different batches of nuclease and different amounts of time nuclease was stored, but this was not systematically analyzed.



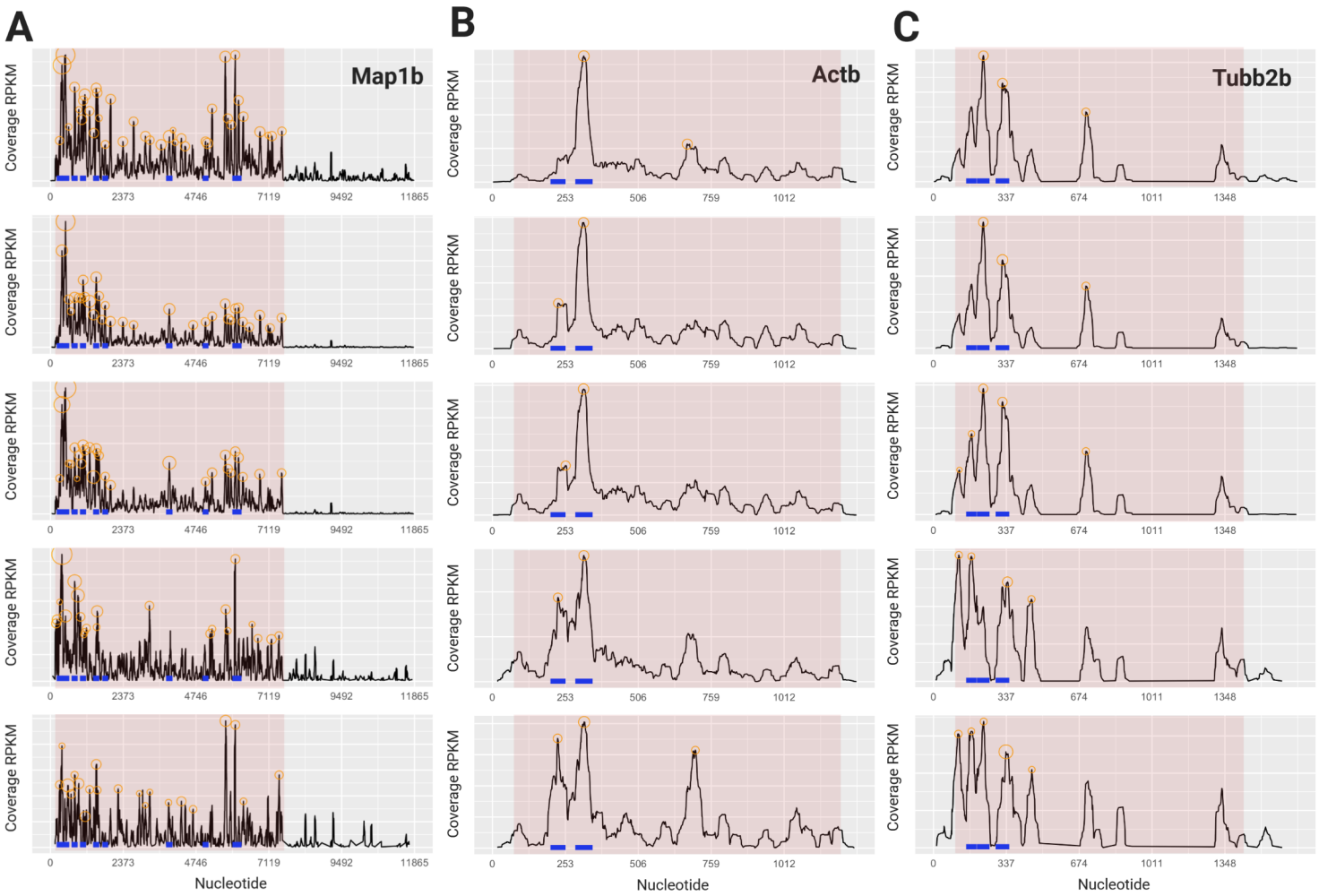
Supplemental Figure 2. No effect of Salt or Chx treatment on RNA reads
(A-B) Heat map for the comparison of biological replicates accomplished in the presence or absence of cycloheximide (Chx) (A) or for the comparison of biological replicates accomplished in the presence or absence of salt (B). Warmer colors indicating higher correlation between groups. Differences between biological samples was equal to or higher than the differences seen with treatment (C-D) Principal component analysis for the comparison of biological replicates accomplished in the presence or absence of cycloheximide (Chx) including RNA-SEQ of starting material (C) or for the comparison of biological replicates accomplished in the presence or absence of salt including RNA-SEQ of starting material (D). Differences between biological samples was equal to or higher than the differences seen with treatment. The footprint reads were clustered separately from the RNA-SEQ of total mRNA. All calculations were done with the Limma program (28).



Supplemental Figure 3. Enrichment and Abundance compared to transported mRNAs
(A) Enrichment and (B) Abundance comparison of footprint reads to mRNAs classified as transported to distal sites as determined (35). For all groups there was a cut-off of 5 RPKM to avoid mRNAs not expressed in the nervous system. P values from comparison to all mRNAs (Students t test with Bonferroni correction for multiple tests (n=2 for all comparisons in figure). Only Significant P values ($p<0.01$ after correction are shown). The N for each comparison group is shown under the group.



Supplemental Figure 4. Comparisons only using neuronal transported mRNAs
(A) Enrichment and (B) Abundance comparison of footprint reads to mRNAs regulated by translation elongation (30) eIF4E phosphorylation (31) mTOR (32) TOP mRNAs (32) and mRNAs upregulated by mGluR with upstream open reading frames (33). (C) Enrichment and (D) Abundance comparison of runoff-resistant mRNAs (19) and mRNAs that are CLIPped by FMRP (17, 34). (E) Enrichment and (F) Abundance comparison of mRNAs translated preferentially by monosomal and polysomal mRNAs in the neuropil (35) and secretory mRNAs (secretory proteins with reviewed annotation from UNIPROT), compared to all mRNAs. (G-H) Enrichment and Abundance comparison of autism-related mRNAs from the SFARI database (syndromic and levels 1-3). The total SFARI group was also divided into ones that are also in the FMRP CLIP group (17,34). and ones that are not. For all groups there was a cut-off of 1RPKM to avoid mRNAs not expressed at this point in development. P values from comparison to all mRNAs (Students t test with Bonferroni correction for multiple tests (n=14 for all comparisons in figure). Only Significant P values (p<0.01 after correction are shown). The N for each comparison group is shown under the group.



Supplemental Figure 5. Read Maps of all 5 biological replicates.
(A-C) Footprint read maps for Map1b (A), Beta-actin (B) and Tubulin 2b (C) for 5 biological replicates. Circles represent consensus peaks of footprint reads mapping to the same sequence, blue lines represent reproducible consensus sequences across biological replicates. Criteria for identifying consensus sequences (blue lines) is that at least 3 of 5 biological replicates have a consensus peak at this site. Red shading represents coding sequence (CDS).