UNIVERSITY OF CALIFORNIA SAN DIEGO

Large-scale activity assignment of RNA-interacting proteins identifies a functional antagonist to fragile X mental retardation protein

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy

in

Bioengineering

by

En-Ching Luo

Committee in charge:

Professor Sheng Zhong, Chair
Professor Eugene Yeo, Co-Chair
Professor Jens Lykke-Andersen
Professor Amy Pasquinelli
Professor Shankar Subramaniam

2020

This dissertation of En-Ching Luo is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____
Co-Chair

_____
Chair

University of California San Diego

2020

DEDICATION

For my family and friends who have supported me in this journey.

En-Ching Luo

EPIGRAPH

Where can wisdom be found?
Where does understanding dwell?
No mortal comprehends its worth;
it cannot be found in the land of the living.
The deep says, "It is not in me";
the sea says, "It is not with me."
It cannot be bought with the finest gold,
nor can its price be weighed out in silver.
It cannot be bought with the gold of Ophir,
with precious onyx or lapis lazuli.
Neither gold nor crystal can compare with it,
nor can it be had for jewels of gold.
Coral and jasper are not worthy of mention;
the price of wisdom is beyond rubies.
The topaz of Cush cannot compare with it;
it cannot be bought with pure gold.
Where then does wisdom come from?
Where does understanding dwell?
It is hidden from the eyes of every living thing,
concealed even from the birds in the sky.
Destruction and Death say,
"Only a rumor of it has reached our ears."
God understands the way to it
and he alone knows where it dwells,
for he views the ends of the earth
and sees everything under the heavens.
When he established the force of the wind
and measured out the waters,
when he made a decree for the rain
and a path for the thunderstorm,
then he looked at wisdom and appraised it;
he confirmed it and tested it.
And he said to the human race,
"The fear of the Lord—that is wisdom,
and to shun evil is understanding."
~ Job 28:12-28 ~

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SUPPLMENTAL FILES

- Extended Data Table 1 is available as supplemental tables online
- Extended Data Table 2 is available as supplemental tables online

ACKNOWLEDGEMENTS

I would like to acknowledge Dr. Gene Yeo for his support and guidance, leading me to explore the beauty of science, giving me opportunities to collaborate with great groups and scientists and teaching me how to be a good scientist. I would like to acknowledge Dr. Stefan Aigner for being a patient mentor, keeping me on track, and giving key criticisms of my work, also sharing his precious advice with me. I couldn't have done any of this without him. I would like to acknowledge my committee members: Dr. Sheng Zhong, Dr. Amy Pasquinelli, Dr. Jens Lykke-Andersen, and Dr. Shankar Subramaniam for all of their intellectual advice and mentorship.

Thank you to Dr. Yan Song, Dr. Shengnan Joy Xiang, and Dr. Chun-yuan Chen for sharing your personal experiences with me and giving me the advice to solve problems. Thank you to Dr. Alex Chaim, Dr. Frederick Tan, and Dr. Meredith Corley for encouraging me when I face challenges. Thank you to Dr. Emily Wheeler, Dr. Mark Perelis, Archana Shankar, Anthony Vu, Ryan Marina, Florian Krach, Steven Blue for spreading happiness around me. Spending time with you guys was the "wonderful experiences" in my grad school. Thank you to everyone in the Yeo lab for your company. This journey was incredible and unforgettable.

Thank you to my dear roommates, Hsiu-I, Tsan-wen, and Peng-Hao, and my friends, Min-Ru, Yi-Fan, Yijia, Jeffrey, Jerry, Michelle, Melody, Carol, Jessie, Thomas for being as my family in the US, supporting me and creating unforgettable memories together. Your company was what got me through grad school.

Thank you to my family, especially my parents for encouraging me to follow my heart, and start adventures overseas. There are no words sufficient to convey my gratitude for how much you have supported me in the last thirty years. Thank you to my mom for showing me

how to be a great researcher, and thank you to my Dad for teaching me how to face the challenges and make decisions. Thank you to my brother for being my best friend. Thank you to Ted for understanding my needs and always be there when I need you.

Chapter 2, in full, was published in Nature structural & molecular biology in 2020 (**Luo EC**, Nathanson JL, Tan FE, Schwartz JL, Schmok JC, Shankar A, Markmiller S, YeeBA, Sathe S, Pratt GA, Scaletta DB, Ha Y, Hill DE, Aigner S, Yeo GW. "Large-scale tethered function assays identify factors that regulate mRNA stability and translation"). Thank you to Stefan Aigner and Gene Yeo for their contribution to this work. The dissertation author is the primary author of this manuscript.

Chapter 3 is currently preparing as a manuscript by **Luo EC**, Lee S, Kosmaczewski SG, Shankar A, Barrett LE, Gao FB, Aigner S, Yeo GW. "UBAP2L is a functional antagonist to fragile X mental retardation protein in fragile X syndrome". The dissertation author is the primary author of this manuscript.

VITA

2011 Bachelor of Science, Life Science, National Tsing Hua University, Taiwan

2014 Master of Science, Physiology, National Taiwan University, Taiwan

2020 Doctor of Philosophy, Bioengineering, University of California San Diego, USA

PUBLICATIONS

Luo EC, Nathanson JL, Tan FE, Schwartz JL, Schmok JC, Shankar A, Markmiller S, YeeBA, Sathe S, Pratt GA, Scaletta DB, Ha Y, Hill DE, Aigner S, Yeo GW. Large-scale activity assignment of RNA-interacting proteins by tethered function assays. *Nature structural & molecular biology* 2020

Coyne AN, Zaepfel BL, Hayes L, Fitchman B, Salzberg Y, **Luo EC**, Bowen K, Trost H, Aigner S, Rigo F, Yeo GW, Harel A, Svendsen CN, Sareen D, Rothstein JD. G4C2 repeat RNA initiates a POM121 mediated reduction in specific nucleoporins in C9orf72 ALS/FTD *Neuron 2020.*

Di Stefano B*, **Luo EC***, Haggerty C, Aigner S, Charlton J, Brumbaugh J, Ji F, Rabano Jimenez I, Clowers KJ, Huebner AJ, Clement K, Lipchina I, de Kort MAC, Anselmo A, Pulice J, Gerli MFM, Gu H, Gygi SP, Sadreyev RI, Meissner A, Yeo GW, Hochedlinger K. The RNA Helicase DDX6 Controls Cellular Plasticity by Modulating P-Body Homeostasis. *Cell Stem Cell*. 2019 (* shared co-first authors)

Markmiller S, Soltanieh S, Server KL, Mak R, Jin W, Fang MY, **Luo EC**, Krach F, Yang D, Sen A, Fulzele A, Wozniak JM, Gonzalez DJ, Kankel MW, Gao FB, Bennett EJ, Lecuyer E, Yeo GW. Context-Dependent and Disease-Specific Diversity in Protein Interactions within Stress Granules. *Cell.* 2018

Luo EC, Chang YC, Sher YP, Huang WY, Chuang LL, Chiu YC, Tsai MH, Chuang EY, Lai LC. MicroRNA-769-3p down-regulates NDRG1 and enhances apoptosis in MCF-7 cells during reoxygenation. *Sci Rep*. 2014

Liu YJ, Lin YF, Chen YF, **Luo EC**, Sher YP, Tsai MH, Chuang EY, Lai LC. MicroRNA-449a enhances radiosensitivity in CL1-0 lung adenocarcinoma cells. *PLoS One*. 2013

IN REVISION

Van Nostrand JL, Hellberg K, **Luo EC**, Van Nostrand EL, Dayn A, Yu J, Shokhirev M, Dayn Y, Shaw RJ. AMPK regulation of Raptor and TSC2 mediate Metformin effects on transcriptional control of anabolism and inflammation. *Submitted to Genes & Development* 2020

Kosmaczewski S, **Luo EC**, Ghosh S, Beccard A, Mazzucato P, Bara AM, Messana A, Lam D, Aigner S, Yeo GW. Molecular convergence between Fragile X syndrome and Down syndrome. *Submitted to Neuron 2020*

Kameda-Smith M, Zhu H, **Luo EC**, Venugopal C, Xella A, Yee B, Xing S, Tan F, Fox R, Brown K, Bakhshinyan D, Adile A, Chokshi C, Gwynn W, Burns I, Subapanditha M, Picard D, Moffat J, Fleming A, Hope K, Provias J, Remke M, Lu Y, Reya T, Reimand J, Wechsler-Reya R,Yeo GW, Singh S. Integrative multi-omic analyses of MSI1 function in brain cancer reveal context-specific downstream targets for drug discovery. *Submitted Nature Genetics, 2020*

Kejiou NS, Ilan L, Aigner S, **Luo EC**, Rabano Ines, Rajakulendran N, Najafabadi HS, Aigner S,Yeo GW, Palazzo AF. Pyruvate Kinase M Links Glucose Availability to Protein Synthesis. *Submitted to BioRxiv* 2019

ABSTRACT OF THE DISSERTATION

Large-scale activity assignment of RNA-interacting proteins identifies a functional antagonist to fragile X mental retardation protein

by

En-Ching Luo

Doctor of Philosophy in Bioengineering

University of California San Diego, 2020

Professor Sheng Zhong, Chair
Professor Eugene Yeo, Co-Chair

RNA binding proteins (RBPs) impact cellular protein levels by regulating messenger RNA (mRNA) levels. Assignment of function to hundreds of emerging, uncharacterized RBPs is a critical bottleneck to a complete understanding of gene expression control. Here, large-scale tethering of nearly a thousand RBPs discovers 50 RBPs that affect reporter RNA turnover and

translation. Enhanced UV crosslinking and immunoprecipitation (eCLIP) identify hundreds of endogenous mRNA targets affected by manipulating levels of more than a dozen candidate RBPs. Among these candidates, we characterize the ubiquitin-associated protein 2-like (UBAP2L) gene. Polysome profiling assays indicate that UBAP2L enhances translation of target mRNAs, likely due to ribosome interactions as supported by eCLIP data. UBAP2L can also be found in complex with fragile X mental retardation protein FMRP, and 52% of UBAP2L mRNA targets are also FMRP targets. UBAP2L depletion in a cortical neuronal model of Fragile X Syndrome (FXS) corrects molecular, cellular and electrophysiological defects relevant to autism spectrum disorder. Reduction of the Drosophila ortholog of UBAP2L in a FXS fly model rescues the neurodevelopmental defects due to loss of FMRP. Our efficient and scalable method identifies proteins involved in RNA metabolism and detailed studies of UBAP2L provides a new therapeutic strategy into human disease.

Chapter 1: Introduction

1.1 Discovery and molecular mechanistic dissection of RNA binding proteins that control mRNA stability and translation

RNA binding proteins (RBPs) involve in RNA processing by regulating RNA Splicing, transportation, turnover, translation, and modification. In the cytoplasm, RBPs control cellular protein levels in three major steps. First, RBPs regulate mRNA steady-state levels and turnover rates. Second, RBPs transport mRNA to subcellular localization. Third, RBPs regulate the translation efficiency. The importance of regulating protein expression levels at these three steps is known, but the mechanisms were far less well studied. Therefore, the current studies in this filed are to understand how RBPs regulate gene expression and control cellular protein levels.

Currently, ~2800 RBPs have been identified in human cells (1-5). However, only a handful of mRNA binding proteins have been shown to directly interact with poly(A)+ mRNA. Studies have shown RBPs interact directly with the 5' and 3' UTR of mRNA to protect mRNA from, or promote mRNA decay by, exonuclease degradation. RBPs bind on the 5' and 3' UTR also affect translation initiation and elongation or termination. Only a few of these RBPs have shown to facilitate the transportation of mRNA in different subcellular localization. For example, FMRP shuttles mRNA into axon terminals of neurons for localized translation. The

misfunction or mutations in these RBPs cause a number of neuron degeneration diseases such as Fragile x syndrome.

Several major questions have now surfaced as urgent to the field: Which RBPs affect mRNA turnover and translation? What are their RNA substrates and the genetic circuits controlled by them? What extent are mRNAs localized to specific subcellular compartments and which RBPs control this distribution? To answer these questions, I proposed to conduct a reporter screen of a large, representative set of ~1000 bona fide RBPs for their ability to affect mRNA stability, translation, and identify cellular mRNA targets and mode of action for a subset of ~15 RBPs. To accomplish these goals, I pursued the following Specific Aims:

- Identify RBPs with novel roles in modulating mRNA stability
- Identify endogenous RNA subtracts of candidates RBPs, and confirm RBP-dependent mRNA turnover and translation

Lastly, I picked up a novel translation regulator UBAP2L and investigate its roles. To characterize the functions of UBAP2L, I analyzed the domain structures of UBAP2L, and confirm the interactions of RNAs and proteins with UBAP2L. Finally, I proposed the potential mechanism of translation regulation of UBAP2L.

1.2 Disease-modifying therapeutic strategies for FXS

The translation repressor FMRP protein is encoded in FMR1 locus. The Silence of the FMR1 gene causes fragile x syndrome (FXS) which is an autism spectrum disorder and the

most common genetic form of intellectual disability. In FXS neurons, the FMR1 gene has repeat expansion insertion in the promoter and 5'UTR of FMR1 gene. Methylation of the repeat expansions silence the FMR1 locus and deplete FMR1 expression.

Loss of FMRP results in exaggerated protein synthesis at the synapses, cortical hyperexcitability, and neuron developmental defects. Despite the FMRP function has been studied for a decade, and the molecular etiology has been understood well, no FDA approved drug exists for the treatment of FXS. Most of the approaches aimed at de-silence FMR1 gene; however, several approaches failed due to lack of effectiveness (e.g., mavoglurant, Novartis; basimglurant, Roche). Therefore, it is crucial to identify alternative methods to pursue novel therapeutic strategies in FXS.

I identified the UBA-domain containing ubiquitin-associated protein 2-like (UBAP2L) protein as an FMRP binding partner and translational enhancer. To pursue a novel therapeutic strategy in FXS, I proposed to rescue the FXS defects by reduction or depletion of translation enhancer UBAP2L. To find out the possibility of the new strategy, several major questions have to answer.

- Confirm the interaction between UBAP2L and FMRP
- Identify the mRNA shared targets for UBAP2L and FMRP
- Examine whether the depletion of UBAP2L restores translation level in FXS neuron
- Examine whether the inhibition of UBAP2L reduce cortical hyperexcitability
- Examine whether the silence of UBAP2L rescue neuron development defect.

1.3 Outline of the dissertation

I will begin this thesis with a description of the tethered screen that identifies novel candidate RBPs which regulate RNA stability and translation (Chapter 2). The results of the screen reveal 50 candidates potentially regulate RNA turnover and translation. I confirmed the screen by modulating the expression of the candidate RBPs and identify a novel translation enhancer UBAP2L. Next, I will show a potential therapeutic strategy in FXS (Chapter 3) This is the second half of my thesis work that is aimed at finding the applications of UBAP2L. I proposed to rescue neuroanatomical deficits in FXS by the reduction of UBAP2L. Each section contains its own detailed introduction and conclusion on the specific subject matter.

Chapter 2: Large-scale tethered function assays identify factors that regulate mRNA stability and translation

En-Ching Luo[1,2,3], Jason L. Nathanson[1,2,3], Frederick E. Tan[1,2,3], Joshua L. Schwartz[1,2,3], Jonathan C. Schmok[1,2,3], Archana Shankar[1,2,3], Sebastian Markmiller[1,2,3], Brian A. Yee[1,2,3], Shashank Sathe[1,2,3], Gabriel A. Pratt[1,2,3], Duy B. Scaletta [1,2,3], Yuanchi Ha[1,2,3], David E. Hill[4], Stefan Aigner[1,2,3], and Gene W. Yeo[1,2,3,*]

[1]Department of Cellular and Molecular Medicine, University of California San Diego, La Jolla, CA, USA

[2]Stem Cell Program, University of California San Diego, La Jolla, CA, USA

[3]Institute for Genomic Medicine, University of California San Diego, La Jolla, CA, USA

[4]Center for Cancer Systems Biology (CCSB), Dana-Farber Cancer Institute, Boston, MA

*Correspondence: geneyeo@ucsd.edu

Funding Source: This work was supported by grants from the NIH (R01HG004659, U19MH107367, R01NS103172 and U41HG009889) to G.W.Y. E.-C.L. was partly supported by a study-abroad graduate-student fellowship from the Taiwanese government. J.C.S. was partially supported by a Natural Sciences and Engineering Research Council of Canada Postgraduate Doctoral scholarship (PGS D - 532649 - 2019). Y.H. was supported by the UCSD

Frontiers of Innovation Scholars Program. G.A.P. was supported by a graduate fellowship from the National Science Foundation. F.E.T. was supported by a postdoctoral fellowship from the American Cancer Society (129547-PF-16-060-01-RMC). S.M. was supported by a postdoctoral fellowship from the Larry L. Hillblom Foundation (2014-A-027-FEL).

Key words: RNA, RNA binding protein, Translation, RNA stability, tethered screen, UBAP2L

2.1 Abstract

The molecular roles of more than a thousand uncharacterized and emerging RNA binding proteins (RBPs) remain unclear, highlighting a major bottleneck to our complete understanding of gene expression regulation. We develop a plasmid resource of 690 RBPs (771 open reading frames) that we subject to luciferase-based 3'UTR tethered function assays (TFAs) to prioritize candidate RBPs for roles in RNA stability or translation. Enhanced UV crosslinking and immunoprecipitation (eCLIP) identifies thousands of endogenous mRNA targets for these prioritized RBPs that respond to protein level changes reminiscent of their effects in TFAs. Among these, we identify that the ubiquitin-associated protein 2-like (UBAP2L) protein interacts with RNA via its RGG domain and cross-links to coding mRNA and ribosomal RNA. Fusion of UBAP2L to RNA-targeting CRISPR/Cas9 demonstrates programmable translational enhancement. Polysome profiling indicates that UBAP2L enhances translation of target mRNAs, particularly global regulators of translation such as initiation and

elongation factors. Our tethering survey rapidly assigns molecular activity of proteins, such as UBAP2L, to specific steps of mRNA metabolism.

2.2 Introduction

The fate of the transcriptome ultimately determines the status and health of the cell, and RNA-binding proteins (RBPs) control the post-transcriptional processing of these mRNA transcripts. Following 5'-end capping, splicing, 3'-end cleavage and polyadenylation, mammalian mRNAs are exported to the cytoplasm, where RBPs control their turnover, subcellular localization and the efficiency with which they are translated. Dysfunction of RBPs is linked to dozens of multisystemic diseases, cancer and neurological disorders (6-9). Despite their association with disease and although the importance of regulating gene expression at these cytoplasmic stages of the mRNA life cycle is well appreciated, only a small fraction of the over 2,000 RBPs identified thus far have known RNA targets and molecular roles (10-15).

Cross-linking and immunoprecipitation (CLIP)-based approaches have enabled detailed studies of individual RBPs through the transcriptome-wide identification of their binding sites (1-5). Molecular functions can then be inferred from integrative analyses of its bound transcripts and region-level binding preferences in conjunction with transcriptome-wide changes in splicing levels or ribosome association of mRNAs upon RBP depletion or overexpression (3, 16-21). While powerful, these strategies as a means to elucidate RBP molecular function are not easily scaled to thousands of RBPs.

As a complementary and orthogonal approach that is also more scalable and efficient, we turn to the well described tethered function assay (TFA) as a preliminary means to assign molecular functions to RBPs without requiring prior knowledge of their functional protein domains or natural RNA targets (22). Here, the RBP is fused to an exogenous RNA binding moiety, such as the bacteriophage MS2 coat protein (MCP) and co-expressed with a reporter mRNA harboring the cognate RNA structural motif, such as the MS2 genomic RNA stem-loop. The effect of the protein 'tethered' to the reporter mRNA is then evaluated. As TFAs have been crucial in elucidating the roles of RBPs in splicing, stability, transport and translation (23, 24), we reasoned that they are a pragmatic choice for assigning RBP function in a high-throughput format. We acknowledge that, similar to any high-throughput screen, it is not without limitations and caveats such as the dependence on reporter design and cell-line used in the screen, as will be discussed further below. Nevertheless, firstly, as we show in this work, positively-scoring candidates from TFA screens can be rapidly prioritized for detailed transcriptomic studies such as CLIP-seq and depletion or overexpression studies. Secondly, our plasmid resource serves as a starting point to explore differences in gene expression regulation with other reporters and cell-lines that will invariably be leveraged in future screens.

To accomplish our objective, we developed an open reading frame (ORF) library of 690 RBPs consisting of 771 isoforms, each fused to MCP. Using two sets of reporters that direct their recruitment to luciferase mRNA via an 3' untranslated region (3'UTR) as a practical convenience for inserting MS2 stem-loops (that do not affect translating ribosomes in the coding region), we report the first and largest-scale tethering screen to date to assign an effect on reporter RNA levels to known and predicted RBPs. We chose to utilize luciferase as a primary readout as the screen can be conducted in 96 or 384-well plate-format in a relatively

cost-effective manner. Our screen revealed hundreds of fusion RBPs that significantly affect luciferase as a readout in two reporter contexts. Of the 50 candidate RBPs that scored positively in both reporters with stringent thresholds, we subjected 14 to enhanced CLIP (eCLIP) analyses (1) to identify their transcriptome-wide endogenous RNA targets. Perturbation of the levels of these candidates revealed that regulation of their natural substrates was largely consistent with our reporter findings. The generality of their function on both reporter and endogenous targets strongly highlights the validity of utilizing TFAs as efficient means to discover molecular roles for RBPs.

2.3 Results

2.3.1 Generation of resource of RBP open-reading frames fused to MS2 coat protein and tethered function assays

We assembled a collection of RBP expression constructs using in-house bioinformatics tools to extract genes annotated to contain RNA-binding domains as predicted by PFAM(25) and PRINTS (26). We extended this set with mRNA-bound putative RBPs identified experimentally in two different studies which used UV-cross-linking and oligo(dT) capture followed by mass spectrometry (7, 8). We initially acquired 881 open-reading frames (ORFs) of known and predicted RBPs from both commercial sources and through in-house cloning efforts and sub-cloned them into two constructs using Gateway-mediated cloning: one that directs expression of the RBPs as fusion proteins with the V5 epitope tag C-terminally appended, and one with an additional bacteriophage MS2 coat protein (MCP) domain at the C-

terminus. To ensure that these ORFs express readily as full-length proteins, we expressed these 881 ORFs in HeLa cells and assessed for the presence of the C-terminal V5 epitope using western dot blot analysis (**Figure 2.1A**). Therefore, we obtained our final library of 690 unique RBPs from 771 expressible ORFs (**Figure 2.1B; Extended Data Table 1**). Overall, ~40% of the RBPs in our final collection contain known canonical RNA-binding motifs, while the remainder may associate with RNA through other interaction domains or binding modes (**Figure 2.1C**), similar to previous reports (7, 8). Highlighting the need for assessing the roles of RBPs in RNA metabolism, Gene Ontology (GO) analysis showed that ~60% of the RBPs in our collection have no known RNA-related functions (**Figure 2.1D**). Thus, we have assembled a resource of representative 'tethered' and 'untethered' RBP expression libraries comprising the majority of all predicted and/or experimentally identified RNA-binding proteins (10-15).

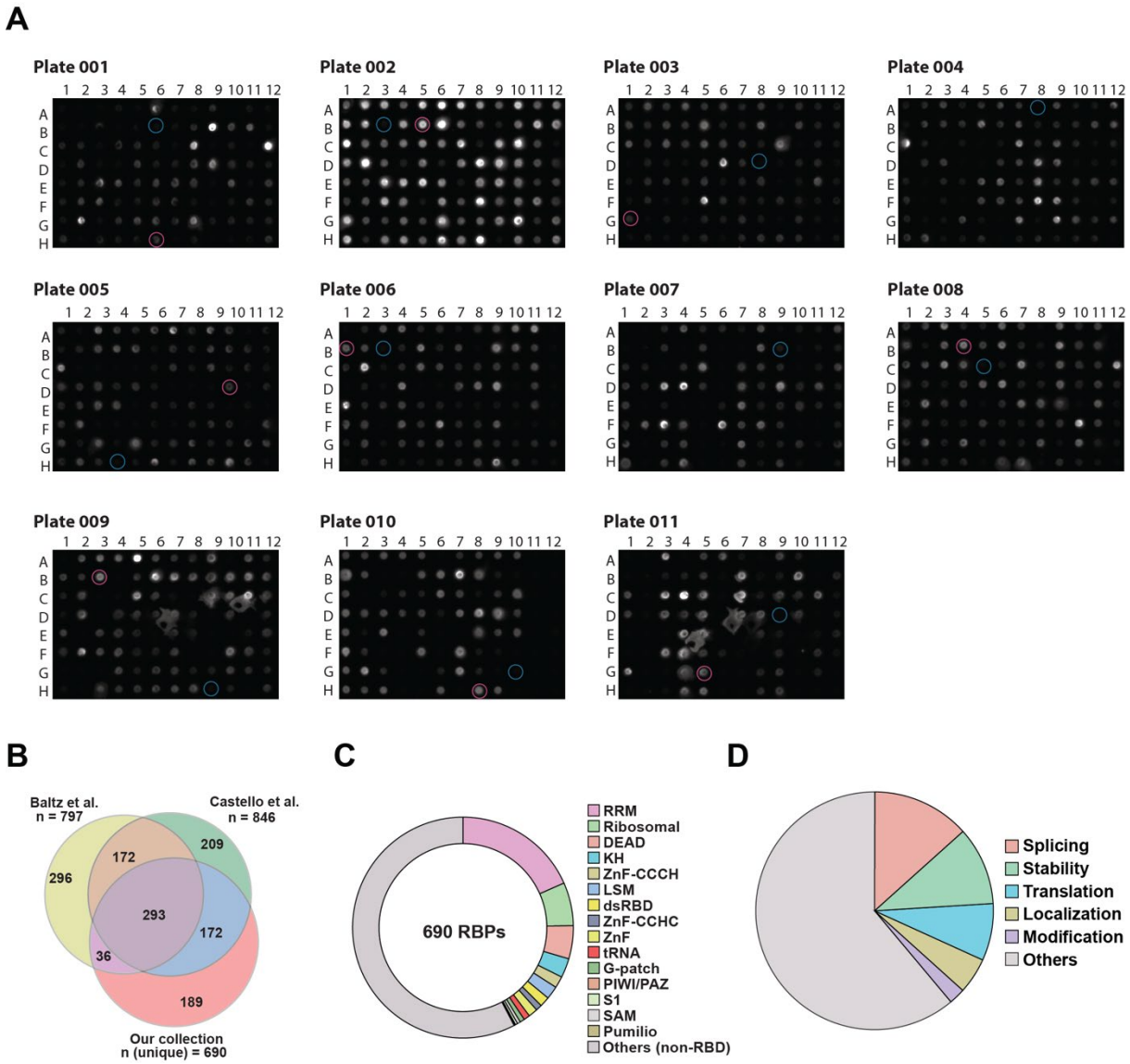


Figure 2. 1 MCP-V5-tagged RBP ORFs

(A) Western dot blot analysis of transiently expressed MCP-V5-tagged RBP ORFs in HeLa cells using a V5 antibody. Blue circles denote negative controls (no plasmid), red circles denote positive controls (CNOT7-V5-MCP). The order of wells and fold changes over negative controls are listed in **Extended Data Table 1**.

(B) Our collection of 771 open reading frames (ORFs) for 690 unique RBPs and their overlap with those identified experimentally by Baltz et al.(7) and Castello et al.(8).

(C) Distribution of known classical and non-classical RNA-binding domains in RBPs represented in our library.

(D) Distribution of molecular categories for RNA-related functions of RBPs represented in our library.

Next, we constructed a set of luciferase reporter plasmids that measure the effect of RBP recruitment to the 3'UTR on reporter expression. F-Luc-6MS2 encodes firefly luciferase followed by 6 MS2 hairpin sequences inserted into the 3'UTR context of *HBB* (β-globin). To evaluate potential reporter context dependencies, we also generated a corresponding *Renilla* luciferase construct. Matched constructs lacking MS2 sequences served as negative controls (**Figure 2.2A**). To validate our system, we co-introduced each reporter into HeLa cells along with constructs expressing MCP-fused and unfused versions of ZFP36 (also known as Tristetraprolin, TTP), an RBP activator of AU-rich element (ARE)-mediated RNA decay (27), enhanced GFP (EGFP), or the FLAG peptide. As expected, ZFP36 but not enhanced GFP (EGFP) or the FLAG peptide, dramatically reduced protein levels of the luciferase reporter in a manner that depended on the presence of the tether but not the identity of the luciferase protein. This demonstrated that tethered ZFP36 can recruit functional CCR4-NOT deadenylase complexes, which contain the Caf1 subunit CNOT7 (an RNase), to the reporter. Tethering of CNOT7 itself recapitulated this finding, indicating that productive recruitment is not limited to sequence-specific RBPs (such as ZFP36), but extends to effector RBPs (such as CNOT7) (**Figure 2.2B**).

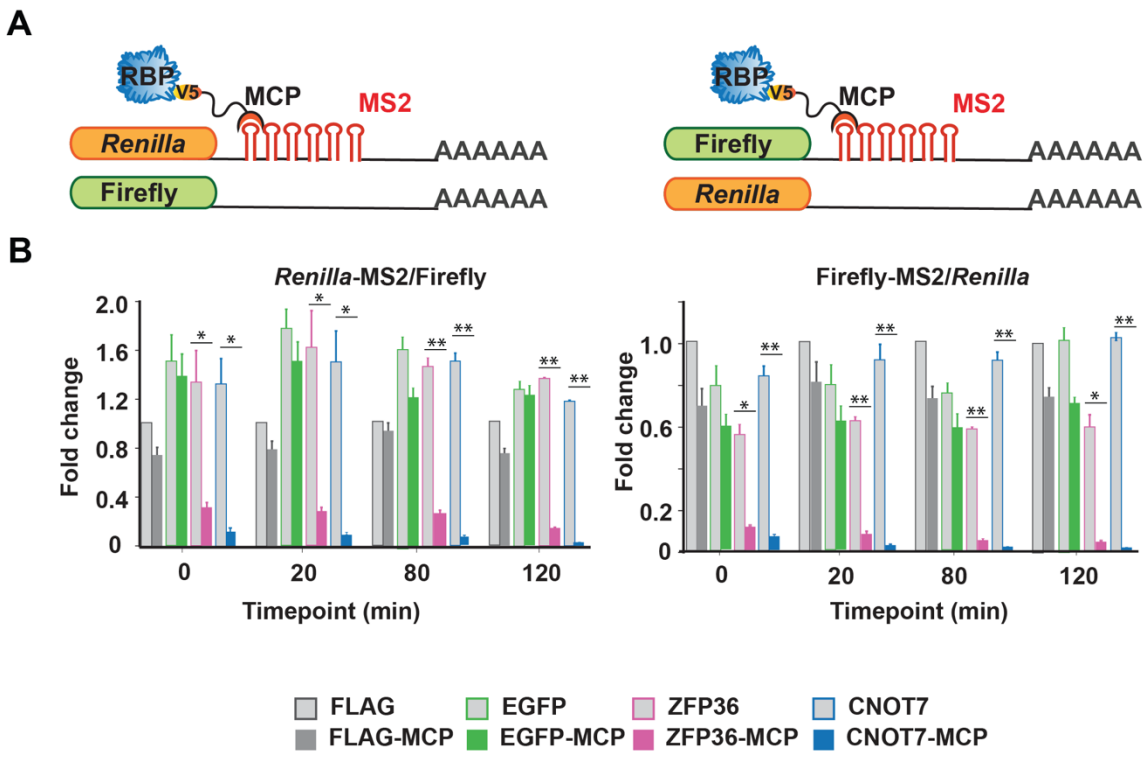


Figure 2. 2 Schematic of tethered luciferase reporter.

(**A**) Schematic of luciferase reporters. The coding region for either firefly (top) or *Renilla* (bottom) luciferase contain 6 MS2 stem-loop structures in the 3'UTR. The complementary reporters lacking MS2 hairpins were used as internal controls to normalize reporter signals. RBPs fused C-terminally to the MS2 coat protein (MCP), which recognizes MS2 hairpins with high affinity, are co-expressed with the reporters in a HeLa cell line.

(**B**) Bar plots showing the activity of the luciferase reporters in the presence of co-expressed known negative regulators of RNA stability (CNOT7, blue bars; and ZFP36, pink bars) or negative controls (the FLAG peptide, gray bars; EGFP, green bars; without MCP fusion, light shading; with MCP fusion, dark shading). Values are expressed as ratio of the median luciferase activity in the presence of the indicated RBPs, relative to that untagged FLAG controls at timepoint 0. Left and right panels in correspond to top and bottom reporter pairs in (b), respectively. Error bars denote mean±SD (n=3 replicate transfections). * $p<0.005$, ** $p<0.0005$ (two-tailed Student's *t*-test).

2.3.2 Large-scale tethered function screen reveals RBPs that affect reporter luciferase levels

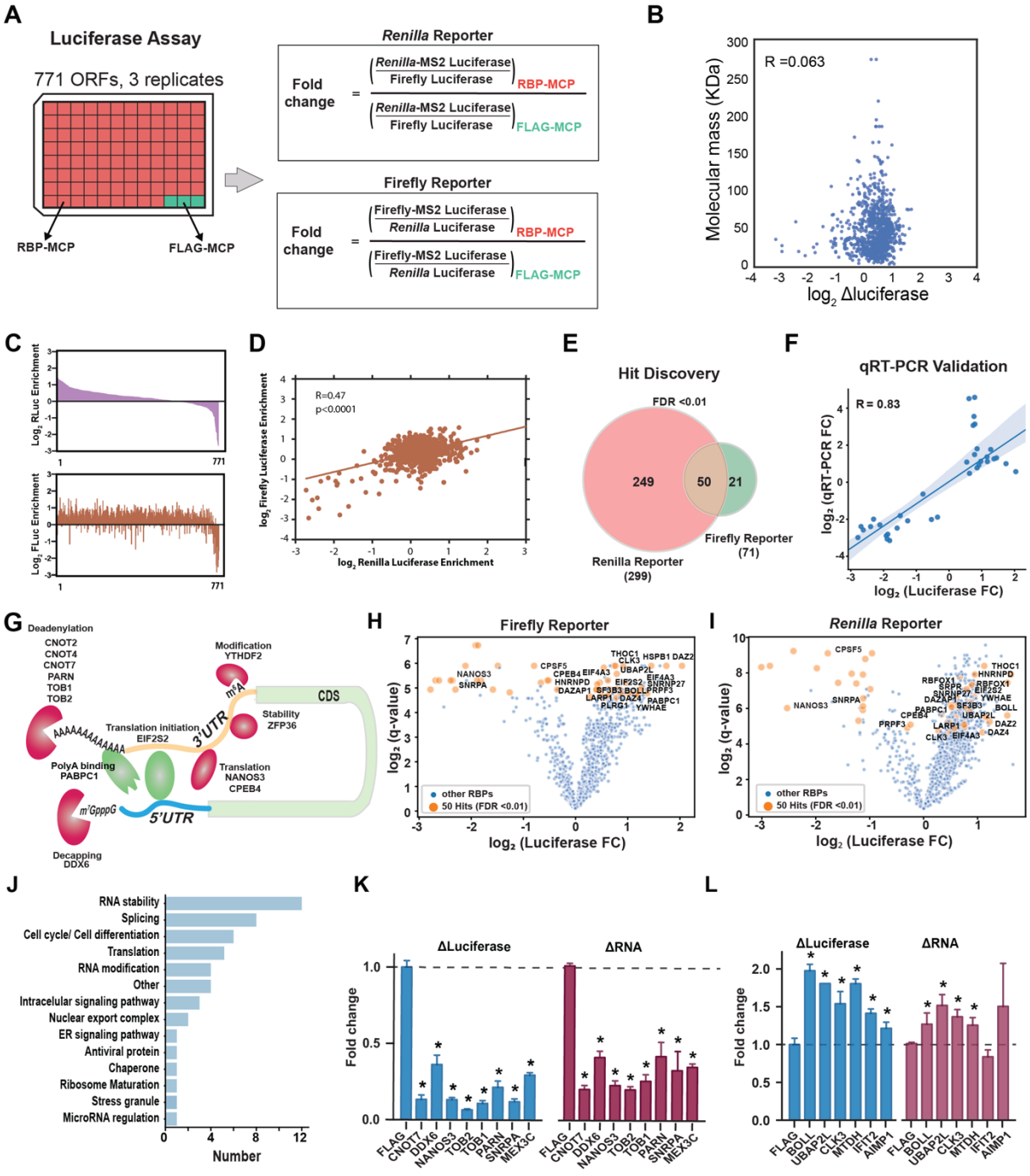
Next, we screened the 771 ORFs in triplicate using our two dual luciferase reporter systems described above, and the FLAG expression construct as a negative control (**Figure 2.3a, *left***). We calculated the effect of RBP recruitment to the tethering reporter as the fold change in luciferase activity relative to FLAG control, after normalization to the respective untethered reporters (**Figure 2.3A, *right*; Extended Data Table 2**). Supporting the validity of our screening approach, we confirmed that the effect was not correlated with RBP size (R=0.063), indicating that steric hindrance is unlikely to account for these observations (**Figure 2.3B**). Although the magnitude of the effect on reporter transcript abundance generally depended on the reporter system (R=0.47), the luciferase enrichment levels were significantly correlated (p<0.0001) between them (**Figure 2.3C-D**). We prioritized candidates from each reporter assay by using multiple t-tests at a threshold p<0.05 and calculated false discovery rates (FDR) for each comparison using the Benjamini, Krieger & Yekutieli procedure(28). We identified 299 and 71 RBPs with an estimated FDR <0.01 in *Renilla* and firefly reporters, respectively, indicating that reporter contexts do factor in the regulatory impact of tethered RBPs. 50 RBPs were recovered from both reporter contexts (**Figure 2.3E**). As an independent measure of the effect that the tethered RBP exerted on the reporter, we measured both luciferase transcript levels by reverse transcription quantitative PCR (RT-qPCR) for 35 of the 50 RBPs. In general, the change in reporter translation levels by luciferase was positively correlated (R=0.83) with changes in reporter transcript levels by qRT-PCR (**Figure 2.3F; Extended Data Tables 3 and 4**). Among the strongest candidate negative regulators were RBP components of both deadenylation-dependent and -independent exonuclease decay pathways, including ZFP36, as

well as members of the CCR4-NOT deadenylase complex (CNOT2, CNOT4, CNOT7, TOB1, and TOB2), the 3'-to-5' exonuclease PARN, and the decapping activator DDX6, which is recruited to the 5' cap via interaction with the CCR4-Not complex(29). As another positive control, we also confirmed that YTHDF2, a member of the YTH domain family of N^6-methyladenosine binding proteins, which recruit target RNAs to degradation bodies(30), exerts a negative effect on target mRNA levels. The results of our screen also confirmed several known negative regulators of translation, such as NANOS3 specific to germ cells (31), and CPEB4, which binds polyadenylation elements in the 3' UTR and negatively regulates translation initiation by interacting with the translation initiation factor eIF3 (ref. (32)). Interestingly, EIF2S2, with roles in promoting translation initiation, emerged as positive regulator of translation when recruited to the 3' UTR. We speculate that recruitment of this protein to the 3' UTR brings it into proximity to the mRNA cap and 5'UTR, similar to DDX6 and CPEB4 and consistent with the closed-loop model of translation(33, 34) (**Figure 2.3G**). Nevertheless, while these proteins are previously proposed as translational regulators, we also see changes in luciferase mRNA upon tethering. This appears consistent with the close coupling between translation and mRNA stability (35-37).

In summary, the screen revealed candidate regulators previously annotated to be linked to post-transcriptional gene regulation of metabolic processes, cell cycle, cell differentiation (DAZ family proteins BOLL, DAZ2 and DAZ4; DAZAP and NANOS3; refs. (38-40)), stress granule-associated proteins (UBAP2L(41-43)), factors involved in translation (EIF2S2, LARP1, PABPC1, and CPEB4; refs. (32, 44-46)), ER proteins (SRPR(47)), and heat shock proteins (HSPB1). Our screen also identified eight annotated splicing factors (CLK3, CPSF5, PLRG1, PRPF3, RBFOX1, SF3B3, SNRNP27, and SNRPA; refs. (48-50)) and two nuclear

export complex proteins (HNRNPD and THOC1; refs. (51, 52)) as candidates (**Figure 2.3H-J**). We next reconfirmed 14 RBPs with significant effects (8 that enhanced and 6 that repressed luciferase mRNA and protein levels) for further investigation (**Figure 2.3K-L**).

Figure 2. 3 A large-scale tethered function screen identifies RBPs regulating stability and translation.
(A) Experimental and analysis workflow. The screen was conducted on 771 MCP-tagged ORFs in two reporter contexts. Levels of MS2-tagged luciferase reporters were normalized to untagged co-transfected controls reporters. The effect of RBP recruitment was calculated as the ratio of normalized luciferase levels in the presence of MCP-tagged RBPs relative to that of MCP-FLAG control.
(B) Scatter plot of RBP size and luciferase effect. R, Pearson correlation coefficient.
(C) Luciferase activities from two different reporter constructs. Bar graphs showing $\log_2$-fold changes of the activity of *Renilla* (top) or firefly (bottom) luciferase reporters in presence of the MCP-fusion ORFs over FLAG control. Each vertical line represents a tethered ORF.
(D) Scatter plot of luciferase activities from the two reporter constructs. Values are expressed as $\log_2$-fold changes of the mean luciferase activity in the presence of MCP-fused ORFs over FLAG controls. R, Pearson correlation coefficient.
(E) Hit discovery. RBPs with effects at estimated FDR <0.01 in both reporter assays were considered candidate regulators.
(F) qPCR validation of reporter levels for 35 candidate RBP regulators. Means (n=3 independent measurements) of $\log_2$-transformed fold-changes of reporter mRNA levels, calculated analogously to (d), were plotted against the corresponding $\log_2$-transformed fold-changes of reporter luciferase levels. The line represents the least-squares linear regression fit. Shaded areas denote the 95% confidence interval. R, Pearson correlation coefficient.
(G) Examples of our 50 candidate RBP regulators that are known to affect RNA stability and translation.
(H–I) Volcano plots showing the distribution of fold changes for 50 RBP hits from the **(h)** firefly and (**i**) *Renilla* reporter assays. RBPs with FDR <0.01 are in orange. q-values were calculated as described(28).
(J) Gene ontology (GO) classification of candidate RBP regulators by manual curation.
(K–L) Validation of the **(K)** 9 negative and **(L)** 6 positive candidate regulators of RNA stability and/or translation by repeat luciferase (blue bars) and RT-PCR (red bars) measurements. Values were calculated as in (f). Error bars denote mean±SD for n=4 replicate transfections. *p<0.05 (two-tailed Student's t-test) vs. FLAG control.


A
Luciferase Assay
771 ORFs, 3 replicates
RBP-MCP
FLAG-MCP
Renilla Reporter
Fold change = (Renilla-MS2 Luciferase / Firefly Luciferase)RBP-MCP / (Renilla-MS2 Luciferase / Firefly Luciferase)FLAG-MCP
Firefly Reporter
Fold change = (Firefly-MS2 Luciferase / Renilla Luciferase)RBP-MCP / (Firefly-MS2 Luciferase / Renilla Luciferase)FLAG-MCP
B
R =0.063
Molecular mass (KDa)
log2 Δluciferase
C
Log2 RLuc Enrichment
Log2 FLuc Enrichment
D
R=0.47
p<0.0001
log2 Firefly Luciferase Enrichment
log2 Renilla Luciferase Enrichment
E
Hit Discovery
FDR <0.01
249
50
21
Renilla Reporter (299)
Firefly Reporter (71)
F
qRT-PCR Validation
R= 0.83
log2 (qRT-PCR FC)
log2 (Luciferase FC)
G
Deadenylation
CNOT2
CNOT4
CNOT7
PARN
TOB1
TOB2
Modification
YTHDF2
m6A
CDS
Stability
ZFP36
Translation initiation
EIF2S2
3'UTR
AAAAAAAAAAAAA
PolyA binding
PABPC1
Translation
NANOS3
CPEB4
m7GpppG
5'UTR
Decapping
DDX6
H
Firefly Reporter
log2 (q-value)
log2 (Luciferase FC)
other RBPs
50 Hits (FDR <0.01)
I
Renilla Reporter
log2 (q-value)
log2 (Luciferase FC)
other RBPs
50 Hits (FDR <0.01)
J
RNA stability
Splicing
Cell cycle/ Cell differentiation
Translation
RNA modification
Other
Intracelular signaling pathway
Nuclear export complex
ER signaling pathway
Antiviral protein
Chaperone
Ribosome Maturation
Stress granule
MicroRNA regulation
Number
K
ΔLuciferase
ΔRNA
Fold change
FLAG
CNOT7
DDX6
NANOS3
TOB2
TOB1
PARN
SNRPA
MEX3C
L
ΔLuciferase
ΔRNA
Fold change
FLAG
BOLL
UBAP2L
CLK3
MTDH
IFIT2
AIMP1

2.3.3 Enhanced CLIP identifies endogenous RNA targets of candidate regulators

We focused on RBPs for which roles in RNA stability and/or translation are not known (UBAP2L, SNRPA, CLK3, MTDH, AIMP1, and IFIT2) and RBPs with known roles but where transcriptome-wide binding sites and preferences have not been described (CNOT7, DDX6, NANOS3, TOB1/2, PARN, MEX3C, and BOLL) (**Figure 2.4A**). We investigated their endogenous mRNA targets and their transcriptome-wide binding sites using enhanced cross-linking immunoprecipitation followed by sequencing (eCLIP) (1). Briefly, HEK293T cells were subjected to UV-crosslinking, lysis and RNA fragmentation, and protein-RNA complexes were immunoprecipitated using validated RBP-specific antibodies (53) (**Figure 2.4B**). We also transiently transfected cells with plasmids expressing V5-tagged fusions of those candidate RBPs which are not expressed in HEK293T cells or do not have RBP-specific antibodies (**Figure 2.4C**). In total, we generated eCLIP datasets in duplicates for the 14 candidate proteins, each replicate consisting of an RBP eCLIP (IP) library and a paired size-matched input (SMInput)(1) library. Libraries were sequenced to at least 14M (million) reads (average of 27M), of which at least 1M (average of 7M) mapped uniquely to the human genome. All libraries passed our routine quality control metrics (19). We observed an average Pearson correlation coefficient above 0.5 between replicates (**Figure 2.4D**). In summary, eCLIP libraries were successfully performed, and yielded reproducible RBP-specific global binding profiles.

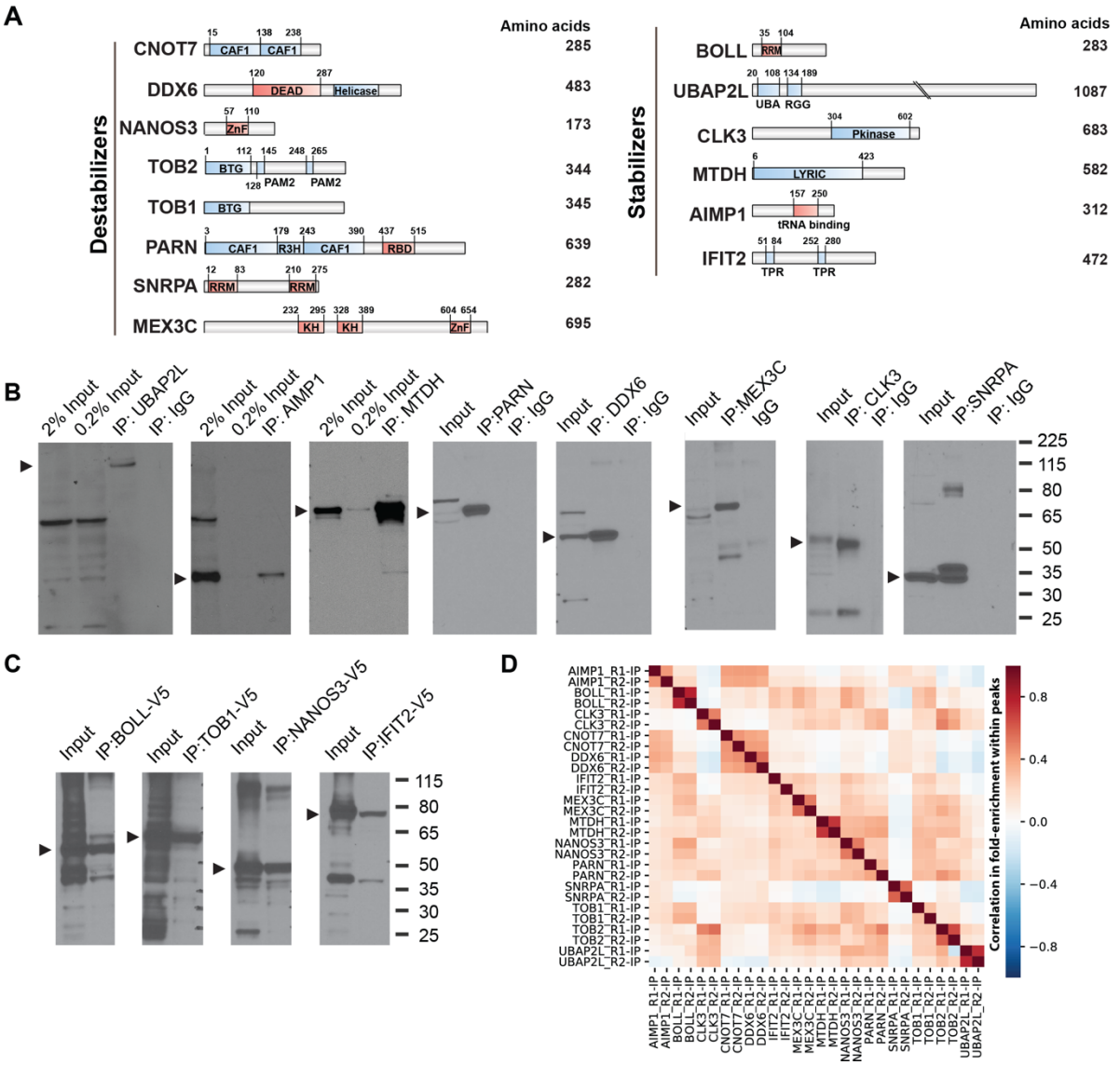


Figure 2. 4 CLIP identifies endogenous RNA targets

(**A**) Domain structures of 14 candidate RBPs with RNA destabilizing (left) and stabilizing (right) effects in the tethering assay, with lengths of their polypeptide chains.

(B-C) In-line western blots of eCLIP immunoprecipitations of candidate RBPs. (B) Extracts from HEK293T cells or (b) from HEK293T transfected with the indicated V5-tagged RBP ORFs immunoprecipitated with non-immune (IgG) control antibodies, and western blot analysis using either RBP-specific (B) or anti-V5 (C) antibodies. The molecular weight (in kDa) of standards are indicated on the right. Arrowheads indicate the calculated molecular weight for each RBP or RBP-V5 fusion protein.

(**D**) Heatmap of the Pearson correlation coefficients of fold-enrichment of eCLIP peaks for the indicated 14 RBPs analyzed in duplicate.

Next, we determined transcript binding region specificities using two distinct metrics, namely read density and binding cluster enrichment. Read density enrichment within 5' and 3'UTRs and coding regions (CDS) of annotated protein coding genes were computed by the fold enrichment in the IPs normalized to their paired SMInputs for target transcripts. To illustrate, BOLL, a germ-cell specific RBP with some documented roles in mRNA stabilization and translation enhancer activity, displayed a strong preference for 3'UTR association. Surprisingly, IFIT2 (Interferon Induced Protein With Tetratricopeptide Repeats 2), which is known to inhibit expression of viral messenger RNAs, robustly displayed a strong 3'UTR preference in human mRNAs (**Figure 2.4A**). The helicase DDX6 was enriched for binding within 5'UTRs and 3'UTRs, consistent with its role in the assembly of the decapping complex and the closed-loop model of translation (33, 34). Unexpectedly, TOB family member TOB1, which recruits the catalytic subunits of the CCR4-NOT deadenylase complex to target mRNAs (54), and 3'-to-5' exonuclease PARN which degrades poly(A) tails, showed preferences for 5'UTRs in addition to 3'UTRs and CDS, suggesting unexpected roles for these proteins (**Figure 2.4B**). UBAP2L (Ubiquitin-associated protein 2-like) showed strong enrichment across CDS exons and 5' UTR (**Figure 2.5C**).

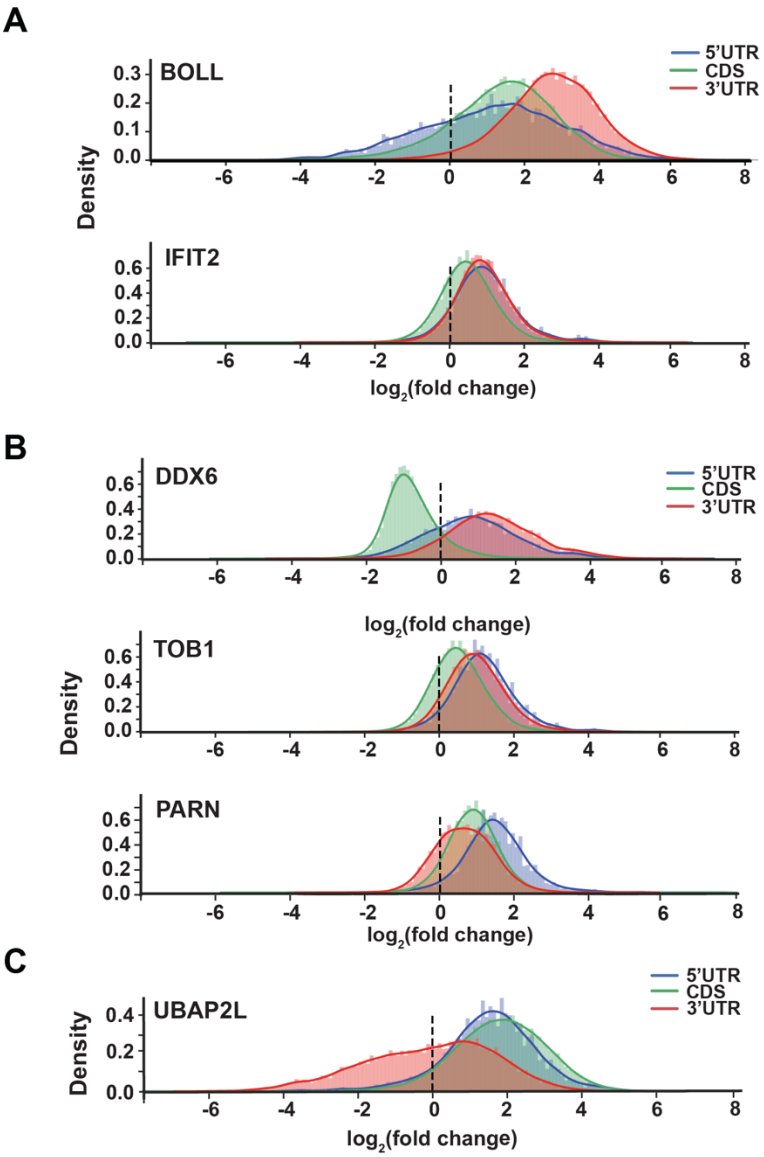
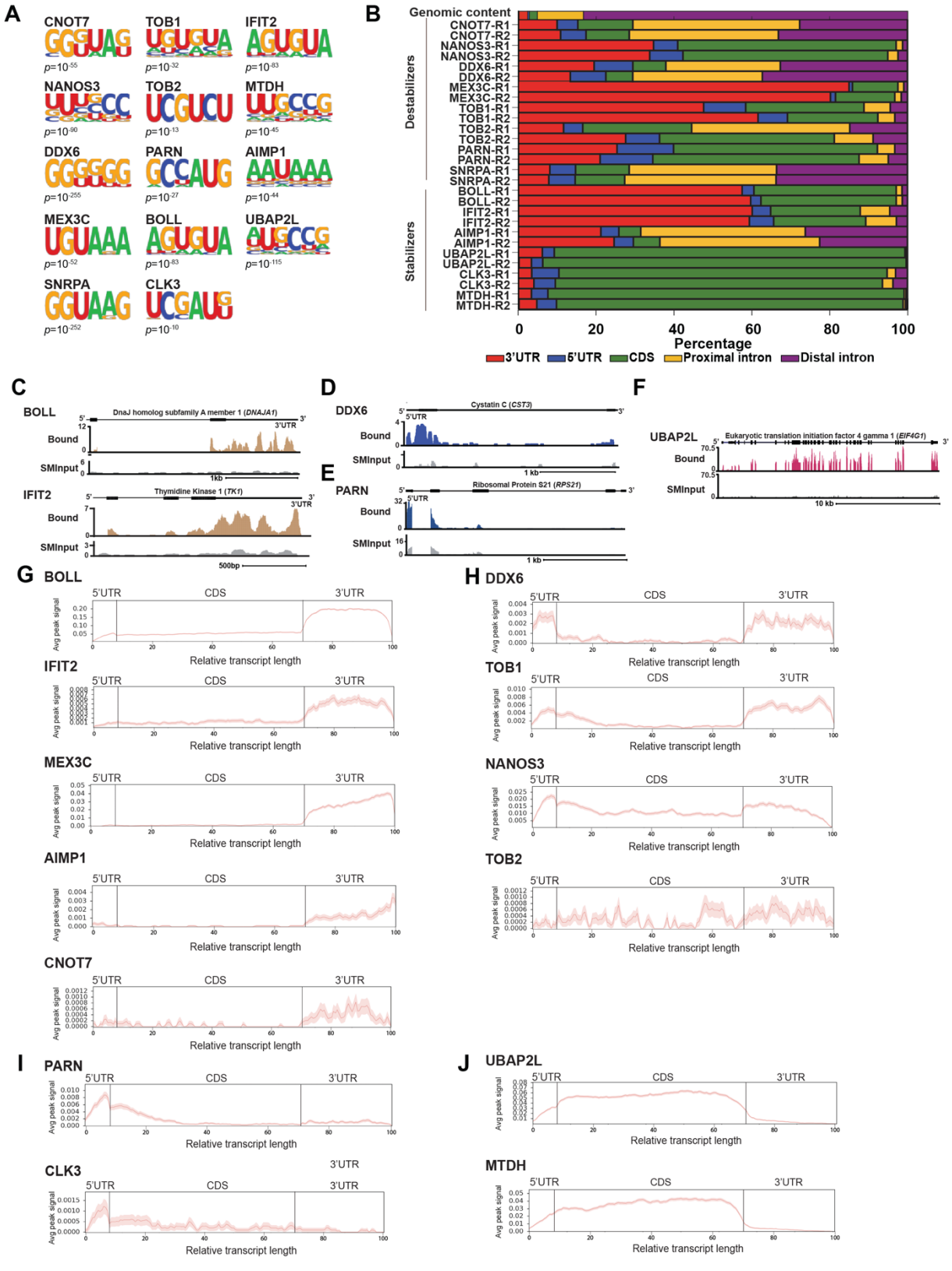


Figure 2. 5 Regional analysis shown the binding preferences of candidate RBPs
(**A-C**) Histograms showing region-based fold-enrichment of read densities, normalized to paired SMInput controls for **(A)** BOLL and IFIT2, which show read density enrichment in 3'UTRs; **(B)** DDX6, TOB1, and PARN which show read density enrichment in 5'UTRs; and **(C)** UBAP2L, which shows read density enrichment in CDS and 5'UTR.

To identify binding sites at higher resolution, clusters were discovered by the CLIPper algorithm(55) and reproducible binding sites were defined as clusters that overlapped between both replicates. Cluster enrichment was computed by calculating the ratio of read densities between IPs and SMInputs within a reproducible cluster and significant clusters were defined as $p \leq 10^{-2}$ (Fisher's exact test for read numbers <5; χ^2 test for read numbers ≥5) and ≥4-fold-enriched over SMInput(1). Significant clusters were enriched for specific sequence motifs (**Figure 2.6A**), some of which reflect expected preferences. For example, the SNRPA motif GGUAAG resembles the 5' splice site consensus (GGURAG), and the helicase DDX6 motif GGGGGG is consistent with its binding preferences to G-quadruplex RNA (56). Interestingly, the binding motif identified for BOLL (AGUGUA) partially overlaps with the pumilio response element (PRE) UGUANAUA, consistent with complex formation of DAZ family proteins with PUM2 on RNA targets(57, 58). Binding cluster enrichment analyses of these RBPs which focus on the binding sites with high signal (**Figure 2.6B**) generally agree with read density enrichment analyses. Meta-gene plots further reinforce that BOLL and IFIT2 (**Figure 2.6C**), as well as MEX3C, AIMP1 and CNOT7 are 3'UTR-preferring binders (**Figure 2.6F**). DDX6 (**Figure 2.6D**), TOB1, NANOS3 and TOB2 appear to have 5' and 3'UTR, but lower CDS, preferences in the meta-gene plots (**Figure 2.6G**). PARN (**Figure 2.6E**) and CLK3 are enriched at the 5'UTR, peaking near the start codon (**Figure 2.6I**). UBAP2L and MTDH binding clusters were predominantly in CDS (**Figure 2.6F**). Indeed, UBAP2L clusters were dispersed across exonic regions (**Figure 2.6J**). Overall, our analyses not only revealed previously unrecognized binding maps and preferences for RBPs known to affect mRNA stability and translation, but also revealed novel RNA interactomes of candidate RBPs.

Figure 2. 6 A large-scale tethered function screen identifies RBPs regulating stability and translation.
(A) De novo sequence motifs in significant eCLIP peaks of the indicated RBP candidates enriched above background, with associated binomial p-value.
(B) Bar graphs showing eCLIP binding cluster distribution across transcript regions for the 8 destabilizers and 6 stabilizers. Peak assignment was performed using stringent enrichment criteria (≥4-fold-enrichment and $p \leq 10^{-2}$ versus SMInput). The region distribution of the entire transcriptome annotated in GENCODE v19 is indicated at the top.
(C-F) Example genome browser track views of eCLIP read densities (in reads per million, RPM) and corresponding SMInput read densities for (C) BOLL and IFIT2, which show peak enrichment in 3'UTRs, (D) DDX6 and (E)PARN, which show peak enrichment in 5'UTRs, and (F) UBAP2L, which shows peak enrichment across exons.
(G-J) Meta-gene maps showing the distribution of eCLIP peak densities at target transcripts. The x-axis indicates the relative length of each regions. Dark red lines indicate the average number of significantly enriched peaks (≥4-fold-enriched and $p \leq 10^{-2}$ versus SMInput) of eCLIP peak densities at all transcripts for (G) BOLL, IFIT2, MEX3C, AIMP1, and CNOT7, which show peak enrichment in 3'UTR; (H) DDX6, TOB1, NANOS3, TOB2, which show peak enrichment in 5'UTR/ 3'UTR; (I) PARN and CLK3, which show peak enrichment in 5'UTR; and (J) UBAP2L and MTDH, which show peak enrichment in CDS. Light shaded areas denote the 95% confidence interval.


A
CNOT7
p=10-55
TOB1
p=10-32
IFIT2
p=10-83
NANOS3
p=10-90
TOB2
p=10-13
MTDH
p=10-45
DDX6
p=10-255
PARN
p=10-27
AIMP1
p=10-44
MEX3C
p=10-52
BOLL
p=10-83
UBAP2L
p=10-115
SNRPA
p=10-252
CLK3
p=10-10
B
Genomic content
Destabilizers
Stabilizers
CNOT7-R1
CNOT7-R2
NANOS3-R1
NANOS3-R2
DDX6-R1
DDX6-R2
MEX3C-R1
MEX3C-R2
TOB1-R1
TOB1-R2
TOB2-R1
TOB2-R2
PARN-R1
PARN-R2
SNRPA-R1
SNRPA-R2
BOLL-R1
BOLL-R2
IFIT2-R1
IFIT2-R2
AIMP1-R1
AIMP1-R2
UBAP2L-R1
UBAP2L-R2
CLK3-R1
CLK3-R2
MTDH-R1
MTDH-R2
0
20
40
60
80
100
Percentage
3'UTR
5'UTR
CDS
Proximal intron
Distal intron
C
BOLL
DnaJ homolog subfamily A member 1 (DNAJA1)
Bound
SMInput
1kb
IFIT2
Thymidine Kinase 1 (TK1)
500bp
D
DDX6
Cystatin C (CST3)
1 kb
E
PARN
Ribosomal Protein S21 (RPS21)
F
UBAP2L
Eukaryotic translation initiation factor 4 gamma 1 (EIF4G1)
10 kb
G
BOLL
IFIT2
MEX3C
AIMP1
CNOT7
H
DDX6
TOB1
NANOS3
TOB2
5'UTR
CDS
3'UTR
Avg peak signal
Relative transcript length
I
PARN
CLK3
J
UBAP2L
MTDH

2.3.4 Integration of eCLIP and RNA-seq data defines regulatory classes of RBPs and transcripts

To gain insight into how our candidate RBPs affect transcriptome-wide mRNA levels, we depleted or exogenously expressed them in HEK293T cells and performed RNA-seq analysis. Specifically, we either depleted RBPs by lentiviral transduction of short-hairpin RNAs (shRNAs) (**Figure 2.7A**), or ectopically expressed ORFs of those candidate RBPs which are not expressed in HEK293T cells or which do not have RBP-specific shRNAs (**Figure 2.7B**). For each RBP, we either performed two independent transductions of two different targeting shRNAs and two non-targeting shRNAs, or performed two independent transfections with a plasmid directing expression of the RBP as a V5-tagged fusion, with the FLAG construct as a control. We selected polyA$^+$ RNA, prepared sequencing libraries and sequenced them to a depth of >32M (or >26M uniquely mapped) reads.

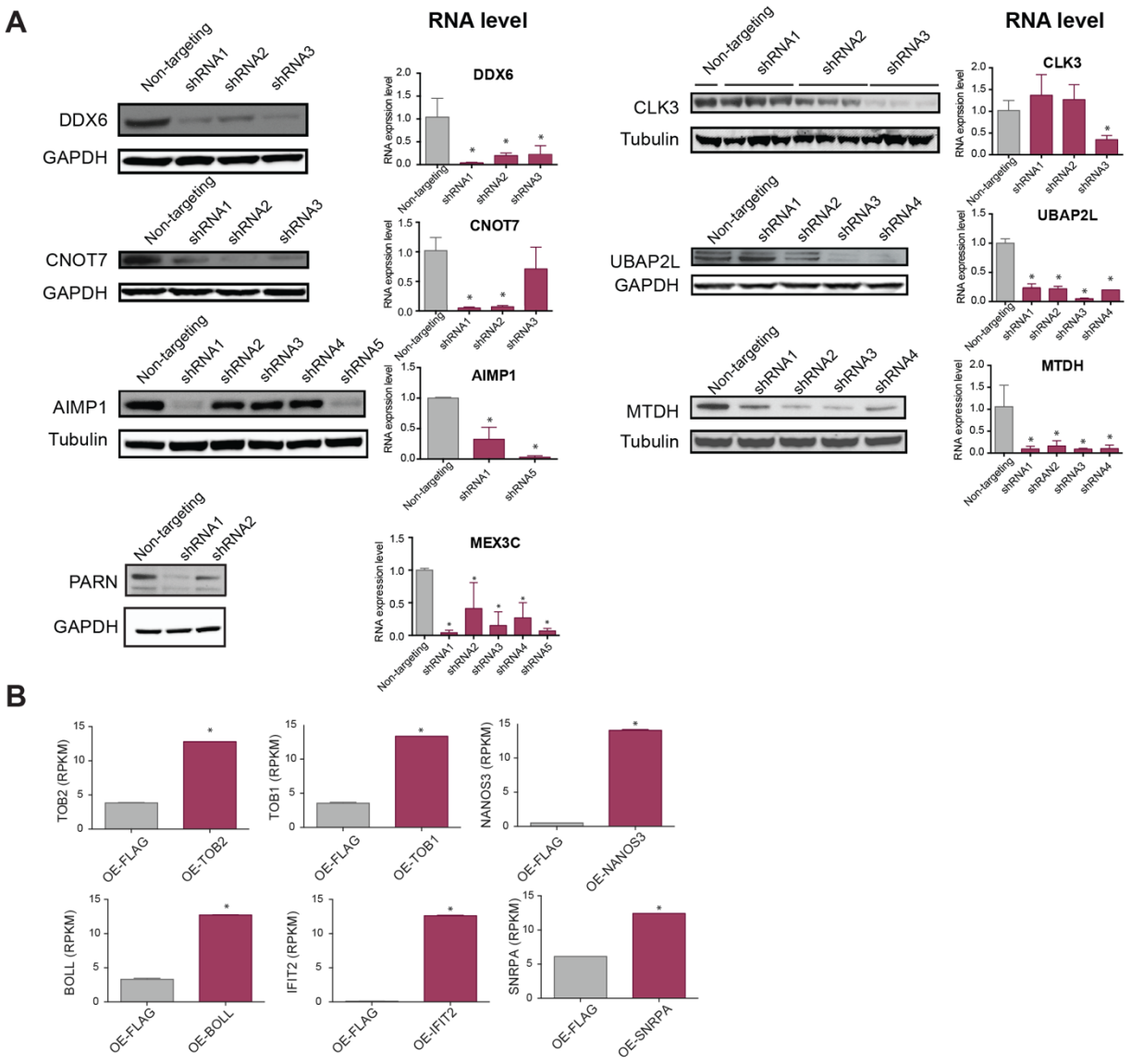


Figure 2. 7 shRNA-mediated depletion or overexpression of RBPs

(**A**) shRNA-mediated depletion of RBPs in HEK293T cells using 3-5 distinct shRNAs for each RBP, as indicated, compared to non-targeting shRNA control. Western blots, with GAPDH or tubulin serving as loading controls, as indicated. Bar graphs indicating RBP transcript levels determined by qRT-PCR, normalized to levels of 18S rRNA. Data are shown as mean±SD (n=3 replicates). Asterisks denote significance at $p<0.05$ determined by two-sided Student's t-test.

(**B**) Overexpression of RBPs in HEK293T cells. Bar plots showing transcript levels (RPKM) for each RBP following transfection of RBP expression constructs or FLAG vector control. Asterisks denote significance at $p<0.05$ determined by two-sided Student's t-test. Data are shown as mean±SD (n=3 replicates).

To assess the effect of a candidate RBP on transcript levels, we measured the number of significantly up- or down-regulated genes upon knockdown or overexpression. In general, our manipulations of RBP levels resulted in a largely unperturbed population of transcripts, typically 80% at threshold of statistical significance (≥1.23-fold, FDR-corrected $p \leq 0.05$ versus non-targeting shRNA or FLAG control). This indicates that our candidate RBPs affect specific sets of target transcripts, instead of having effects on global transcript stability. When only considering those transcripts that were bound by the respective RBP, as measured by eCLIP (≥1 significantly enriched cluster per transcript), we observe higher numbers of targets that change in the direction anticipated by the tethering assays, than in the opposite direction, for candidate destabilizers (negative regulators) DDX6 and PARN (**Figure 2.8A**), as well as SNRPA **(Figure 2.8B**), and candidate stabilizers (positive regulators) UBAP2L and CLK3 (**Figure 2.8C**), as well as BOLL and IFIT2 (**Figure 2.8D**). In other words, knockdown of specific predicted destabilizers led to more up-regulated genes, whereas overexpression of destabilizers led to more down-regulated genes, and reciprocal effects were observed in the alterations of stabilizing RBPs.

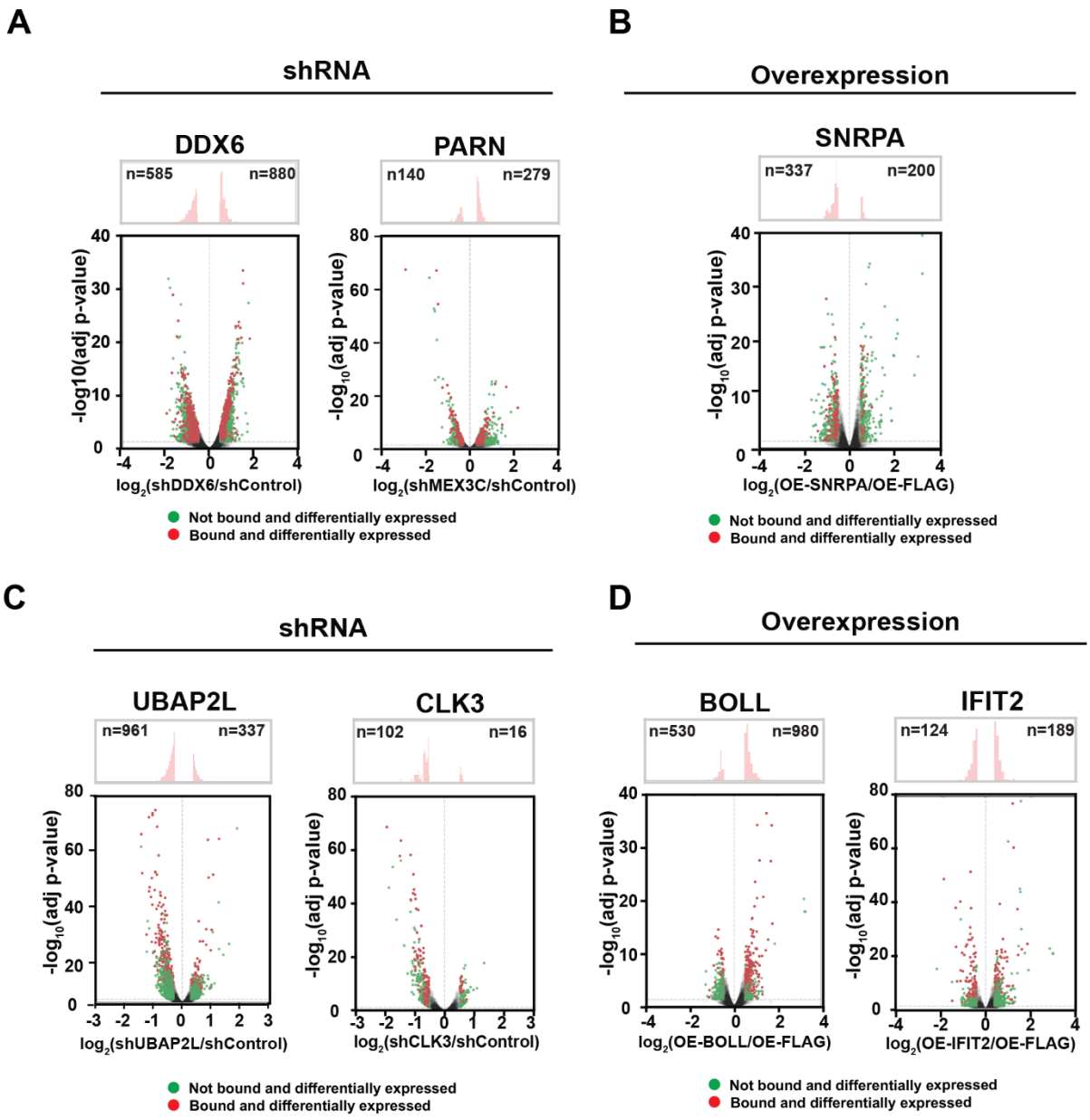


Figure 2. 8 Integration of eCLIP and RNA-seq data

(A–D) Volcano plots showing the distribution of fold changes in transcript levels upon modulation of (A-B) destabilizers and (C-D) stabilizers, with distribution histograms shown at the top. (A) Depletion of DDX6 (*left*) and PARN (*right*). (B) Overexpression of SNRPA. (C) Depletion UBAP2L (*left*), and CLK3(*right*). (D) overexpression of BOLL (*left*) and IFIT2 (*right*).Transcripts with $\log_2$(fold change) ≥1.23 and FDR-corrected $p \leq 0.05$ are in color, with red and green denoting transcripts with or without at least 1 significant RBP binding peak (≥4-fold-enriched and $p \leq 10^{-2}$ versus SMInput in eCLIP), respectively.

We also confirmed that the fraction of bound targets in the genes changing in the anticipated direction was statistically significantly enriched relative to unbound targets (**Figure 2.9A,B**). In fact, we observed significant correlation between different eCLIP cluster fold enrichments of IP over SMInput and change in transcript levels, for both candidate destabilizers (e.g. DDX6 and SNRPA; **Figure 2.9C,D**) and candidate stabilizers (e.g. UBAP2L and BOLL; **Figure 2.9E,F**). This indicates that our candidate RBPs directly engage hundreds of previously unknown endogenous target mRNAs to affect transcript levels in the predicted direction. For example, knockdown of the destabilizer PARN increased transcript levels of *RPS21* mRNA, a PARN-bound transcript (**Figure 2.9G**). Conversely, depletion of the stabilizer CLK3 reduced the abundance of its target *NELFCD* mRNA (**Figure 2.9H**). Interestingly, when we further evaluated which genic regions bound by the RBPs are most correlated with transcript levels, UBAP2L binding within CDS was the most enriched (**Figure 2.9I**). In general, we conclude that the majority of our candidate RBPs affect mRNA levels of their endogenous RNA targets, in agreement with our tethering results.

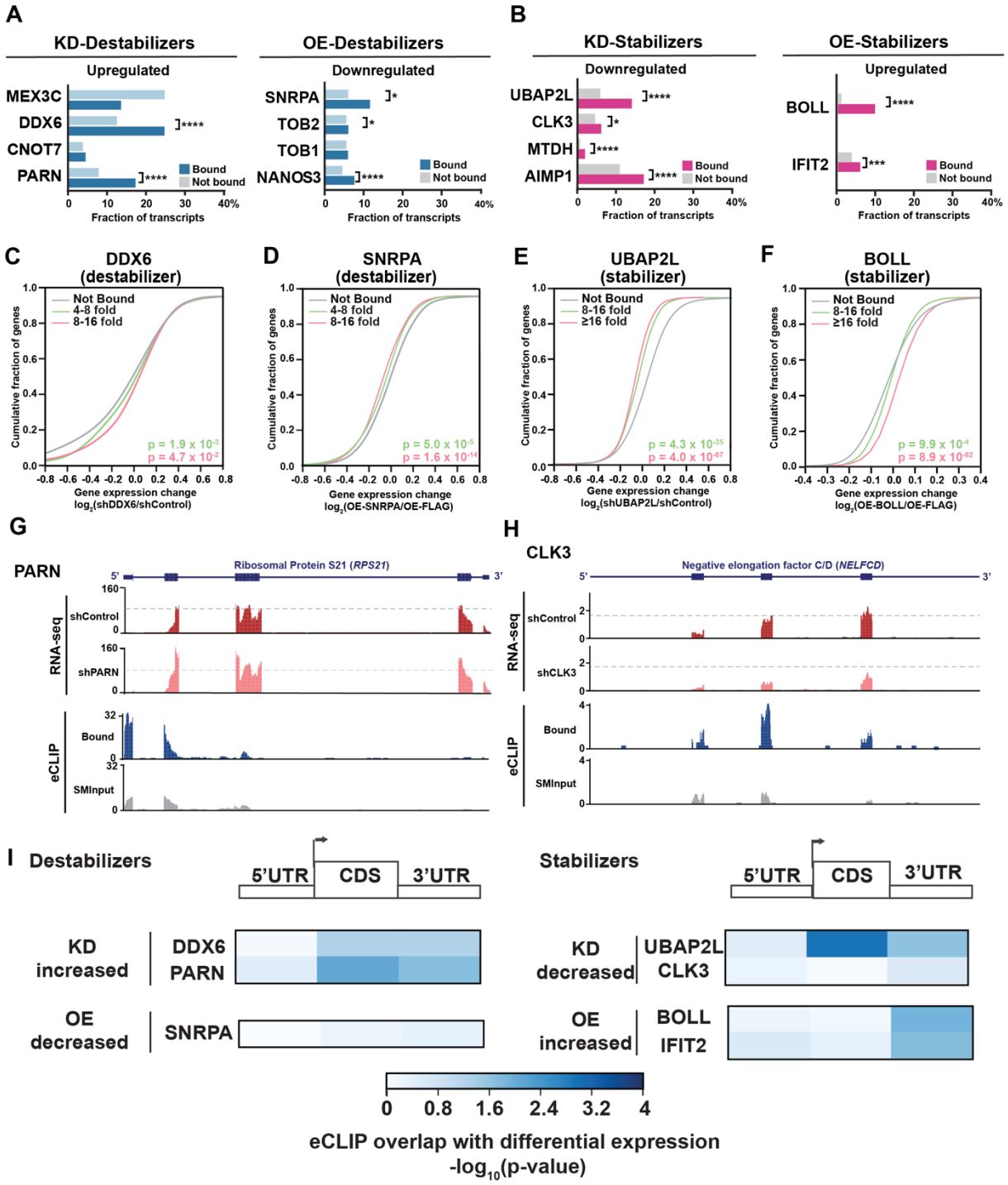
Figure 2. 9 Integration of eCLIP and RNA-seq data defines regulatory classes of RBPs and transcripts

(A-B) Bar plots showing the percentage of overlap between genes significantly up- or downregulated [$\log_2$(fold change) ≥1.23 and FDR-corrected p≤0.05] and significantly bound (≥4-fold-enriched and p≤10^{-2} versus SMInput in eCLIP) upon knockdown (KD) or overexpression (OE) of candidate **(A)** destabilizers and **(B)** stabilizers. *p<0.01, **p<10^{-2}, ***p<10^{-3} ****p<10^{-4} (hypergeometric test) vs control (non-targeting shRNA or FLAG overexpression, as appropriate).

(C–F) Cumulative distribution plots of transcript $\log_2$-transformed fold changes of overexpression versus vector control or shRNA-mediated knockdown vs non-targeting control, as indicated, for the destabilizers **(C)** DDX6 and **(D)** SNRPA, and the stabilizers **(E)** UBAP2L and **(F)** BOLL. Distributions are shown for transcripts showing the indicated significant read enrichments over SMInput (p≤10^{-2}) from eCLIP analysis (4-8 fold, green; 8-16 fold, red) or transcripts that are not significantly bound ('Not Bound', p>10^{-2} or <4-fold enrichment; gray). p-values for the distributions were calculated compared to 'Not Bound' using a two-tailed Mann–Whitney U test.

(G–H) Genome browser views from shRNA-mediated knockdowns showing RNA-seq reads (shRNA knockdown in maroon and non-targeting shRNA control in pink) and eCLIP reads (IP in blue and SMInput in gray) for **(G)** PARN at the *RPS21* locus and **(H)** CLK3 at the *NELFCD* locus. y-axes denote read density in reads per million (RPM).

(**I**) Heatmap showing significance in differential expression of genes significantly differentially expressed [$\log_2$(fold change) ≥1.23 and FDR-corrected p≤0.05] and significantly bound (≥4-fold-enriched and p≤10^{-2} versus SMInput in eCLIP) vs all unbound genes, upon knockdown (KD) or overexpression (OE) of candidate RBPs in each region. Significance was calculated using a two-tailed Mann–Whitney U test.


A
KD-Destabilizers
Upregulated
MEX3C
DDX6
CNOT7
PARN
Bound
Not bound
Fraction of transcripts
OE-Destabilizers
Downregulated
SNRPA
TOB2
TOB1
NANOS3
B
KD-Stabilizers
Downregulated
UBAP2L
CLK3
MTDH
AIMP1
OE-Stabilizers
Upregulated
BOLL
IFIT2
C
DDX6 (destabilizer)
Not Bound
4-8 fold
8-16 fold
p = 1.9 x 10-3
p = 4.7 x 10-2
Cumulative fraction of genes
Gene expression change
log2(shDDX6/shControl)
D
SNRPA (destabilizer)
p = 5.0 x 10-5
p = 1.6 x 10-14
log2(OE-SNRPA/OE-FLAG)
E
UBAP2L (stabilizer)
8-16 fold
≥16 fold
p = 4.3 x 10-35
p = 4.0 x 10-67
log2(shUBAP2L/shControl)
F
BOLL (stabilizer)
p = 9.9 x 10-4
p = 8.9 x 10-82
log2(OE-BOLL/OE-FLAG)
G
PARN
Ribosomal Protein S21 (RPS21)
RNA-seq
shControl
shPARN
eCLIP
Bound
SMInput
H
CLK3
Negative elongation factor C/D (NELFCD)
shCLK3
I
Destabilizers
5'UTR
CDS
3'UTR
KD increased
DDX6
PARN
OE decreased
SNRPA
Stabilizers
KD decreased
UBAP2L
CLK3
OE increased
BOLL
IFIT2
0 0.8 1.6 2.4 3.2 4
eCLIP overlap with differential expression
-log10(p-value)

2.3.5 UBAP2L increases mRNA polysome association and promotes translation

Among the candidates we analyzed, UBAP2L had the highest CDS read density enrichment, suggesting a role in translation. However, such a function for UBAP2L had not been described. We measured global protein synthesis rates in cells lacking UBAP2L with the SUnSET assay, which uses incorporation of puromycin (a structural analog of aminoacyl-transfer RNA) to label newly synthesized proteins (59). HEK293T cells biallelically deleted for *UBAP2L* by CRISPR/Cas9-mediated genome editing (**Figure 2.10A**) showed a ~40% reduction in protein synthesis (**Figure 2.10B**), indicating that UBAP2L promotes global translation.

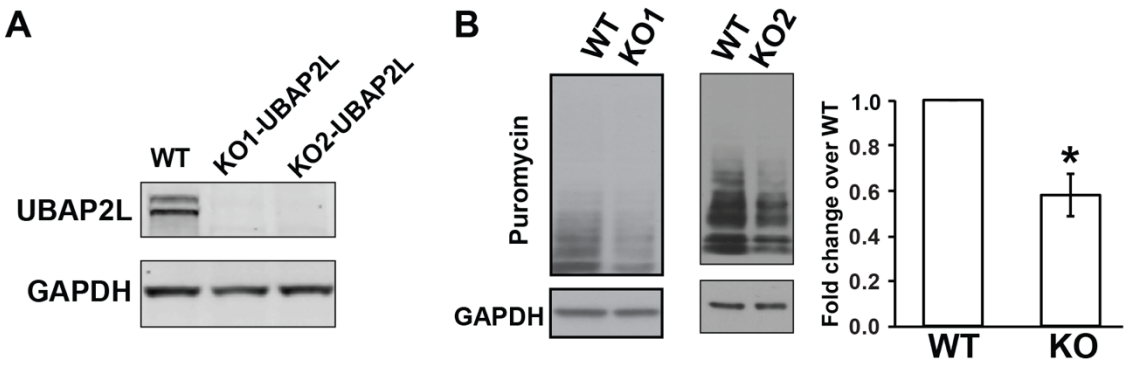


Figure 2. 10 Translation monitoring using puromycin incorporation.
(A) Western blots of extracts from control (WT) HEK293T cells and two independent clonal isolates with CRISPR-mediated disruption of *UBAP2L*.
(**B**) (*Left*) Representative anti-puromycin western blot of extracts from puromycin-treated WT and two independent KO clonal cells. GAPDH served as loading control. (*Right*) Densitometric quantitation of blots from of n=3 independent experiments. *p<0.05 vs. WT (two-tailed Student's *t*-test).

We next performed sucrose gradient centrifugation of HEK293T lysates to examine the association of UBAP2L with ribosomes. We detected enrichment of Eukaryotic Elongation factor 2 (EEF2) in 40S, 60S and monosome but not polysome fractions, as expected(60). UBAP2L co-fractionated with monosomes and polysomes, suggesting a role for UBAP2L in translation (**Figure 2.11A**). In order to rule out the possibility that this observation is due to the presence of UBAP2L in non-ribosomal complexes of similar buoyant density, we treated cells with puromycin to release polysomes from transcripts. Puromycin treatment led to accumulation of 80S monosomes, as expected, and levels of UBAP2L in fractions corresponding to polysomes were strongly reduced (**Figure 2.11B**). We also treated cell lysates with EDTA to disassemble 80S monosomes into 40S and 60S ribosomal subunits and found that, similarly, UBAP2L was depleted from fractions corresponding to monosomes (**Figure 2.11B**). These results strongly suggest that UBAP2L directly interacts with translating ribosomes.

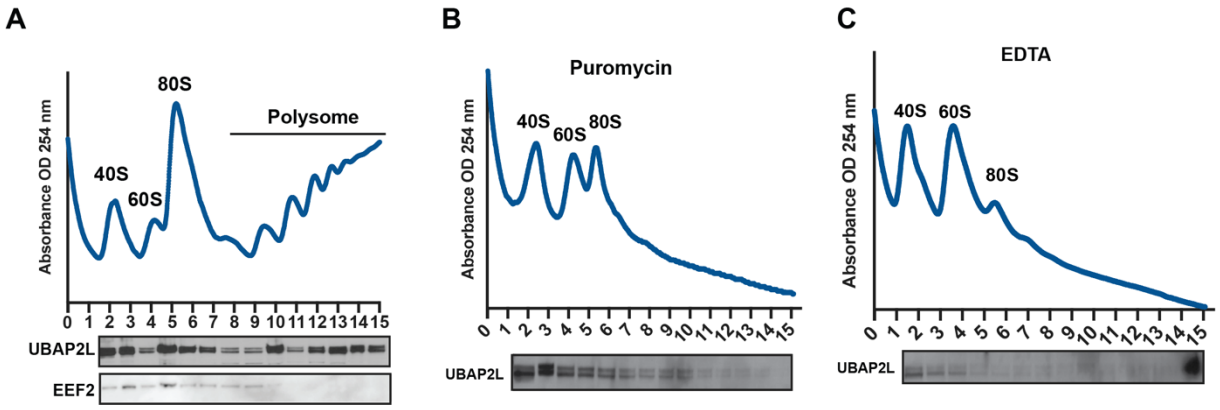


Figure 2. 11 Polysome profile shows UBAP2L associate with ribosome

(A) Polysome profile of UBAP2L. (*Top*) Absorbance (at 260 nm) plot of a HEK293T cell lysate fractionated through a 10-50% a sucrose gradient. (*Bottom*) Western blots of UBAP2L and EEF2 from corresponding sucrose fractions.

(B–C) Polysome profile of UBAP2L after **(B)** treatment of cells with 0.5 mM puromycin, and **(C)** treatment of lysates with 30 mM EDTA. (*Top*) Absorbance (at 260 nm) plot of a HEK293T cell lysate fractionated through a 10-50% a sucrose gradient. (*Bottom*) Western blots of UBAP2L from the corresponding fractions.

To identify specific transcripts subject to UBAP2L-mediated translational regulation, we performed polysome profiling in cell lysates from two independent UBAP2L knockout clonal isolates and from two control samples (**Figure 2.12A**). From two independent fractionations per line, we isolated polyA$^+$ mRNA from a portion of the input lysates and from pooled polysome fractions, and prepared and sequenced RNA-seq libraries. We considered all transcripts with RPKM ≥1 in inputs. We observed that UBAP2L knockout resulted in a larger number of transcripts with changes in pooled polysome fractions compared to changes in input RNA abundance. Most of the transcripts (82%) were down-regulated in pooled polysome fractions, but a similar fraction of transcripts were up- (55%) or down- (45%) regulated in input RNA, indicating the changes in pooled polysome fractions were independent of RNA abundance (**Figure 2.12B**). In the aggregate, these results suggest that UBAP2L predominantly acts at the translational level. As a measure of ribosome association, we computed the ratio of transcript RPKMs in polysome pools over input for all transcripts. We found a significant decrease ($p<10^{-300}$; Mann–Whitney *U* test, two-tailed) in mean transcript polysome-enrichment in both UBAP2L knockout lines compared to the controls (**Figure 2.12C**). Replicate analyses showed excellent correlation between the cell lines (**Figure 2.12D**). When we isolated those transcripts that changed in the same direction in both knockout lines, we found that overall nearly 10-fold more transcripts were reduced in translation (90.6%; n=8,784) than enhanced (9.4%; n=908) (**Figure 2.12E**). Even more strikingly, 97% of the 4,789 UBAP2L exon target transcripts, identified by eCLIP, showed significant downregulation in polysome association upon UBAP2L knockout (**Figure 2.12F**). We also measured a subset of target transcripts by qRT-PCR, which confirmed the magnitude of translational downregulation (**Figure 2.12G**).

To investigate how depletion of UBAP2L affected global translation, we evaluated the gene function attributes of UBAP2L direct targets. We observed a significant enrichment (FDR <0.05) in protein translation and ribosome biogenesis terms by Gene ontology (GO) analysis (**Figure 2.12H**). We also revealed that UBAP2L depletion decreased polysome association on mRNAs encoding translation initiation factors, elongation factors, and poly(A) binding proteins (**Figure 2.12I**). In fact, western blot analysis of these UBAP2L targets confirmed decreased protein levels of translation and elongation factors, such as Eukaryotic Translation Initiation Factor 4 Gamma 1 (EIF4G1), Eukaryotic Initiation Factor 3 Subunit B (EIF3B) DEAD-Box Helicase 54 (DDX54), and Eukaryotic Translation Elongation Factor 2 (EEF2) in cells lacking UBAP2L (**Figure 2.12J, K**). Taken together, these results suggest that UBAP2L enhances translation by directly binding mRNA substrates and by increasing translation of genes involved in global protein synthesis.

Figure 2. 12 UBAP2L promotes translation

(**A**) Polysome profiles of HEK293T cells (WT, n=2) and UBAP2L knockout HEK293T cells (KO, n=4) fractionated through 10-50% sucrose gradients. Light-colored lines indicate means from each set (WT, light blue; KO, pink), dark shaded areas (WT, blue; KO, red) denote standard deviations.

(**B-C**) Global transcript association with polysomes in UBAP2L knockout cells. (B) Scatter plots of log2-transformed RPKM ratios of input transcript levels (x-axis) and polysome transcript levels (y-axis) between the UBAP2L knockout HEK293T lines and WT samples. RPKM values from the two replicates were averaged prior to analysis and transcripts with average RPKM ≥1 were considered. Numbers and percentages of transcripts in each quadrant are indicated. (C) Cumulative distribution plots of log_2-transformed transcript levels (RPKM ≥1) in pooled polysome fractions from the two UBAP2L knockout HEK293T lines and WT control, normalized to levels in the respective input lysates. p-values were calculated using a two-sample Kolmogorov-Smirnov test.

(**D**) Scatter plots showing correlation of log_2-transformed ratios of input-normalized polysome transcript levels (RPKM) between the two UBAP2L knockout HEK293T lines. R, Pearson correlation coefficient.

(**E**) Bar graph showing log_2-transformed ratios of input-normalized polysome transcript levels (RPKM) between the two UBAP2L knockout lines (KO) and control (WT). Only transcripts with RPKM ≥1 in all three samples were considered (n=9,692). RPKM levels for the two KO lines were averaged.

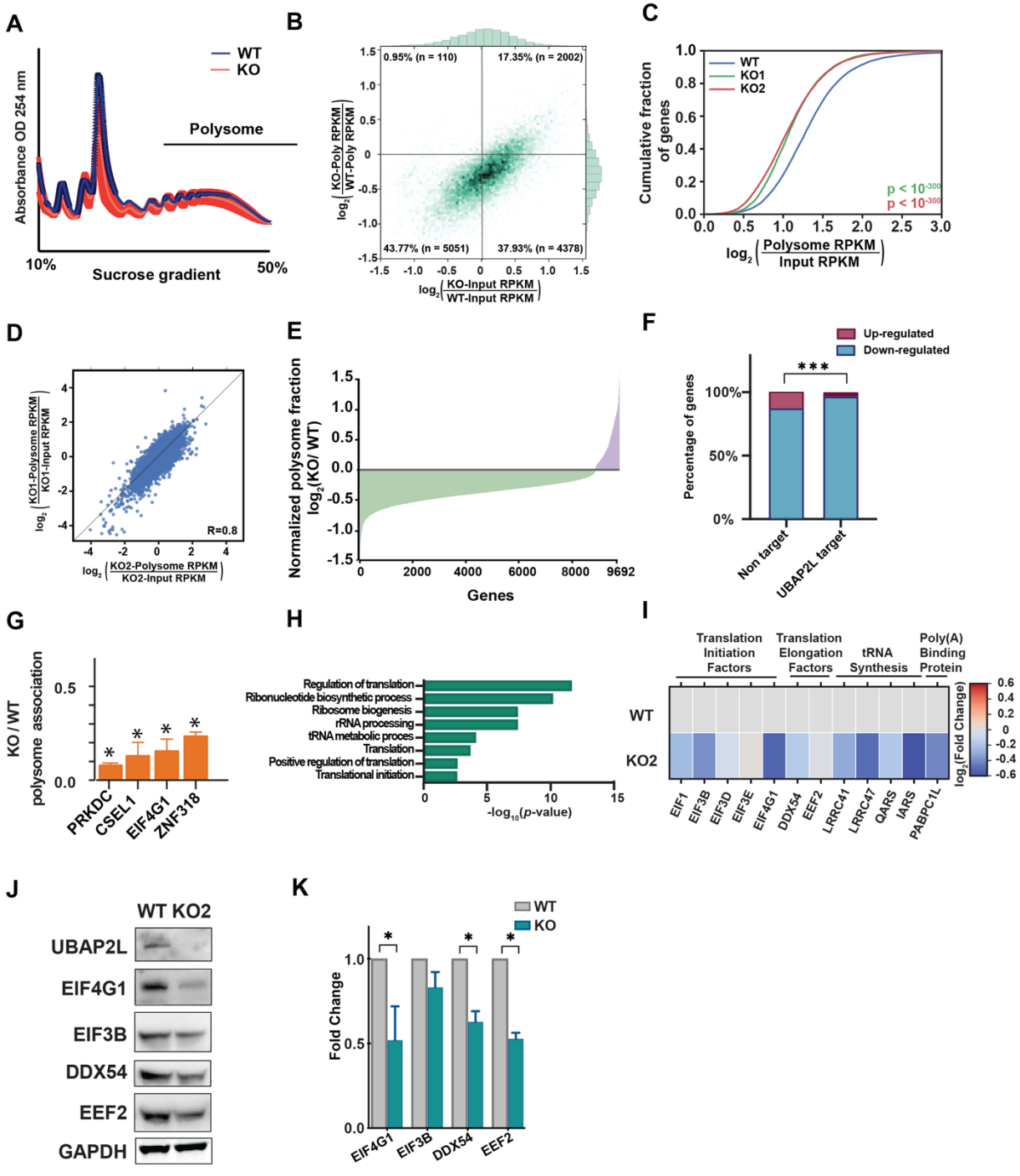
(**F**) Bar graph showing the percentage of regulated transcripts in UBAP2L targets, and non-targets. *p<0.0001 (χ^2 test with Yates' correction)

(**G**) Quantitative RT-PCR validation of reduced polysome association for the indicated transcripts. Transcript levels in inputs and polysome fractions were measured for KO and WT samples. KO/WT ratios of input-normalized polysome association of transcripts were then calculated.

(**H**) Gene ontology (GO) analysis for UBAP2L exon target transcripts (n=4,789). Significantly enriched GO terms were determined by Fisher's exact test at a false discovery rate of p<0.01. Shown are GO terms that are related to mRNA translation.

(**I**) Heat map showing log2-transformed polysome association ratio between UBAP2L knockout line (KO2) and control (WT) for the indicated translation regulators.

(**J-K**) (J) Representative western blots of UBAP2L, EIF4G1, EIF3B, DDX54, and EEF2 in UBAP2L knockout cells. GAPDH served as a loading control. (K) Densitometric quantitation of blots from of n=3 independent experiments. *p<0.05 vs. WT (two-tailed Student's t-test).


A
WT
KO
Polysome
Absorbance OD 254 nm
10%
Sucrose gradient
50%
B
0.95% (n = 110)
17.35% (n = 2002)
43.77% (n = 5051)
37.93% (n = 4378)
log2(KO-Poly RPKM / WT-Poly RPKM)
log2(KO-Input RPKM / WT-Input RPKM)
C
WT
KO1
KO2
Cumulative fraction of genes
p < 10-300
p < 10-300
log2(Polysome RPKM / Input RPKM)
D
R=0.8
log2(KO1-Polysome RPKM / KO1-Input RPKM)
log2(KO2-Polysome RPKM / KO2-Input RPKM)
E
Normalized polysome fraction log2(KO/ WT)
Genes
F
Up-regulated
Down-regulated
***
Percentage of genes
Non target
UBAP2L target
G
KO / WT polysome association
PRKDC
CSEL1
EIF4G1
ZNF318
H
Regulation of translation
Ribonucleotide biosynthetic process
Ribosome biogenesis
rRNA processing
tRNA metabolic proces
Translation
Positive regulation of translation
Translational initiation
-log10(p-value)
I
Translation Initiation Factors
Translation Elongation Factors
tRNA Synthesis
Poly(A) Binding Protein
WT
KO2
log2(Fold Change)
EIF1
EIF3B
EIF3D
EIF3E
EIF4G1
DDX54
EEF2
LRRC41
LRRC47
QARS
IARS
PABPC1L
J
WT KO2
UBAP2L
EIF4G1
EIF3B
DDX54
EEF2
GAPDH
K
WT
KO
Fold Change
EIF4G1
EIF3B
DDX54
EEF2

2.3.6 Programmable RNA-targeting CRISPR-mediated recruitment of UBAP2L promotes translation

In order to assess the dependence of UBAP2L-mediated translational regulation on direct binding to its target mRNA, we employed a fluorescence-activate cell sorting (FACS)-based reporter assay using UBAP2L fused to RNA-targeting RCas9 (RCas9)(61, 62) (**Figure 2.13A**). As a control, we performed our assay with RCas9-fused eukaryotic translation initiation factor 4E-binding protein 1 (EIF4EBP1, or 4EBP1), an inhibitor of translational initiation (**Figure 2.13B**). HEK293T cell lines expressing a fusion of RCas9 and UBAP2L, RCas9 and 4EBP1, or Cas9 only were derived via transposase-mediated piggyBAC genomic integration of plasmid constructs. We transfected a second, tripartite construct expressing a reporter that stably expresses RFP transcripts not regulated by RCas9, a guide RNA (gRNA), and a tetracycline-inducible YFP transcript harboring the guide RNA target sequences. We designed 7 different guide RNAs targeting locations across the YFP transcript (5' UTR, CDS, and 3'UTR), and a non-targeting guide RNA. We then measured post-transcriptional regulation as changes in the normalized YFP/RFP fluorescence ratio between Cas9-fusion and Cas9 only cells by using analytical flow cytometry. Due to the random nature of piggyBAC-mediated integration in terms of construct integration sites and numbers, regulation for various rCas9 construct levels (CFP) and reporter construct levels (RFP) can be quantified across thousands of data points (cells). With this highly sensitive and quantitative assay, we observed that the effect of UBAP2L on YFP reporter expression was dependent on UBAP2L directed to targeting sites within the 3'UTR and coding regions (**Figure 2.13C**). In contrast, significant 4EBP1-mediated reporter repression was only observed when 4EBP1 was targeted to the 5' UTR, as expected (**Figure 2.13D**). Normalized YFP mRNA levels were not significantly different

between RCas9-UBAP2L and RCas9 expressing cells transfected with gRNA 2 (which elicited the strongest increase with RCas9-UBAP2L), indicating that UBAP2L's positive effect on reporter expression was not due to upregulation of reporter mRNA (**Figure 2.13E**). Our UBAP2L-RCas9 results indicates a programmable means to enhance translation and further corroborate our observations from eCLIP and tethering in another orthogonal manner.

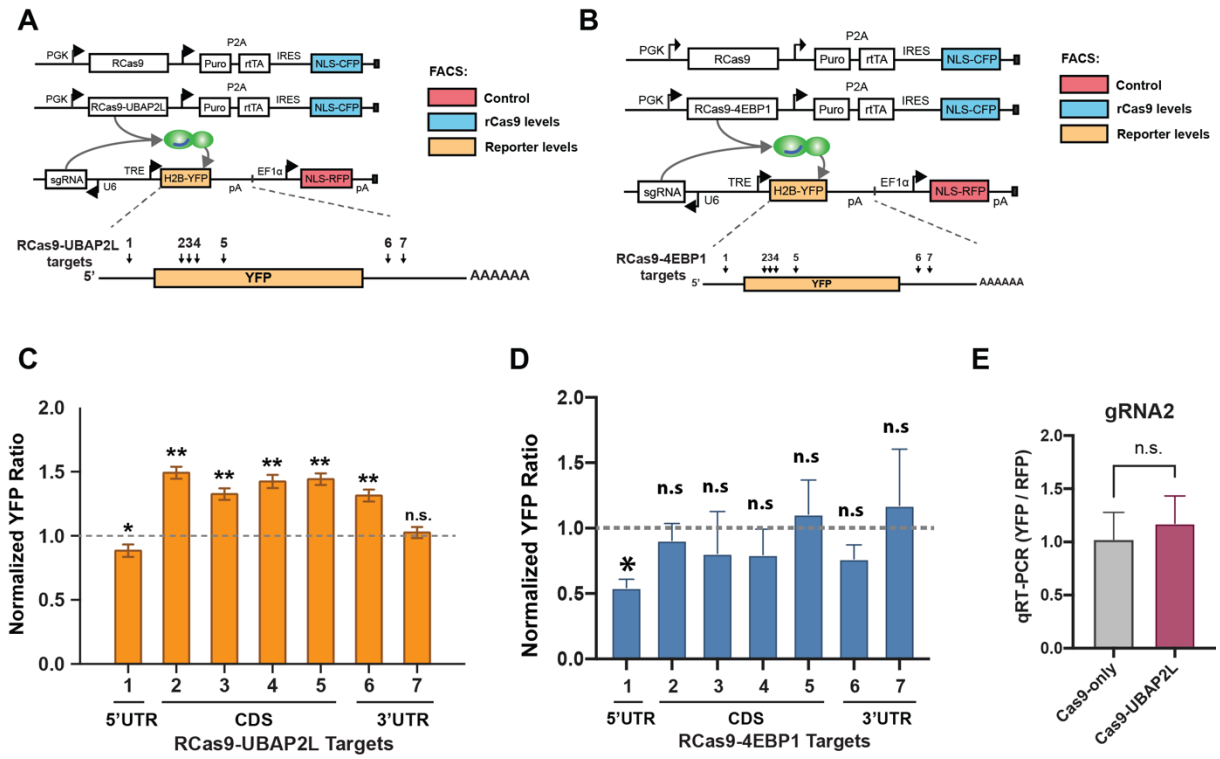


Figure 2. 13 Quantitative fluorescence-activated cell sorting (FACS)-based reporter assay

(A–B) Quantitative fluorescence-activated cell sorting (FACS)-based reporter assay for mRNA translation. Transgene expression constructs. RCas9 is expressed from a tetracycline responsive element (TRE) reporter. A constitutive promoter drives a polycistronic transcript containing puromycin *N*-acetyl transferase (Puro) and the reverse tetracycline (tet)-controlled transactivator (rtTA) separated by a P2A self-cleaving peptide, as well as CFP fused to a nuclear localization signal (NLS) preceded by an internal ribosome entry site (IRES). A second construct drives rCas9 fused to (A) UBAP2L (B) 4EBP1 in same plasmid backbone. rCas9 and rCas9-UBAP2L constructs were integrated into the genome at random copy number to establish stable cell lines. A third reporter construct harbors a U6 promoter driven single guide (sg)RNA targeting the indicated sites in the YFP reporter, which contains of a YFP fused to histone H2B driven by a tet-inducible promoter, and NLS-fused RFP driven by the EF1α promoter. The reporter construct was transiently transfected into rCas9 and rCas9-UBAP2L-expressing lines, and the expression levels of the three reporters were measured by FACS.

(**C-D**) Bar graph showing mean YFP levels in (C) rCas9-UBAP2L (D) rCas9-4EBP1 expressing cells, normalized to rCas9 expressing cells, on each targeting site. Error bars denote SD from n=2,000 rCas9-UBAP2L or rCas9-4EBP1 and n=2,000 rCas9 expressing cells per site. *p<0.05, **p<0.0001; n.s., not significant (p=0.7) (two-tailed Student's *t*-test).

(E) Bar graph showing ratios of YFP/RFP mRNA levels in rCas9-UBAP2L expressing cells, normalized to rCas9 expressing cells, in the presence of the gRNA targeting site 2. Transcript levels were measured by qRT-PCR and calculated with the $\Delta\Delta C_T$ method. Error bars denote mean±SD for n=3 replicates. n.s., not significant (*p*=0.5; two-tailed Student's *t*-test).

2.3.7 UBAP2L binds to RNA via the RGG domain and crosslinks to the expansion segments of the ribosome

To gain deeper molecular insight into the mechanisms by which UBAP2L enhances mRNA translation, we determined which protein domains mediate UBAP2L's interaction with RNA. UBAP2L is predicted to contain only two structured domains: a ubiquitin-associated (UBA) domain and an Arg-Gly-Gly repeat (RGG) domain, a common RNA and protein binding domain. Using inducible lentiviral vectors, we expressed UBAP2L, or truncated versions lacking either the UBA domain (ΔUBA), the RGG domain (ΔRGG), or both (**Figure 2.14A**), in UBAP2L knockout HEK293T cells. We then performed UV-crosslinking, immunoprecipitation, RNA fragmentation and radiolabeling to visualize RNA bound to UBAP2L (**Figure 2.14B**). Deletion of the RGG domain resulted in dramatically reduced recovery of RNA, indicating that the interaction between UBAP2L and RNA is mainly mediated by the RGG domain (**Figure 2.14C**).

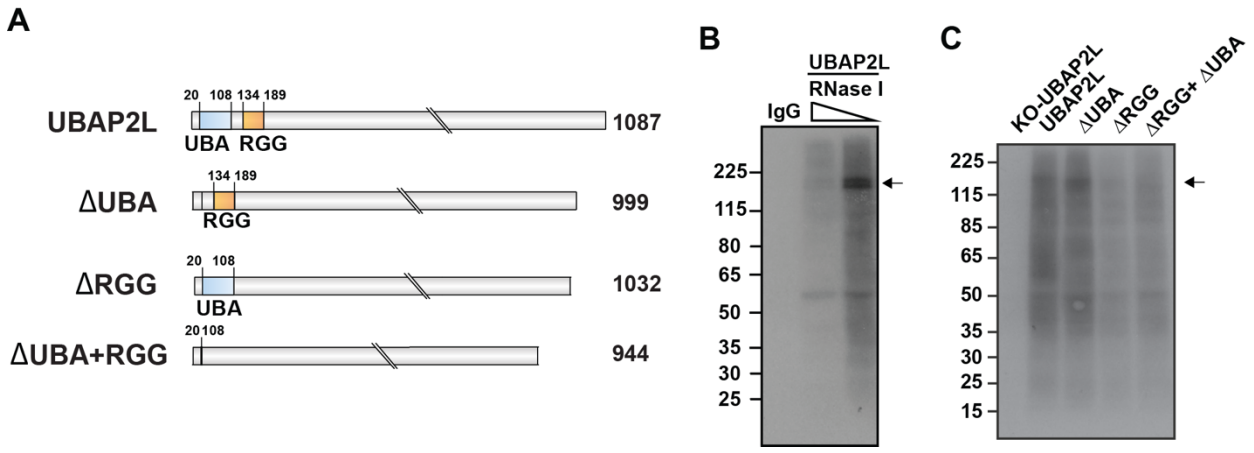


Figure 2. 14 UBAP2L binds RNA through RGG region

(**A**) Domain structures of UBAP2L constructs inducibly expressed in UBAP2L knockout HEK293T cells. The ubiquitin-associated domain (UBA; blue bar) and arginine-glycine-rich region (RGG; orange bar) are indicated.

(**B**) Autoradiograph of UBAP2L-RNA complexes immunoprecipitated from UV cross-linked HEK293T cells treated with increasing concentrations of RNase I, radiolabeled and separated on SDS polyacrylamide gel. Arrow indicates the expected molecular weight of UBAP2L.

(**C**) Autoradiograph of UBAP2L-RNA complexes immunoprecipitated from lysates of UV cross-linked UBAP2L knockout cells (KO-UBAP2L) expressing the indicated constructs, treated with RNase I, radiolabeled and separated on SDS polyacrylamide gel. Arrow indicates the expected molecular weight of UBAP2L.

Given that UBAP2L cofractionated with monosomes and polysomes in sucrose gradients, we reasoned that UBAP2L may interact directly with functional ribosomes. We first confirmed that UBAP2L is indeed localized to the cytoplasm (**Figure 2.15A**). We next examined two UBAP2L eCLIP datasets using a repeat-family centric mapping strategy, which maps reads to consensus transcripts from repetitive and recurrent genomic loci, including ribosomal RNA (rRNA) genes. Remarkably, reads from *rRNAs* constituted the largest fraction, with 72% in replicate 1 and 65% in replicate 2, while mRNA reads totaled 22% and 24%, respectively (**Figure 2.15B**). Closer inspection showed that reads were most highly enriched over SMInput at the expansion segments (ES) 15L, 27L of 28S rRNA, and E7S of 18S rRNA (**Figure 2.15C-G**), which are located at the solvent-exposed surface of ribosomes and are thought to engage with RBPs to modulate translation (63). As a further measure of the confidence of fold-enrichment, we utilized an information-theoretic metric, relative entropy, which scales each enrichment with the strength of evidence (i.e. read depth) at each peak (64). We confirmed that the peaks at ES15L, ES27L and ES7S (and an additional peak at ES31L) contained high information content (**Figure 2.15H**). In contrast, the mean of 446 other RBPs (64) shows very limited information content as a reflection of their limited specificity for binding the rRNAs. As a further confirmation, we performed RNA immunoprecipitation followed by RT-PCR on ES7S, ES7L, ES15L, and ES31L. Indeed, UBAP2L immunoprecipitated rRNA targets, indicating UBAP2L interacts with ribosomes on expansion segments (**Figure 2.15I**). This is consistent with a recent UBAP2L IP-mass spectrometry study that recovered peptides from 15 ribosomal proteins (65), further supporting a UBAP2L-ribosome interaction.

To assess the spatial arrangement of UBAP2L and the ribosome, we mapped these interactions onto the cryo-electron microscopy structure of the mammalian ribosome (66). The top ribosomal proteins that co-immunoprecipitate with UBAP2L (65) cluster in the 60S subunit (**Figure 2.15J**). In addition, ES31L, which is highly enriched for UBAP2L binding, lies close to the region of the 60S subunit, which is occupied by tRNA in the exit site (E site) during protein synthesis (**Figure 2.15K**). Collectively, these data support a model in which UBAP2L's function is associated with its interactions with the ribosome.

Figure 2. 15 UBAP2L binds directly to the ribosome

(A) Immunofluorescence images showing UBAP2L (green) in HEK293T cells. DAPI (blue) marks nuclei. Scale bar, 10 μm.

(**B**) Pie chart showing fractions of two replicates of UBAP2L eCLIP reads from HEK293T cells unambiguously mapping to mRNAs, ribosomal RNAs, and other repeat families.

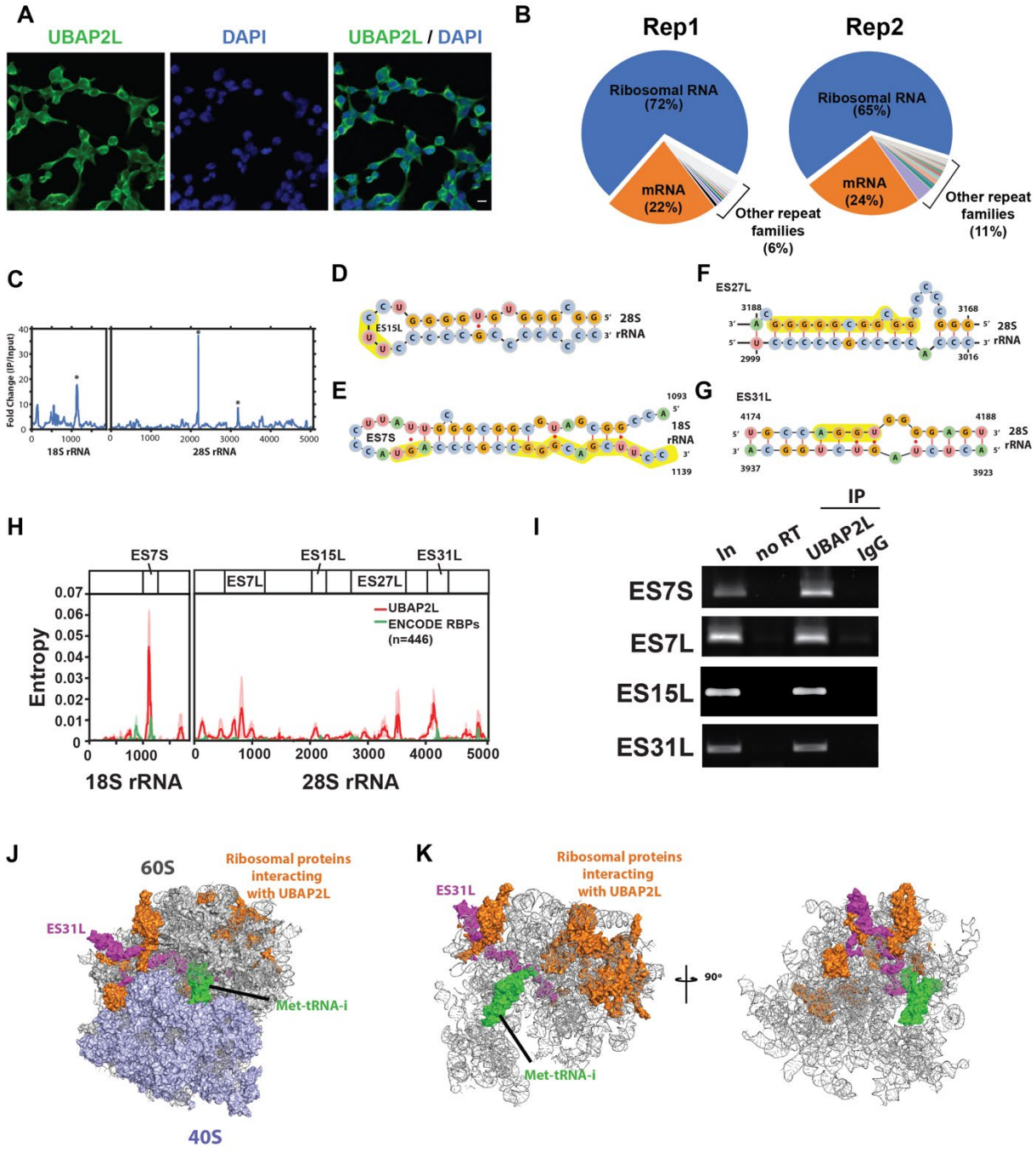
(C) Line plot of UBAP2L binding sites on rRNAs. Fold-enrichment of reads for IP over SMInput is plotted against the nucleotide positions of 18S and 28S rRNAs. * denotes relative entropy ≥0.01.

(D–G) Location of UBAP2L binding sites on rRNA. (D) ES15L, (E) ES7S, (F) ES27L, and (G) ES31L. Nucleotides with significant binding are highlighted in yellow.

(**H**) Locations of UBAP2L binding sites on rRNAs. Line plots showing the Kullback-Leibler divergence (relative entropy) for UBAP2L in HEK293T cells (red line) and the mean of 446 other RBPs analyzed by the ENCODE consortium (green line)(64) on 18S and 28S rRNAs. Lines show the mean of relative entropy, with light shaded areas indicating 10%-90% confidence intervals.

(I) RNA immunoprecipitation (RIP) of UBAP2L-RIP and RT-PCR in HEK293T cell lysates. The RIP assay was performed using anti-UBAP2L antibody or rabbit non-immune IgG. RT-PCR was performed on using primer sets within UBAP2L target regions ES7S, ES7L, ES15L, and ES31L.

(**J-K**) Model of the interactions of UBAP2L on the human ribosome structure(66). (J) Surface view with 60S ribosomal subunits (RNA and protein) shaded in gray and lavender, respectively. The Met-tRNA is highlighted in green. Ribosomal proteins previously identified as UBAP2L interactors by immunoprecipitation and mass spectrometry(65) are highlighted in orange. The expansion segment ES31L is highlighted in magenta. (g) View as in (J) with non-highlighted proteins removed. (K, *right*) View as in (K, left) rotated 90° around the *z*-axis.


A
UBAP2L
DAPI
UBAP2L / DAPI
B
Rep1
Rep2
Ribosomal RNA (72%)
mRNA (22%)
Other repeat families (6%)
Ribosomal RNA (65%)
mRNA (24%)
Other repeat families (11%)
C
Fold Change (IP/Input)
18S rRNA
28S rRNA
D
ES15L
28S rRNA
E
ES7S
18S rRNA
F
ES27L
28S rRNA
G
ES31L
28S rRNA
H
ES7S
ES15L
ES31L
ES7L
ES27L
UBAP2L
ENCODE RBPs (n=446)
Entropy
18S rRNA
28S rRNA
I
IP
In
no RT
UBAP2L
IgG
ES7S
ES7L
ES15L
ES31L
J
60S
Ribosomal proteins interacting with UBAP2L
ES31L
Met-tRNA-i
40S
K
Ribosomal proteins interacting with UBAP2L
ES31L
Met-tRNA-i
90°

2.4 Discussion

Tethering of RBPs, followed by genome-wide validation, reveals new candidate regulators and rules

To assign molecular functions to the rapidly growing number of predicted RNA binding proteins, we have developed comprehensive resources and methods enabling large-scale RBP tethered function assays. Using two pairs of 3'UTR reporters interrogating nearly 700 RBPs, we discovered at least 50 RBPs with significant positive and negative effects on both reporters. Extrapolating to the over 2,000 RBPs encoded in the human genome (10-15), we speculate that over a hundred may have yet unrecognized roles in regulating RNA metabolism by either mRNA stability and/or translation.

Highlighting the practicality of our approach, our results point to previously unknown mRNA targets and novel molecular mechanisms for several RBPs, all of which warrant future investigation. To illustrate, SNRPA (or U1A) is a dual-function protein functioning as a component of the spliceosomal U1 snRNP essential for the recognition of the 5' splice site, and, in a snRNP-free form, couples splicing to polyadenylation (67). In its latter role, SNRPA has been shown to interact with stem-loop structures in 3'UTRs called polyadenylation inhibitory elements (PIEs) to inhibit poly(A) polymerase (68). Consistent with this, SNRPA emerged as an RNA destabilizer in our tethering assay, its overexpression led to significant downregulation of its targets, and SNRPA binding was enriched, expectedly, in intronic regions (59% of significant peaks) and also 3'UTRs (7.3%), including the PIE in its own transcript. Prior to our study, only a handful of SNRPA 3'UTRs had been identified. We identified 552 of peaks in 344 SNRPA 3'UTR targets, suggesting that this mode of regulation is surprisingly widespread.

In the future, it will be interesting to evaluate if these also contain PIE structures or if SNRPA recognizes other sequence or structural elements.

IFIT2 is an interferon-induced protein that blocks translation via sequestration of eIF3 which is essential for cap-dependent translational initiation. In conjunction with other IFIT family members, which act by binding viral cap structures and sequestering viral proteins and RNAs, IFIT2 is an important element of the antiviral response. Although silencing and overexpression studies have suggested additional roles for IFIT2 in modulating the host response to viral infection, the molecular mechanisms remain unclear (69). We find that IFIT2 is highly enriched for binding at the 3'UTR of endogenous RNAs (2538 of peaks in 2,100 genes) but only regulates stability of a minor fraction of them (14.5%). In our tethering assay, IFIT2 recruitment to the 3'UTR enhanced reporter translation without significantly affecting reporter mRNA stability. We speculate that IFIT2 may mainly act at the translational level and perhaps regulates the host antiviral response by increasing translation of relevant endogenous RNAs.

AIMP1 is multifunctional protein acting as an auxiliary factor of aminoacyl-tRNA synthetase complexes and as an intracellular and extracellular signaling molecule promoting inflammation and suppressing tumorigenesis. However, a role for this protein in mRNA stability and translation has not been described (70). Surprisingly, we find that AIMP1 binds a large number of endogenous mRNAs (4241 peaks in 2,206 genes), of which approximately equal numbers are up- and downregulated. Our results point to another facet of the multiple mechanisms and pathways regulated by AIMP1.

Lastly, MTDH is a key oncogene promoting metastasis and resistance to chemotherapy, acting through multiple mechanisms. In the nucleus, MTDH is thought to activate the transcription factor NFκB, thereby promoting transcription of NFκB target genes. In the

cytoplasm, MTDH is thought to promote drug resistance by increasing polysome association of the multidrug resistance gene 1 (MDR1) mRNA (71). Interestingly, MTDH has been shown to prevent stress granule formation and promote global protein translation via regulation of eIF4G phosphorylation, required for translational initiation (72). In agreement with these studies, our tethering assay showed that MTDH promotes translation to a larger extent than mRNA stability. MTDH knockdown led to changes in levels of only 90 transcripts, consistent with a predominant role as a translational modulator.

Future studies using a battery of genome-wide assays such as transcription inhibitors, metabolic labeling, nuclear/cytoplasm export and subcellular localization will be informative to further detail the mechanisms underlying how these candidate stabilizers and destabilizers function.

UBAP2L: a new global translation factor

Our studies revealed UBAP2L as a novel translational regulator that acts both locally on direct targets, with also global consequences. We present evidence pointing to UBAP2L as an RNA binding protein that interacts with RNA via its RGG domain, and crosslinking analysis suggests that UBAP2L interacts with the expansion segments of the ribosome. We observe a limited number of sites on rRNA to which UBAP2L crosslinks, supporting a specific set of interactions. Intriguingly, all interactions map to the eukaryote-specific expansion segments, flexible RNA structures protruding from the surface of the ribosome that are not fully occupied by ribosomal proteins. A recent report showed that binding of methionine amino peptidase to expansion segment 27L controls translational fidelity (63). UBAP2L thus joins a likely growing number of ribosome-associated proteins that modulate translation through binding to these

rRNA scaffolds. Association of UBAP2L with the ribosome is consistent with our eCLIP observations that UBAP2L binds to the coding regions of thousands of transcripts.

Our results indicate that UBAP2L acts to enhance translation, as deletion of UBAP2L in human cells leads to a reduction of polysome association of thousands of target transcripts. This translational enhancement activity is transferable to a target mRNA via RNA-targeting Cas9 fused to UBAP2L, indicating UBAP2L recruitment to a mRNA substrate is sufficient to increase its translation. Indeed, when UBAP2L is recruited to the CDS of the reporter via RCas9, UBAP2L increases reporter translation by ~35-50%, depending on the location within the coding region (or 3'UTR). While moderate, this magnitude is in agreement with the magnitude of the decrease in protein synthesis (~40%) and in transcript polysome association of UBAP2L targets upon UBAP2L knockdown (median, 31%). Incidentally, this orthogonal approach also further expands *in vivo* applications of RNA-targeting CRISPR/Cas (RCas9) (61, 62, 73).

Finally, our transcriptome-wide analyses reveal that UBAP2L affects a significant number of mRNA targets. Interestingly, mRNAs targeted by UBAP2L are themselves enriched for central regulators of translation, and protein synthesis, revealing a role for UBAP2L in modulating protein homeostasis globally through directly and indirect means. Our current working model proposes that UBAP2L is dynamically recruited to translating ribosome-mRNP complexes to enhance translation on many targets, including translational regulators to affect global protein synthesis. Future studies will be needed to further evaluate the structural basis of UBAP2L-RNA interactions.

Considerations of tethered function assay screens

Even though our screen to identify new regulators of mRNA, and in the case of UBAP2L, protein levels is generally successful, we do caution that tethered function assays, as with all large-scale screens, are designed to prioritize positively-scoring candidates quickly but false negatives and positives will undoubtedly occur. The sensitivity of such screens depends on the specific reporters and cell-lines chosen. For instance, our reporter mRNA is subject to endogenous polyA tailing which recruits polyA binding proteins (PABPs). Hence, additional tethering of PABPs is unexpected to further increase mRNA translation. Additionally, productive RBP recruitment to the reporter may fail for a number of reasons(23, 24, 74). Some RBPs require RNA binding at a specific site within the transcript region or in a specific orientation, or the ability to transit along or cycle to and from their RNA substrate. RBPs may also require additional cofactors for activity that are not expressed in the cell line utilized for the assay. Also, it is hard to predict if the fused RBP will function similarly to the cognate RBP. Thus, the absence of an effect in the assay is generally difficult to interpret. Conversely, forced recruitment may lead to artifactual signals with little relevance for endogenous RBP recruitment to natural RNA substrates. For example, binding of a protein per se, independent of its function, may block access or movement of RNA processing machinery.

Despite these caveats inherent in such screens, our large-scale approach and our follow-up of over a dozen candidate RBPs for their endogenous roles allow us to systematically address some of these questions for the first time. We find that the vast majority of RBPs resulted in reporter activity changes in the same direction (up- or downregulation) for both reporters we used, suggesting that for RBP recruitment to 3'UTRs, reporter context dependence may be more limited than we expect. We also observed that RBP size had no significant correlation with the magnitude of an effect, suggesting that overall, steric hindrance may represent but a minor

confound in large-scale 3'UTR tethering assays. Lastly, analysis of the effects of depletion and ectopic expression of 14 RBPs on the stability of their endogenous RNA targets largely recapitulated RBP effects from the tethering assay, but also revealed additional roles.

Overall and in conclusion, our results provide proof-of-principle for the utility of large-scale 3'UTR tethering assays for identification of candidate RBPs involved in affecting mRNA, and with deeper validation, protein levels. We anticipate that our RBP-MCP fusion library and screening methods will now enable massively parallel tethered function assays aimed at elucidating the roles of RBPs in other RNA metabolic processes, such as splicing, transport and A-to-I editing. In light of new experimental and computational approaches that have unearthed hundreds of candidate novel RBPs (7, 8, 75) and non-canonical RNA binding domains (76), we speculate that high-throughput approaches such as ours will be of increasingly critical importance to assign molecular functions to these newly identified bona fide RBPs.

Future perspectives

Our results provide a brief picture of RNA stability and translation. According to our eCLIP data, RBPs have their own binding preferences and specific subset of target genes, indicating they regulate RNA stability and translation in different manners. For future following work, we should dissect the mechanism of 50 candidate RBPs individually, and investigate which protein complex that the RBP interacts with, and find out how and when the complex recognizes its target genes. Since mRAN and protein levels are dynamic balance in the cells, disruption of the balance causing cancer or neuron degeneration diseases. Therefore finding out how the protein complexes coordinate together to maintain homeostasis in cells is a crucial question.

For UBAP2L regulation, our results show UBAP2L associate with ribosome, and enhance translation. It is important to find out how UBAP2L regulates translation on ribosome? To answer this question, it is necessary to have UBAP2L crystal structure on the ribosome. This could tell whether UBAP2L promotes conformation changes in ribosome and then enhances translation. It is also important to find out whether UBAP2L involves in other pathways and have multiple roles in the cells. Since UBAP2L has a UBA and RGG domains, UBA domain might play a crucial role in ubiquitination and protein degradation, which could also regulate protein expression levels.

2.5 Materials and Methods

2.5.1 Generation of MS2 coat protein tagged RBP expression plasmids

The majority of ORF clones were obtained in pENTR vectors from the CCSB human ORFeome collection (77) (Dana Farber Cancer Institute) or the DNASU Plasmid Repository (Arizona State University). Some ORFs were purchased in standard expression clones, amplified by PCR (Phusion polymerase, NEB) with oligonucleotide primers containing attB recombination sites, and recombined into pDONR221 using BP clonase II (Thermo Fisher). ORFs were then recombined into a custom pEF DEST51 destination vector (Thermo Fisher) engineered to direct expression of the ORFs as fusion proteins with a V5 epitope tag and the MS2 coat protein appended C-terminally and under the control of the EF1-alpha promoter, to create ORF-V5-MS2BP constructs. The identity of all cDNA clones was verified by Sanger sequencing. **Extended Data Table 1** lists all ORFs and relevant information.

2.5.2 Gene ontology analysis

Panther was used for gene ontology (GO) analysis(78, 79). For library RBPs, the following GO terms related to RNA processing were used: Splicing ('RNA splicing'), Stability (' RNA stabilize', ' RNA stability', ' RNA stabilization', ' RNA decay', ' RNA turnover', ' RNA deadenylation', ' RNA cleavage', ' RNA cleaving', ' RNA degradation'), Translation ('translation'), Localization (' RNA localization', ' RNA transport', ' RNA localize', ' RNA export'), and Modification ('RNA methylation', 'RNA modification'). Significant GO terms were determined by Fisher's Exact Test after FDR at $p<0.01$ and sorted by fold enrichment. For GO analysis of UBAP2L-regulated genes, significantly enriched GO terms were determined by Fisher's exact test after FDR at $p<0.01$ and sorted by fold enrichment.

2.5.3 Dot blots

HeLa cells were grown in antibiotic-free DMEM (Life Technologies) media with 10% FBS. 50 ng plasmid was prepared for transfection using Lipofectamine 3000 reagent (Thermo Fisher). After 15 minutes incubation, plasmid was transferred to 96-well tissue culture plates coated with poly-D-lysine hydrobromide (Sigma-Aldrich). 3.5×10^5 cells were subsequently plated in each well. After 48 h, cells were washed with phosphate-buffered saline (PBS) and lysed in lysis buffer (50 mM tris-HCl, 100 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate; pH 7.4) with protease inhibitor cocktail III (EMD Millipore). Lysate was transferred to nitrocellulose membrane pre-wetted with PBS using the Bio-Dot Microfiltration Apparatus (Bio-Rad) following manufacturer instructions. Membranes were blocked in blocking buffer [tris-buffered saline (TBS) containing 5% (w/v) dry milk powder] for 30 min,

then probed with rabbit anti-V5 tag antibody (Bethyl) in blocking buffer for 16 h at 4 °C. Membranes were washed 3 times with TBS and probed with secondary HRP-conjugated antibody in blocking buffer for 1 h at room temperature. Signal was detected by Pierce ECL Substrate (Thermo Fisher) and collected using the Azure c600 Imager (Azure Biosystems). Dot intensity was calculated using the ImageJ Gel Analyzer, treating each column of the blot as a gel lane. Fold-change was calculated for each sample over the highest intensity negative control well from the membrane containing the sample. Positive detection was called for fold-changes >1.0.

2.5.4 Generation of luciferase reporter assay constructs

Standard restriction enzyme cloning was used to generate reporter constructs directing expression of firefly (FLuc) or *Renilla* (RLuc) luciferase fused to protein destabilizing domains CP1 and PEST (DD), with a 3'UTR consisting of that from the human β-globin (HBB) 3'UTR, under control of the tetracycline response element promoter (pTET2). Six MS2 hairpin structures(80) were inserted into the 3'UTR to generate FLuc-6MS and RLuc-6MS by standard restriction enzyme cloning. All constructs were sequence-verified.

2.5.5 Luciferase reporter screen

For time course analyses, Tet-Off Advanced HeLa cells (Clontech) were grown in DMEM (Thermo Fisher) with tetracycline-free FBS (10%; Clontech), Penicillin-Streptomycin (1×; Thermo Fisher), and G418 (100 µg/ml; Corning). Prior to transfection, G418 was removed. A 6:1:1 mix of RBP-MCP, firefly-MS2 (or *Renilla*-MS2), and *Renilla* (or firefly) luciferase

reporter (transfection control) constructs were diluted in 150 mM NaCl and mixed for transfection with polyethyleneimine (Polysciences) at a ratio of 1 µg DNA to 4 µg PEI. Cells were transfected at 50-60% cellular confluency, with a total of 125 ng and 250 ng DNA for 48-well plates and 24-well plates, respectively, and grown in the absence of G418. 48 h post-transfection, reporter transcription was suppressed by the addition of tetracycline (1 µg/ml; Sigma). Cells were lysed after 20, 80 and 120 min, lysed and luciferase activity measured with the Dual-Luciferase Reporter Assay System (Promega), following the manufacturer's directions, in a microplate reader. Values were expressed as the ratio of the mean luciferase activity of MS2-tagged over untagged reporters from 3 replicates. For the screen and validations, transfections were done as for the time course assay, and luciferase activities were measured 48 h post-transfection. **Extended Data Table 2** lists the results of the luciferase assays.

2.5.6 qRT-PCR

Total RNA was isolated by lysing cells in TRIzol and purification with Direct-zol RNA Kits, following the manufacturer's protocols. 0.5-1 µg of total RNA was reverse transcribed using Superscript III with *oligo(dT)*$_{12\text{-}18}$ primers (Thermo Fisher). cDNA was diluted 20-fold in water and target transcripts quantified with Power SYBR Green Master Mix (Thermo Fisher). Three biological replicate samples were analyzed, and qRT-PCR was carried out in three technical triplicates. Mean C_t values were calculated from each triplicate set. Biological replicates were averaged to generate mean fold changes and values expressed as a fold

differences to control samples calculated using the $\Delta\Delta C_t$ method. Significance was assessed by a two-tailed Student's *t*-test.

2.5.7 eCLIP library preparation and sequencing

eCLIP was performed essentially as described(1). Briefly, for each RBP, 3×10^7 HEK293T cells were UV-crosslinked (400 mJ/cm^2, 254 nm) and lysed. Lysates were sonicated and treated with RNase I to fragment RNA. 2% of each lysate sample was stored for preparation of a parallel size-matched input (SMInput) library. The remaining lysates were immunoprecipitated using RBP-specific antibodies. Bound RNA fragments in the IPs were dephosphorylated and 3'-end ligated to an RNA adapter. Protein-RNA complexes from SMInputs and immunoprecipitates (IPs) were run on an SDS polyacrylamide gel and transferred to nitrocellulose membrane. Membrane regions comprising the exact RBP sizes to 75 kDa above were excised and RNA released from the complexes with proteinase K. SMInput samples were dephosphorylated and 3'-end ligated to an RNA adapter. All RNA samples (IPs and SMInputs) were reverse transcribed with AffinityScript (Agilent). cDNAs were 5'-end ligated to a DNA adaptor. cDNA yields were quantified by qPCR and 100-500 fmol of libraries generated with Q5 PCR mix (NEB).

2.5.8 Computational analysis of eCLIP sequencing data

Reads were processed essentially as described(1). Briefly, reads were adapter-trimmed and mapped to human-specific repetitive elements from RepBase (version 18.05) by STAR (81). Repeat-mapping reads were removed, and remaining reads mapped to the human genome

assembly hg19 with STAR. PCR duplicate reads were removed using the unique molecular identifier (UMI) sequences in the 5' adapter and remaining reads retained as 'usable reads'. Peaks were called on the usable reads by CLIPper (55) and assigned to gene regions annotated in GENCODE (v19) with the following descending priority order: CDS, 5'UTR, 3'UTR, proximal intron, and distal intron. Proximal intron regions are defined as extending up to 500 bp from an exon-intron junction. Each peak was normalized to the size-matched input (SMInput) by calculating the fraction of the number of usable reads from immunoprecipitation to that of the usable reads from the SMInput. Peaks were deemed significant at ≥4-fold enrichment and $p \leq 10^{-5}$ (χ^2 test, or Fisher's exact test if the observed or expected read number in eCLIP or SMInput was below 5). Reproducible clusters were defined as clusters that overlapped in both replicates. Target transcripts were defined as transcripts that contain at least 1 significant reproducible cluster. Code is available on GitHub (https://github.com/YeoLab/eclip).

2.5.9 eCLIP region-based fold-enrichment analyses

Region-based fold-enrichment was calculated as described previously(82). Briefly, usable reads were counted across regions for all annotated transcripts in Gencode v19 (comprehensive). Possible regions included CDS, 5'UTR, and 3'UTR. For each gene, a read was first queried for overlap with coding exons; then with 5'UTR or 3'UTR exons. Reads were then summed across all exons for the gene to obtain final region counts, and a pseudocount of 1 was added to classes for which no reads were observed. Read counts were normalized by the total number of usable reads (RPM normalization). Only regions with at least 10 reads in one

of IP or SMInput, and where at least 10 reads would be expected in the opposite dataset given the total number of usable reads, were considered. The fold enrichment was calculated as the ratio of normalized read counts in IP over SMInput.

2.5.10 Repeat-family centric mapping

Binding to ribosomal rRNA was quantified using a family-aware repeat element mapping pipeline (19). Briefly, reads were mapped to a database of 7,419 multi-copy element transcripts, including the 5S, 5.8S, 18S, and 28S rRNA as well as tRNAs, retrotransposable elements, and numerous other RNAs. Reads mapping to multiple element families were not considered for further analysis. To summarize relative enrichment between IP and input, relative information was defined as the Kullback-Leibler divergence (relative entropy): $p_i \times log_2(\frac{p_i}{q_i})$, where p_i is the fraction of total reads in IP that map to a queried repetitive element i, and q_i is the fraction of total reads in input for the same element. Code is available on GitHub (https://github.com/YeoLab/repetitive-element-mapping)

2.5.11 Meta-gene mapping analyses

Metagene plots were created using the intersection of eCLIP peaks and a set of mRNA regions. To generate the list of each CDS, 5'UTR, and 3'UTR, non-overlapping gene annotations from GENCODE v19 were used. First, lowly expressed transcripts (TPM < 1) were filtered. Then, transcripts with the highest TPM were selected, resulting in a single transcript per gene in the CDS. For each 5'UTR, CDS, and 3'UTR in a gene, the entire set of exons comprising the region was concatenated and overlapped with CLIP peaks, resulting in a list the length of the spliced transcript containing values of 1 if a peak was found at a given position,

or 0 otherwise. Plotted lines represent the number of total peaks found at each position divided by the total number of unique transcripts. The length of each region within the metagene was then scaled to 8%, 62%, and 30%, corresponding to the average length of regions from the highest expressed transcripts in ENCODE HepG2 RNASeq control datasets (64). The peak density was calculated by percentage of peak number at given position. (https://github.com/YeoLab/rbp-maps).

2.5.12 *De novo* motif analysis

HOMER was used to identify *de novo* motifs using the command '*findMotifsGenome.pl* <foreground> hg19 <output location> -rna -S 20 -len 6 -p 4 -bg <background>' Foreground was a bed file of significant peaks; the background was randomly defined peaks within the same annotated region as foreground peaks. Code is available on GitHub (https://github.com/YeoLab/clip_analysis_legacy).

2.5.13 eCLIP correlation analysis

We utilized Pearson correlation statistics to measure the reproducibility between pair-wise comparisons of replicate eCLIP experiments. The read density in peaks was normalized to the size-matched input (SMInput) by calculating the fraction of the number of usable reads from immunoprecipitation to that of the usable reads from the SMInput. The correlation was the comparison of fold-enrichment in both datasets for all peaks.

2.5.14 Lentiviral shRNA knockdowns and transient plasmid transfections

To generate lentiviral particles for RBP knockdown, 3.8×10^6 HEK293XT cells were seeded in 10 cm plates in antibiotic-free DMEM (Gibco) media with 10% FBS. After 24 h, cells were transfected with sequence-verified shRNA plasmids (pLKO.1) and packaging plasmids (pMD2.G: Addgene #12259 and psPAX2: Addgene #12260; both gifts from Didier Trono) using Lipofectamine 3000 (Thermo Fisher). Virus-containing media was harvested, replaced with 15 mL 20% FBS media, and harvested again a further 24 h later. Virus-containing media was pooled. For lentiviral transduction of HEK293T cells, cells were seeded in 6-well plates at 8×10^5 cells/well and grown for 16 h in DMEM media with 10% FBS. Cells were transduced with virus-containing media diluted 1:1 in fresh media. After 24 h, media was replaced with fresh media containing 2 µg/µl puromycin. After a further 72 h, cells were harvested and analyzed for shRNA knockdown efficiency by western blot and RT-qPCR, and for RNA-seq analysis.

For RBP overexpression, 3.8×10^6 HEK293XT cells were seeded in 10 cm plates in antibiotic-free DMEM (Gibco) media with 10% FBS at 37 °C. After 24 h, cells were transfected with 24 ug RBP plasmids using Lipofectamine 3000 (Thermo Fisher). After a further 48 h, cells were harvested and analyzed by western blot for successful overexpression, by RNA-seq, and by eCLIP analysis as indicated.

2.5.15 RNA-seq library preparation and analysis

RNA was extracted from cells with TRIzol (Invitrogen). Strand-specific RNA sequencing libraries were prepared from 0.5-3 µg total RNA using the Illumina TruSeq Stranded mRNA Sample Preparation kit (Illumina). Libraries were sequenced on the Illumina

HiSeq 4000 platform at a depth of at least 12×10^6 reads per sample in SE50 mode. RNA-sequencing reads were trimmed using *cutadapt* (v1.4.0) of adaptor sequences and mapped to repetitive elements (RepBase v18.04) using the STAR (v2.4.0i). Reads that did not map to repetitive elements were then mapped to the human genome (hg19). GENCODE (v19) gene annotations and featureCounts (v.1.5.0) were used to create read count matrices. Differential expression was calculated using DESeq2 version 1.10.1 (ref. (83)), individually pairing each knockdown or overexpression experiments with their respective controls. Genes with RPKM values lower than 1 were not used.

2.5.16 Polysome profiling

For lysate preparation, 4×10^7 HEK293T cells were prepared. Before collection, cells were treated with cycloheximide (CHX) at 100 µg/ml for 5 min at 37°C. Culture media was removed, cells were washed twice with cold PBS containing 100 µg/ml CHX (PBS-CHX), resuspended in PBS-CHX by centrifugation at 200×*g* at 4 °C for 5 min, collected in PBS-CHX and snap frozen in liquid nitrogen. Cells were lysed by trituration through a 27 gauge needle in 400 µl polysome lysis buffer (20 mM Tris-HCl pH 7.4, 150 mM NaCl, 5mM $MgCl_2$) with 1× protease inhibitor cocktail (EMD Millipore), CHX 100 µg/ml, 1 mM DTT, 25 U/ml DNase (TurboDNase; Thermo Fisher), 20 U/ml RNase inhibitor (RNaseOUT, Thermo Fisher) and incubation on ice for 30 min. Lysates were clarified by centrifugation at 17,500×*g* at 4 °C for 5 min. 50 µl was reserved for inputs and the remainder used for fractionation.

For puromycin treatment to release polysomes, cells were treated with puromycin at 0.5 mM for 40 min, and then treated with CHX at 100 µg/ml for 5 min. For disassembly of

monosomes into 40S and 60S ribosomal subunits, the polysome lysis buffer was supplemented with 30 mM EDTA.

For fractionation, a 14 ml 10-50% (w/v) sucrose gradient was prepared in polysome buffer. Samples were loaded on the sucrose gradient and centrifuged in a swinging bucket rotor at 35,000×*g* at 4 °C for 3 h. Fractions were collected from the top and UV absorbance monitored using a Gradient Station (BioCamp) equipped with ECONO UV monitor (BioRad). Fractions (500 µl each) were collected using a FC203B (Gilson) fraction collector. Fractions containing polysomes were pooled. Total RNA from the inputs and polysome pools were extracted in TRIzol-LS (Thermo Fisher) and purified with Direct-zol RNA kits (Zymo). RNA sequencing libraries were generated sequenced and reads processed as described above.

For analysis of fractions by western blotting, 2 µg BSA was added to each fraction and protein precipitated by addition of trichloroacetic acid was added to 20% (v/v). Protein was precipitated for 16 h at 4 °C and collected by centrifugation at 15,000×*g* for 20 min at 4 °C. Protein pellets were washed twice with 1 ml ice-cold acetone, centrifuged at 15,000×*g* for 15 min at 4 °C, dried at room temperature, neutralized, re-suspended, and denatured by incubation in 50 mM tris-HCl pH 7.4 at 65 °C for 30 min and 98 °C for 15 min. Western blotting was performed as described above.

2.5.17 Polysome association analysis

The transcript RPKMs of input and polysome fractions were calculated from the read count matrices. Only genes with RPKM ≥1 were considered. Polysome association was measured by calculating the RPKM ratio of transcript levels in polysomes over input. Polysome

association ratios between depletion samples and their respective controls were calculated, $\log_2$-transformed, sorted, and used to calculate cumulative probabilities. *p*-values were calculated using a two-sided Kolmogorov-Smirnov test.

2.5.18 Western blot

Cells were washed with PBS and lysed in lysis buffer (50 mM tris-HCl, 100 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate; pH 7.4) with protease inhibitor cocktail III (EMD Millipore). Lysates were sonicated in a water bath sonicator (Diagenode) at 4 °C for 5 min with 30 s on/off pulses at the 'low' setting. Protein extracts were denatured at 75 °C for 20 min and run at 150V for 1.5 h on 4%–12% NuPAGE Bis-Tris gels in NuPAGE MOPS running buffer (Thermo Fisher). Proteins were transferred to Polyvinylidene Difluoride (PVDF) membrane using NuPAGE transfer buffer (Thermo Fisher) with 10% methanol. Membranes were blocked in blocking buffer [tris-buffered saline containing 5% (w/v) dry milk powder] for 30 min, probed with primary antibodies in blocking buffer for 16 h at 4 °C. Membranes were washed 3 times with TBS and probed with secondary HRP-conjugated antibodies in blocking buffer for 1 hours at room temperature. Signal was detected by Pierce ECL Substrate (Thermo Fisher) and exposure to film.

2.5.19 SUnSET assay

De novo protein synthesis was measured by the SUnSET method(59). Control and homozygous *UBAP2L* deletion HEK293T cells were treated with puromycin (10 µg/ml) for 10 min and then harvested on ice by lysing cells in eCLIP lysis buffer. Protein concentration was

determined with the Pierce BCA protein assay kit (Thermo Scientific). Equal amounts of protein were analyzed by western blotting, as described above. Newly synthesized proteins were detected with an anti-puromycin antibody (1:20,000). Membranes were stripped and re-blotted with mouse anti-GAPDH antibodies (1:8,000) as a loading control. Immunoblots were quantified by densitometric analysis in ImageJ to obtain levels of protein synthesis for each sample.

2.5.20 RCas9-UBAP2L tethered translation assay

HEK293T cells were grown in antibiotic-free DMEM (Life Technologies) media with 10% FBS in 37 °C. Cells were transfected at 50-60% cellular confluency with a 4:1 mix of a PiggyBac transposon vector co-expressing CFP, a PiggyBac transposase vector, and either RCas9-UBAP2L, RCas9-4EBP1, or RCas9 only, using FuGENE HD transfection reagent (Promega). CFP-positive cells (integrants) were collected by FACS, expanded, and transfected again with a PiggyBac transposon vector constitutively expressing RFP, YFP under a tet-inducible promoter, and a gRNA target sequence targeting the YFP reporter. RFP-positive cells were collected by FACS and expanded. Cells were induced with doxycycline (10 ng/ml) for 36 h and quantified by FACS. For each cell the YFP:RFP fluorescence ratio was quantified as a metric of RCas9-UBAP2L or RCas9-4EBP1 mediated post-transcriptional regulation of the target transcript, and CFP fluorescence was used to quantify expression levels of RCas9-UBAP2L, RCas9-4EBP1, or RCas9 only.

2.5.21 Immunocytochemistry

HEK293T were fixed with 4% paraformaldehyde (PFA) for 10 min at room temperature. Cells were permeabilized with PBS with 0.1% Triton X-100 (PBST) and blocked with blocking buffer (5% goat serum in PBST) for 1 hour at room temperature. Cells were then incubated with UBAP2L antibody (1:1,500) in blocking buffer for 16 h at 4 °C and washed with PBST 3 times for 5 min each at room temperature and then incubated with secondary antibody [goat anti-rabbit secondary IgG (H+L) Superclonal Recombinant Secondary Antibody, Alexa Fluor 488 (Invitrogen)] in blocking buffer for 1 hour. After staining, cells were washed again with PBST 3 times for 5 min each at room temperature. Staining of nuclei with DAPI was performed with mounting solution. Images were captured on a Zeiss LSM 710 confocal microscope.

2.5.22 Radiolabeling of RBP-bound RNA fragments

20×10^6 UV-crosslinked HEK293T cells were lysed in 550 µl lysis buffer (50 mM Tris-HCl pH 7.4, 100 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate) with protease inhibitor (Roche). Lysates were sonicated for 5 min (BioRuptor, low setting, 30 s on/off) in an ice-cold water bath. After addition of 2.2 µl Turbo DNase (NEB) and RNase I, diluted 1:3 (high RNase) or 1:25 (low RNase) in low-stringency wash buffer (20 mM Tris-HCl, pH 7.4, 10 mM $MgCl_2$, 0.2% Tween-20), samples were incubated at 37 °C for 5 min with shaking. RNase digestion was stopped with 11 µl Murine RNase inhibitor (NEB) and insoluble material removed by centrifugation (15 min, 15,000 g, 4 °C). Protein-RNA complexes were immunoprecipitated for 16 h at 4 °C with UBAP2L antibody or normal rabbit IgG (Thermo Fisher) pre-coupled to magnetic beads (Dynabeads M-280 Sheep anti-Rabbit IgG, Thermo Fisher). A series of wash steps was employed to ensure stringency, as follows: twice with low stringency wash buffer (see above), twice with high stringency buffer (15 mM Tris-HCl pH 7.4,

5mM EDTA, 2.5 mM EGTA, 1% Triton-X 100, 1% sodium deoxycholate, 0.1% SDS, 120 mM NaCl, 25 mM KCl), twice with high salt wash buffer (50 mM Tris-HCl, pH 7.4, 1M NaCl, 1mM EDTA, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate), twice with low stringency wash buffer, and twice with no-salt wash buffer (50 mM Tris-HCl pH 7.4, 10mM $MgCl_2$, 0.5% NP-40). Protein-RNA complexes were radiolabeled on-bead in 40 µl reactions with T4 polynucleotide kinase (NEB) and 2 µl [γ-^{32}P]ATP (6,000 Ci/mmol, 10 mCi/ml) for 10 min at 37 °C. Beads were washed three times in low-salt wash buffer, resuspended in NuPAGE LDS Sample Buffer (Thermo Fisher) containing 0.1 M DTT. Protein-RNA complexes were denatured at 75 °C for 15 mins and run on 4%–12% NuPAGE Bis-Tris gels in NuPAGE MOPS running buffer (all Thermo Fisher) at 150 V for 1.5 h, wet-transferred to nitrocellulose membrane using NuPAGE transfer buffer (Thermo Fisher) with 10% methanol for 3 h at 200 mA. The membrane was exposed to film for 20 min at room temperature and the film developed

2.5.23 RNA immunoprecipitation and RT-PCR

1.5×10^7 HEK293T cells were washed with PBS and lysed in lysis buffer (50 mM tris-HCl, 100 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate; pH 7.4) with protease inhibitor cocktail III (EMD Millipore). 5% of each lysate sample was stored for preparation of the input RNA sample. The remaining lysates were split into 2 aliquots and immunoprecipitated using 10 µg UBAP2L or IgG antibody and incubated at 4 °C for 8 h. Bound RNA fragments in the IPs were washed 6 times in Net-2 buffer (5 mM tris-HCl, 150 mM NaCl, 0.1% Triton-X 100). Immunoprecipitated RNA was isolated in TRIzol and purification with Direct-zol RNA Kits, following the manufacturer's protocols. 10% of RNA was saved for the no-RT control.

The remaining RNA was reverse transcribed using Superscript III with random primer mix primers (Thermo Fisher). cDNA was diluted 10-fold in water and target transcripts cDNA were amplified by PCR.

2.6 Acknowledgements

We thank members of the Yeo lab, in particular Emily Wheeler and Florian Krach, for helpful discussions. We thank Alexander Palazzo for helpful polysome fractionation advice. E.-C.L. was partly supported by a study-abroad graduate-student fellowship from the Taiwanese government. J.C.S. was partially supported by a Natural Sciences and Engineering Research Council of Canada Postgraduate Doctoral scholarship (PGS D - 532649 - 2019). Y.H. was supported by the UCSD Frontiers of Innovation Scholars Program. G.A.P. was supported by a graduate fellowship from the National Science Foundation. F.E.T. was supported by a postdoctoral fellowship from the American Cancer Society (129547-PF-16-060-01-RMC). S.M. was supported by a postdoctoral fellowship from the Larry L. Hillblom Foundation (2014-A-027-FEL). This work was supported by grants from the NIH (R01HG004659, U19MH107367, R01NS103172 and U41HG009889) to G.W.Y.

Chapter 2, in full, was published in Nature structural molecular biology in 2020 (**Luo EC**, Nathanson JL, Tan FE, Schwartz JL, Schmok JC, Shankar A, Markmiller S, YeeBA, Sathe S, Pratt GA, Scaletta DB, Ha Y, Hill DE, Aigner S, Yeo GW. “Large-scale activity assignment of RNA-interacting proteins by tethered function assays”). Thank you to Stefan Aigner and Gene Yeo for their guidances and contributions to this work. The dissertation author is the primary author of this manuscript.

2.7 Extended Data Table 1

Extended Data Table 1 is available as supplemental tables online

2.8 Extended Data Table 2

Extended Data Table 2 is available as supplemental tables online

Chapter 3: UBAP2L is a functional antagonist to fragile X mental retardation protein in fragile X syndrome

En-Ching Luo[1,2,3], Soojin Lee[6], Sara G. Kosmaczewski[4,5], Archana Shankar[1,2,3], Lindy E. Barrett[4,5], Fen-Biao Gao[6], Stefan Aigner[1,2,3], and Gene W. Yeo[1,2,3,7,*]

[1]Department of Cellular and Molecular Medicine, University of California San Diego, La Jolla, CA, 92093, USA

[2]Stem Cell Program, University of California San Diego, La Jolla, CA, 92093, USA

[3]Institute for Genomic Medicine, University of California San Diego, La Jolla, CA, 92039, USA

[4]Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

[5]Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA 02138, USA

[6]Department of Neurology, University of Massachusetts Medical School, MA 01655, USA

[7]Lead contact

*Correspondence: geneyeo@ucsd.edu

Funding Source: This work was supported by grants from the NIH (R01NS075449, R01HG004659, U19MH107367, and U54HG007005) to G.W.Y., and R21MH109761 to

L.E.B.. E.-C.L. was partly supported by a study-abroad graduate-student fellowship from the Taiwanese government.

Key words: UBAP2L, FMRP, Fragile X syndrome, Translation, RNA, RNA binding protein

3.1 Abstract

Ubiquitin-associated protein 2-like (UBAP2L) protein is a ribosome associated translation enhancer. UBAP2L was found in complex with fragile X mental retardation protein FMRP, and 52% of UBAP2L mRNA targets are also FMRP targets. UBAP2L depletion in a cortical neuronal model of Fragile X Syndrome (FXS) corrects molecular, cellular and electrophysiological defects relevant to autism spectrum disorder. Reduction of the Drosophila ortholog of UBAP2L in a FXS fly model rescues the neurodevelopmental defects due to loss of FMRP. Our efficient and scalable method identifies proteins involved in RNA metabolism and detailed studies of UBAP2L provides a new therapeutic strategy into human disease.

3.2 Introduction

Fragile x syndrome (FXS) is the common inherited cause of intellectual disability and a single gene mutation cause of autism. Research in the past decades has revealed that the absence of Fragile x retardation protein (FMRP) leads to the pathophysiology of FXS. FMRP is an RNA binding protein that regulates translation. The silence of FMRP dysregulates the protein synthesis causing impairment synaptic structure, development, and function (84-86) It

is well known that FMRP interacts with a specific subset of mRNAs to regulate their translation through binding more frequently in the coding region but also within UTRs (84). The coding region binding preference has been shown in many ribosome-associated proteins. Similarly, FRMP has also been characterized to directly associated with polyribosomes(85, 87, 88), indicating FMRP directly regulates translation of its targets while it binds on the ribosome. The PI3K-mTOR pathway is the FMRP targets involved in regulating the protein synthesis machinery indicating FMRP could regulate translation both locally and globally (89, 90). The increase of global translation can cause many FXS associated phenotypes, including exaggerated protein synthesis, hyperelectricity in neurons, and morphological defects in drosophila (91, 92). Therefore, scientists tried to genetic and pharmacological rescue FXS by disruption of FMRP targets. However, the various signaling pathways that are affected by loss of FMRP in FXS complicate the path of finding an effective therapeutic solution. Therefore, our understanding of the basic FMRP biology and the mechanism of FMRP regulating translation becomes a critical question that needs to solve.

Our previous discovery reveals that Ubiquitin associated protein 2 like (UBAP2L) is an RNA binding protein and a novel translation enhancer. Our previous eCLIP data shows that UBAP2L binds dominantly in coding regions which has a similar binding pattern as ribosomes. The polysome profiling track has shown UBAP2L associated with polyribosomes which is consistent with the previous mass-spec study that UBAP2L interacts with ribosomal proteins and ribosome-associated RNA binding proteins such as FMRP(87). UBAP2L could enhance the protein synthesis rate on translation machinery proteins, and globally regulates translation. This translational role is similar to FMRP. Therefore, we are wondering whether UBAP2L could play as an antagonist of FMRP and potentially be a new drug target for FXS.

Interestingly, in this study, we discovered that UBAP2L associates with ribosomes and co-occupies target transcripts of the FMRP. Although protein partners of FMRP had been previously identified (93), to our knowledge, UBAP2L is the first that appears to bind to similar transcript regions as FMRP. FMRP is a well-known RBP that acts as a translation repressor on target transcripts by interacting with elongating ribosomes to facilitate activity-dependent synaptic plasticity and long-term memory formation (85). Loss of FMRP leads to FXS. We hypothesized that if UBAP2L acts as a translational enhancer and antagonistically to FMRP, loss of UBAP2L in FXS models (which lack FMRP) should rescue exaggerated mRNA translation, a hallmark of FXS. Indeed, in a human $FMR1^{-/-}$ stem-cell based model of FXS we demonstrate that UBAP2L depletion antagonizes increased polysome association and normalizes neuronal hyperexcitability caused by loss of FMRP. Genetic loss of UBAP2L also restores neurodevelopmental defects in a *Drosophila* model of the disease. Our results indicate that balanced co-recruitment of UBAP2L and FMRP to transcripts regulates global translation and lays the framework of UBAP2L inhibition as a strategy for understanding and potentially treating FXS. In summary, rapid assignment of function to RBPs can identify new roles that generalize transcriptome-wide, prioritizing RBPs that are critical for our understanding of human biology and disease.

3.3 Results

3.3.1 UBAP2L interacts with FXR family complex

A previous UBAP2L IP-mass spectrometry study recovered peptides from 15 ribosomal proteins (65), further supporting a UBAP2L-ribosome interaction (**Figure 3.1A**). This data also suggested that UBAP2L may interact with all three members of the fragile X related (FXR) family of proteins, which includes the fragile X mental retardation protein (FMRP, encoded by *FMR1*), and its paralogs FXR1 and FXR2. Since these proteins are thought to function as translational repressors, we explored the extent of functional synergy for UBAP2L interaction with these RBPs to modulate translation of target mRNAs. We first determined if UBAP2L and FMRP interact. We found that UBAP2L and FMRP reciprocally co-immunoprecipitate each other both from RNase-treated HEK293T cell lysates (**Figure 3.1C-D**) and in an in vitro binding assay using purified recombinant proteins (**Figure 3.1E-F**). These data indicate that UBAP2L and FMRP can associate in an mRNA-independent manner.

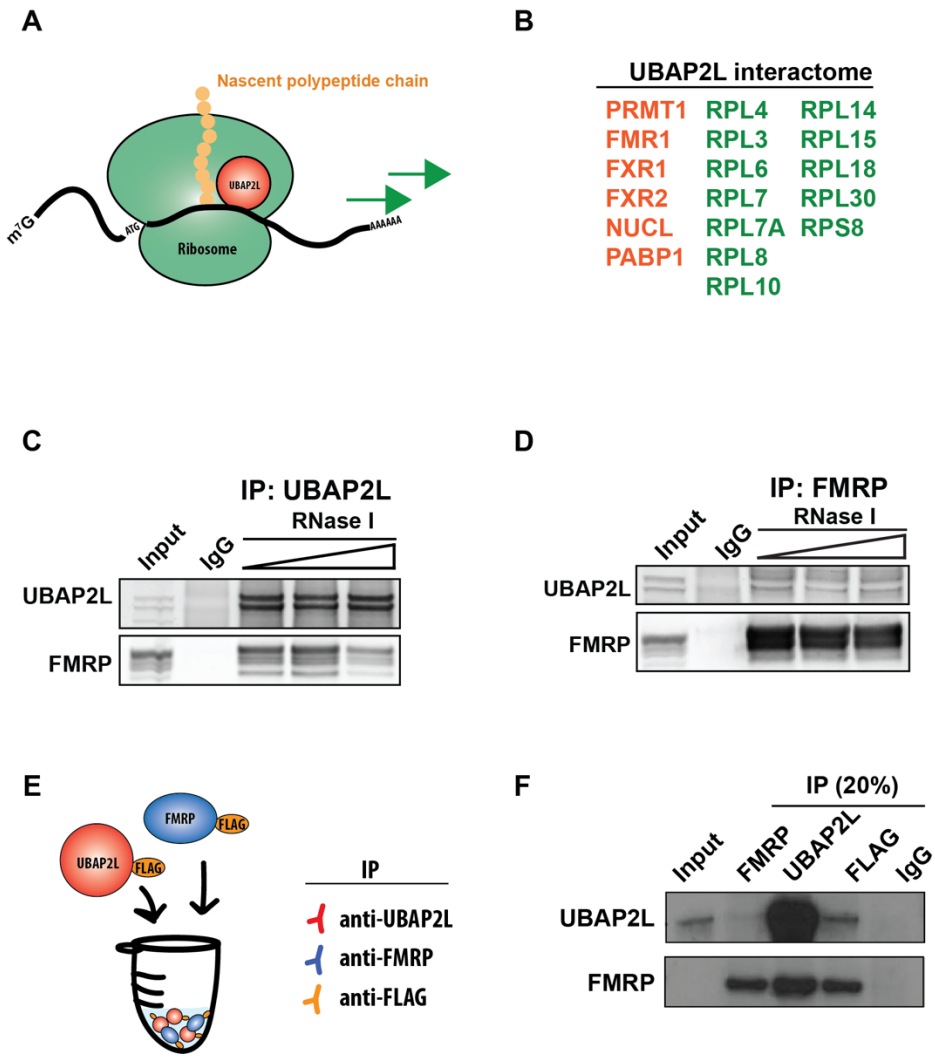


Figure 3. 1 UBAP2L interacts with FMRP

(A) Schematic describing our working model of UBAP2L's role in regulation of global protein synthesis.

(B) The list of proteins interact with UBAP2L in UBAP2L IP-mass spectrometry study(65)

(C-D) Reciprocal co-immunoprecipitation (IP) of UBAP2L and FMRP. HEK293T cell extracts were treated with increasing concentrations of RNase I as indicated, immunoprecipitated with non-immune (IgG) control antibodies and (C) anti-UBAP2L or (D) anti-FMRP antibodies, and western blots performed for both proteins.

(E-F) (E)In vitro co-immunoprecipitation (IP) of UBAP2L and FMRP. (F)Recombinant FLAG-tagged UBAP2L and FLAG-tagged FMRP were incubated and immunoprecipitated under stringent conditions with anti-FMRP, anti-UBAP2L or anti-FLAG antibodies or non-immune (IgG) control antibodies, and western blots performed for UBAP2L and FMRP.

Given their association, we next asked if UBAP2L and FMRP bind similar sequence motifs and mRNA transcripts. We applied the HOMER algorithm to identify sequence motifs recognized by UBAP2L. We found that the most significant motifs were highly enriched in G-rich sequences and 5'-WGGA-3' (W = A or U) (**Figure 3.2A**), which are also previously reported binding motifs of FMRP, recognized by its own RGG domain*(84).* To investigate whether UBAP2L and FXR family proteins regulate similar target gene sets*, we analyzed* published ENCODE eCLIP data(1) for FMRP, FXR1, and FXR2. We considered those transcripts that contain at least one significant (≥ 4-fold-enriched and $p \leq 10^{-3}$ over SMInput) cluster in exonic regions for UBAP2L, FMRP, and FXR2. We found a highly significant overlap of target transcripts among the three RBPs, and between FMRP and UBAP2L ($p = 6.8 \times 10^{-409}$), with 52% of UBAP2L exon targets shared with FMRP (**Figure 3.2B-C**).

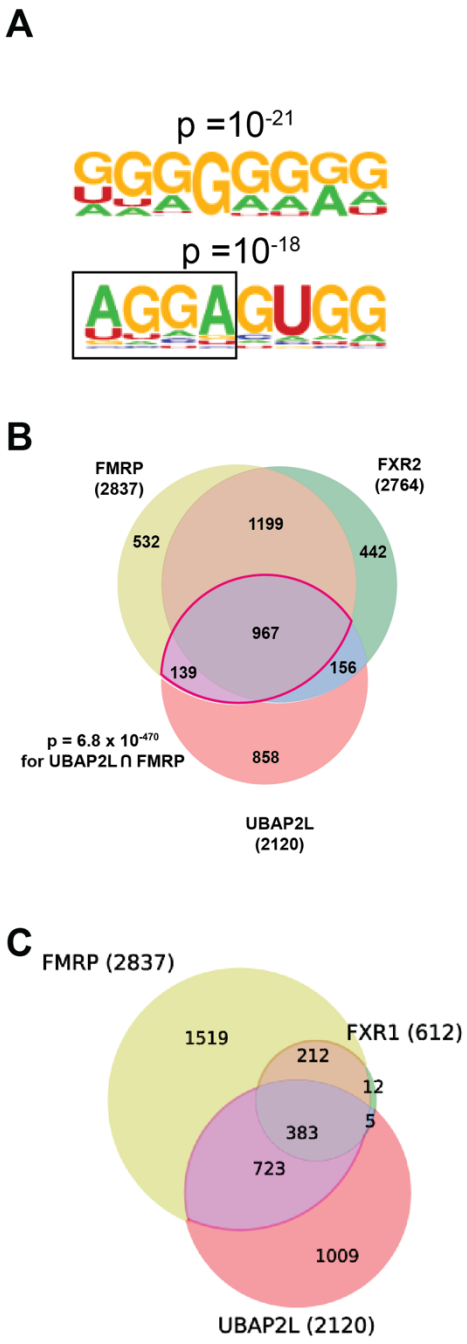


Figure 3. 2 UBAP2L engages FMRP and FXR1/2 targets

(A) De novo sequence motifs of UBAP2L enriched above background within the transcriptome or specific genic regions with associated binomial p-value.

(B-C) Venn diagram showing overlap in exon target transcripts between UBAP2L, FMRP and (B) FXR2 (C) FXR1. p-value was calculated by hypergeometric test using genes expressed in HEK293T cells as background.

Closer inspection revealed largely overlapping binding read densities along many transcripts, for example, the ENO1 transcript (**Figure 3.3A**). In addition, gene ontology (GO) terms related to protein translation were also significantly enriched for UBAP2L exon targets shared with FMRP (**Figure 3.3B**). Our results indicate that UBAP2L forms protein complexes with FMRP, and likely with other members of the FXR family of proteins.

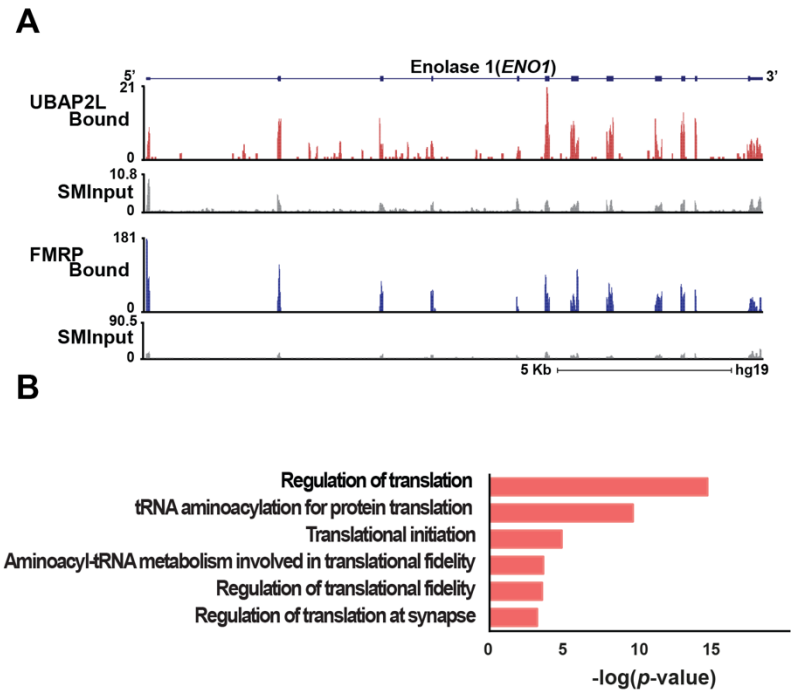


Figure 3. 3 UBAP2L and FMRP shared targets involved in regulation of trnaslation.
(A) Example genome browser views of UBAP2L, FMRP eCLIP data at the *ENO1* locus. *y*-axes denote read density in reads per million (RPM).
(B-C) Gene ontology analysis for UBAP2L exon targets shared with FMRP (n=1106). Significantly enriched GO terms were determined by Fisher's exact test at a false discovery rate of $p < 0.01$. Shown are GO terms that are related to mRNA translation.

Since FMRP is a ribosome-associated protein(88, 94), we assessed the spatial arrangement of FMRP, UBAP2L, and the ribosome. We used eCLIP data from human brain(95) to *identify the binding sites of* UBAP2L and FMRP on rRNA. The relative entropy measurements show that the UBAP2L and FMRP binding sites on ES15L and ES31L overlap, and both bind ES27L (**Figure 3.4A**). We next mapped these interactions onto the X-ray structure of the mammalian ribosome(66) and superimposed the electron density of ribosome-bound FMRP(94). Most ribosomal proteins that co-immunoprecipitate with UBAP2L(65) cluster in close proximity to those that interact with FMRP. In addition, ES31L, which is highly enriched for UBAP2L binding, lies on the same face of the ribosome (**Figure 3.4B**). Collectively, these data support a model in which UBAP2L associates with the ribosome to co-regulate translation in complex with FMRP and discount a scenario in which FMRP and UBAP2L compete for a single ribosomal binding site.

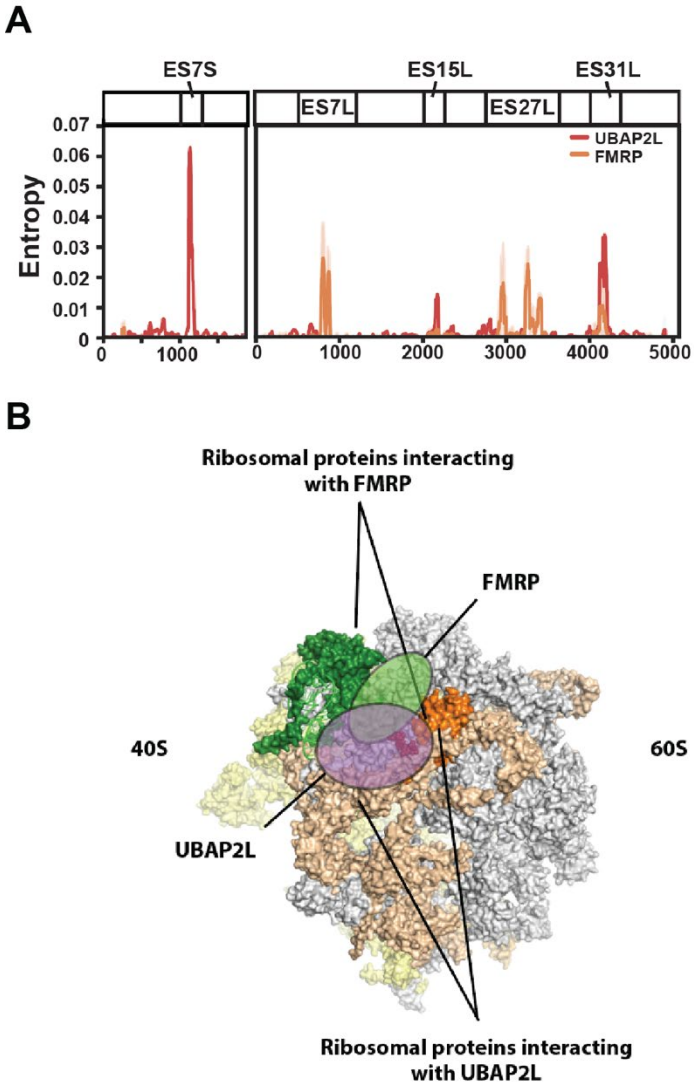


Figure 3. 4 UBAP2L and FMRP shared targets involved in regulation of translation.
(A) Line plots showing the Kullback-Leibler divergence (relative entropy) for UBAP2L in HEK293T cells (red line) and FMRP in human brain (orange line) on 18S and 28S rRNA. The solid line shows the mean of relative entropy, and the shaded areas indicates 10%-90% confidence intervals.
(B) Model of the interactions of UBAP2L and FMRP superimposed on the electron density of the human ribosome. The 40S ribosome is shaded in yellow, and 60S ribosome is shaded in gray. Ribosomal subunit RPL5, previously reported to directly interact with FMRP by chemical cross-linking, is highlighted in green. Ribosomal subunits previously identified as UBAP2L interactors by immunoprecipitation and mass spectrometry (IP-MS) are highlighted in orange. RPL18, previously reported to directly interact with FMRP by chemical cross-linking and with UBAP2L by IP-MS, is highlighted in orange. The location of the UBA2P2L-interacting expansion segment 31L is highlighted in red. The light green denotes the binding location of FMRP, based on previous cryo-electron microscopy studies of the *Drosophila* ribosome and an N-terminal fragment of the *Drosophila* FMRP ortholog. The yellow circle indicates the potential UBAP2L binding site.

3.3.2 UBAP2L counteracts exaggerated translation of FMRP-regulated transcripts in a human stem-cell based model of fragile X syndrome

Classical fragile X syndrome (FXS), the most common monogenic cause of autism spectrum disorder, is caused by loss of FMRP expression due to silencing of its gene FMR1 due to trinucleotide repeat expansions in its promoter and 5'UTR region. By promoting stalling of ribosomes, FMRP is thought to act as a translational brake on hundreds of genes, including positive regulators of translation and synapse components, leading to excessive global protein synthesis in its absence. Thus, the major pathological mechanism of FXS is exaggerated protein synthesis (85, 96). Given our results that UBAP2L enhances translation and regulates FMRP RNA targets, we were interested to see if UBAP2L depletion could counteract increased translation in a stem-cell based model of FXS (**Figure 3.5A**).

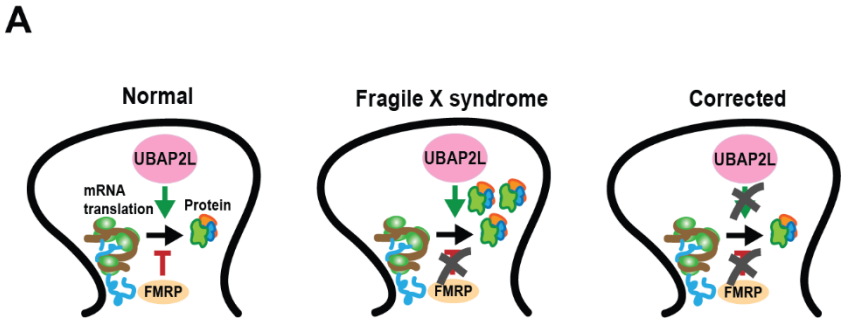


Figure 3. 5 UBAP2L inhibition as a potential therapeutic strategy.
(A) UBAP2L inhibition as a potential therapeutic strategy for the treatment of fragile X syndrome. Exaggerated protein synthesis at the synapse may be normalized by inhibition of UBAP2L.

We first disrupted FMR1 in human pluripotent stem cells (hPSCs) using CRISPR/Cas9 genome editing, generating a pair of otherwise isogenic hPSC lines. We chose the H1 hPSC line because it had been whole-exome sequenced (97) and genotyped to confirm absence of enrichment of marker SNPs associated with psychiatric disease (98). In these FMRP knockout lines, we then depleted UBAP2L by lentiviral delivery of two different shRNAs, with a non-targeting shRNA serving as a control (**Figure 3.6A**). Western blots confirmed the complete absence of FMRP in FMRP knockout lines and substantial stable depletion of UBAP2L in shRNA-targeted hPSC lines, compared to their respective controls (**Figure 3.6B**). UBAP2L-depleted and FMRP knockout hPSC lines exhibited no overt cellular defects. We then differentiated all lines to cortical neurons using a directed protocol (99), which is based on a modified version of that of Zhang et al. (100). Importantly, this approach has been successfully used in other studies of genes implicated in autism (101, 102).

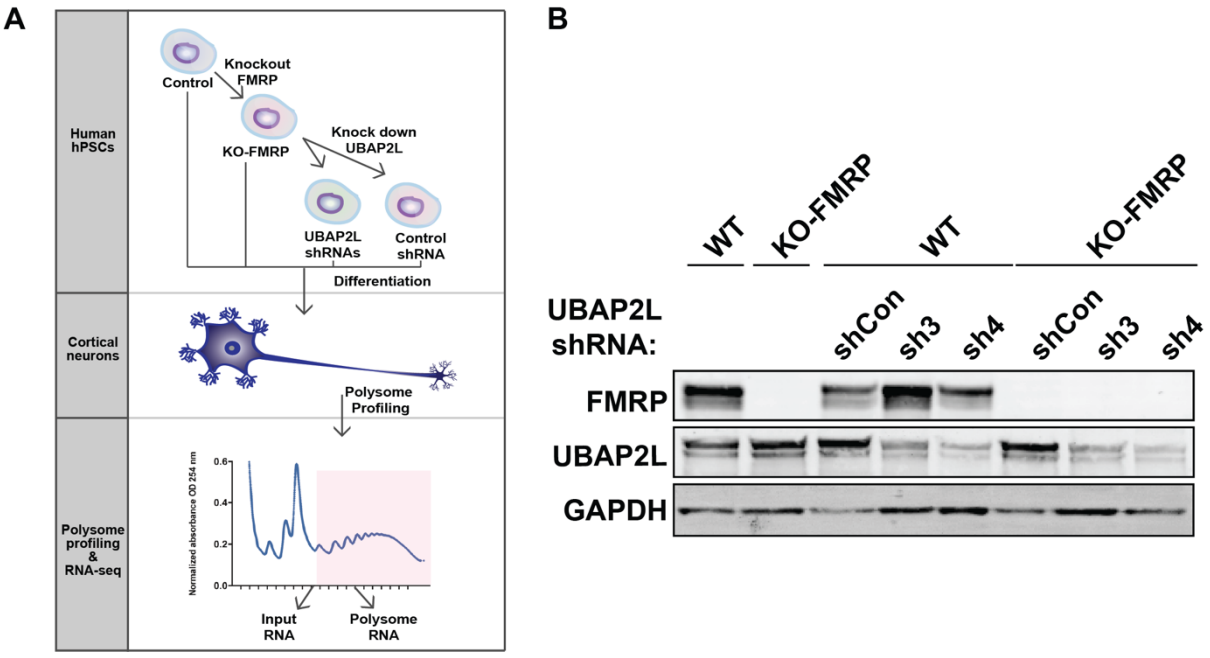


Figure 3. 6 KO-FMRP hPSCs were transduced with UBAP2L-targeting shRNA.
(A) Schematic of the experimental design. H1 hPSCs were CRISPR-engineered to harbor a disruption in *FMR1* (KO-FMRP). KO-FMRP hPSCs were transduced with UBAP2L-targeting shRNA or non-targeting control shRNA lentiviral vectors and cells selected for stable integrants. hPSCs were differentiated to cortical neurons for polysome profiling and RNA-seq analysis.
(B) Western blots of UBAP2L and FMRP following knockout of FMRP and knockdown of UBAP2L in hPSCs. GAPDH served as a loading control.

We performed polysome profiling by measuring the ratios of transcript levels in polysome fractions over inputs, as before. We performed two independent experiments; duplicate samples showed excellent correlation ($R_{control}$ = 0.908, $R_{KO:FMRP}$ = 0.875, $R_{UBAP2L\ shRNAs}$ = 0.912, $R_{Control\ shRNA}$ = 0.904 for input-normalized polysomal transcript levels; **Figure 3.7A** to **D**).

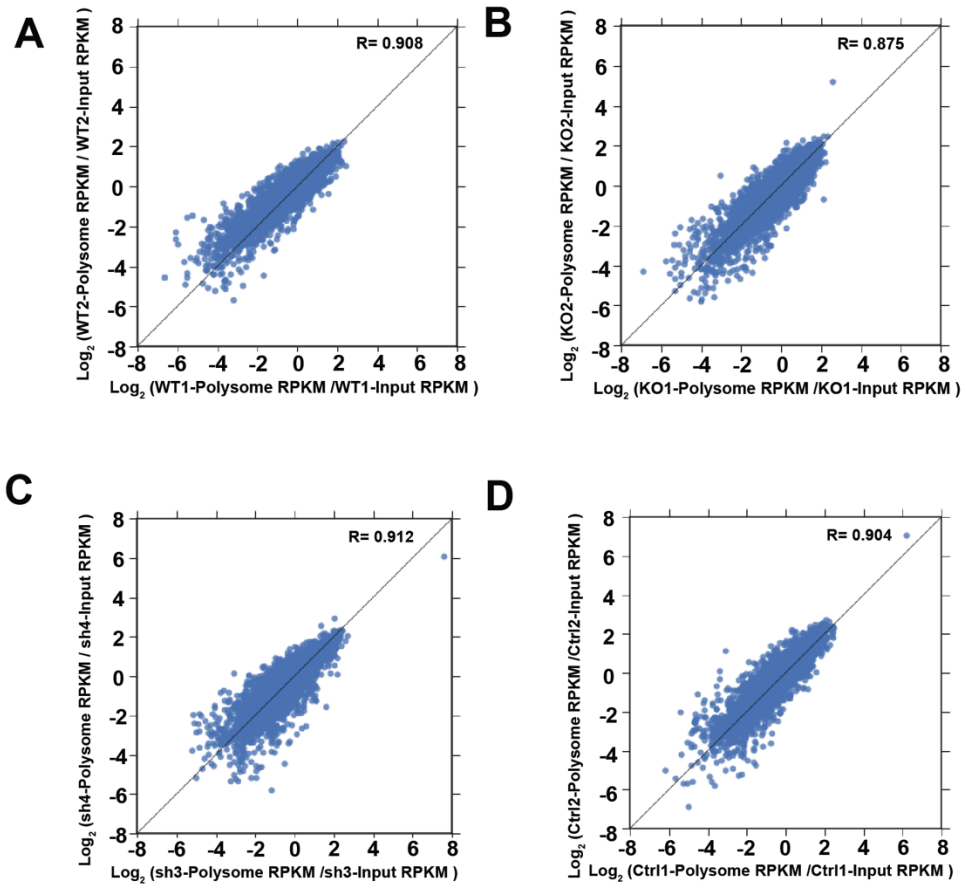


Figure 3. 7 Replicate concordance of polysome data from FMRP knockout, UBAP2L knockdown, and control neurons

(**A–D**) Scatter plots showing correlation of log_2-transformed ratios of input-normalized polysome transcript levels (RPKM) between the (A) two WT lines, (B) two FMRP knockout lines, (C) two FMRP knockout lines treated with UBAP2L-targeting shRNAs (sh3 and sh4), and (D) two FMRP knockout lines treated with the non-targeting shRNA. R values denote Pearson correlation coefficients.

We then investigated changes in polysome association of FMRP target transcripts upon FMRP knockout. Previous FMRP iCLIP analyses in mouse brain had identified 842 FMRP targets (85), 765 of which had human homologs and were detected at RPKM ≥1 in our input samples for polysome analysis. In FMRP knockout neurons we observed a significant ($p = 1\times10^{-22}$; Mann–Whitney U-test, two-tailed) decrease in polysome association of these FMRP targets, but not in non-targets, compared to control (**Figure 3.8A**). These results suggest that FMRP knockout removes stalled ribosomes from these transcripts, thereby decreasing ribosome occupancy, as expected.

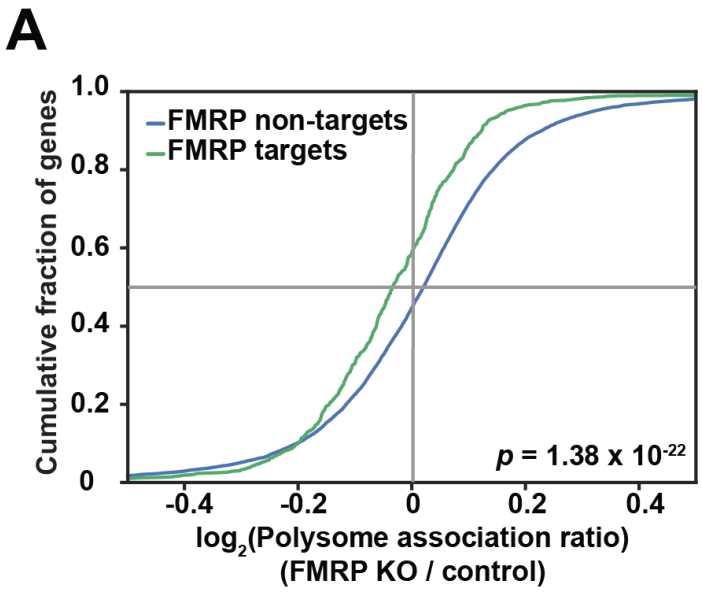


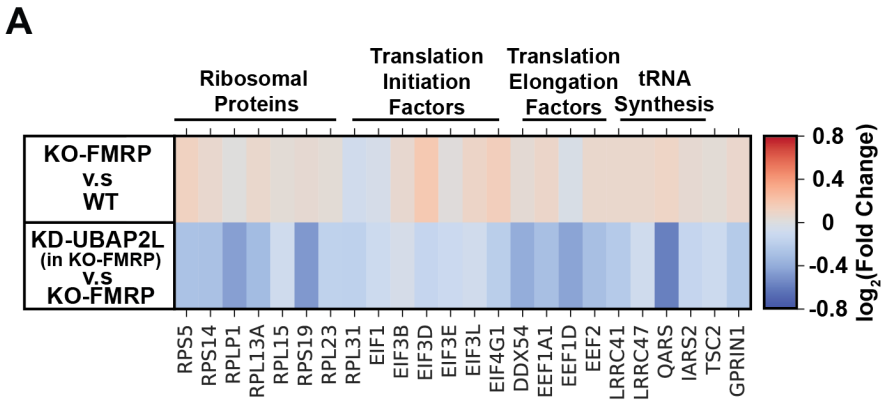
Figure 3. 8 polysome association changes in FMRP target and non-targets
(**A**) Cumulative distribution plots of $\log_2$-transformed input-normalized polysome transcript level (RPKM ≥ 1 in inputs) ratios between FMRP knockout and control neurons for FMRP targets (n = 765) and non-targets (n=11,202). *p*-values were calculated using a two-tailed Mann–Whitney U-test.

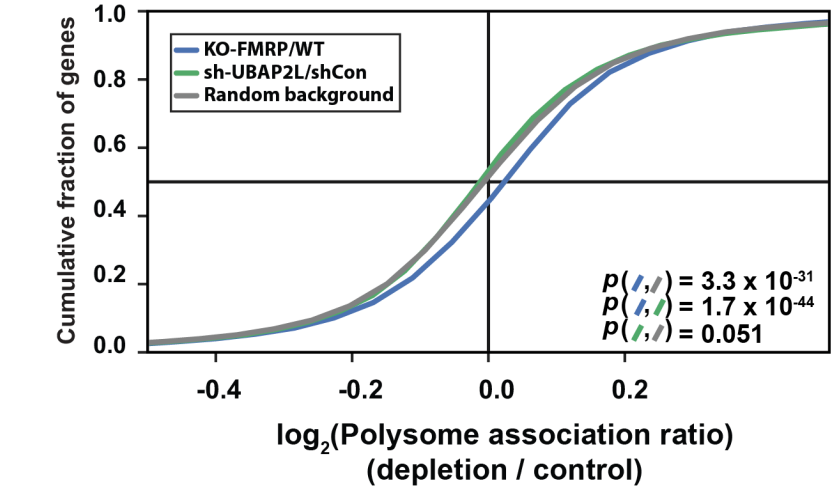
We next investigated the changes in polysome association ratio for the UBAP2L and FMRP shared targets. We calculated the ratio of polysome association between FMRP knockout neurons and control neurons. We found that genes involved in protein synthesis, including ribosomal proteins, translation initiation factors, elongation factors and tRNA synthesis proteins, exhibited increased polysome association upon FMRP loss. In contrast, in the FMRP knockout neurons, UBAP2L depletion decreased the polysome association ratio for these targets (**Figure 3.9A**). Given the decrease in genes involved in protein synthesis, we then asked if UBAP2L knockdown rescues excessive global protein synthesis in FMRP knockout neurons. We examined global changes in polysome association upon FMRP knockout (**Figure 3.9B**). FMRP knockout led to a significant ($p = 3.3 \times 10^{-31}$; Mann–Whitney U-test) increase in polysome association (blue arrow in insert), suggesting limited global translational de-repression. In contrast, on the FMRP knockout background, UBAP2L knockdown led to a significant decrease in global polysome association ($p = 1.7 \times 10^{-44}$; Mann–Whitney U-test). The polysome association ratios in UBAP2L knockdown neurons were similar to random background, suggesting polysome association ratios in FMRP knockout neurons were rescued by UBAP2L knockdown. Consistently, this held true for the set of autism candidate genes from http://gene.sfari.org (**Figure 3.9C**).

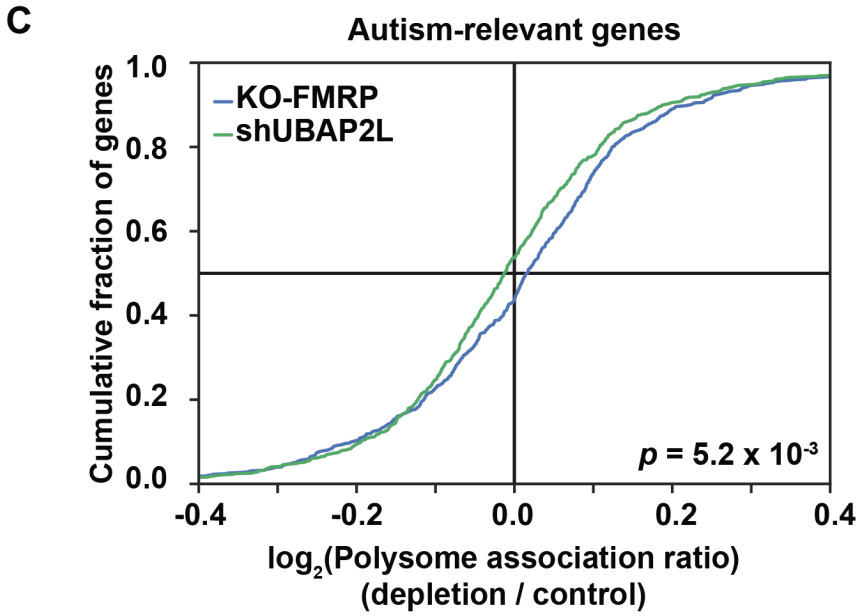
Figure 3. 9 UBAP2L counteracts exaggerated translation of FMRP-regulated transcripts in a human stem-cell based model of fragile X syndrome

(**A**) Heat map showing log_2-polysome association ratio between KO-FMRP and WT neurons, and between UBAP2L-targeting shRNAs (sh3 and sh4) and control shRNA (shCon) on KO-FMRP neurons.

(**B–C**) Cumulative distribution plots of log_2-transformed input-normalized polysome transcript level (RPKM ≥ 1 in inputs) ratios between KO-FMRP and control neurons (blue line), and between shUBAP2L and control shRNA treated neurons (green line), and randomly selected background gray line) for (E) all genes (n = 11,965) and (F) autism-relevant genes (n=743). *p*-values were calculated using a two-sample Mann–Whitney U-test.


A
Ribosomal Proteins
Translation Initiation Factors
Translation Elongation Factors
tRNA Synthesis
KO-FMRP v.s WT
KD-UBAP2L (in KO-FMRP) v.s KO-FMRP
0.8
0.4
0
-0.4
-0.8
log2(Fold Change)
RPS5
RPS14
RPLP1
RPL13A
RPL15
RPS19
RPL23
RPL31
EIF1
EIF3B
EIF3D
EIF3E
EIF3L
EIF4G1
DDX54
EEF1A1
EEF1D
EEF2
LRRC41
LRRC47
QARS
IARS2
TSC2
GPRIN1


B
KO-FMRP/WT
sh-UBAP2L/shCon
Random background
Cumulative fraction of genes
1.0
0.8
0.6
0.4
0.2
0.0
p(/,/) = 3.3 x 10-31
p(/,/) = 1.7 x 10-44
p(/,/) = 0.051
-0.4
-0.2
0.0
0.2
log2(Polysome association ratio)
(depletion / control)


C
Autism-relevant genes
KO-FMRP
shUBAP2L
Cumulative fraction of genes
1.0
0.8
0.6
0.4
0.2
0.0
p = 5.2 x 10-3
-0.4
-0.2
0.0
0.2
0.4
log2(Polysome association ratio)
(depletion / control)

Finally, we evaluated if UBAP2L had reciprocal effects on polysome association of individual transcripts. For each transcript detected at RPKM ≥1 in all inputs, we calculated the ratios of polysome association of FMRP knockout to control, and of UBAP2L shRNA knockdown to controls. This analysis revealed that overall, UBAP2L acts as a translational activator on transcripts repressed by FMRP (**Figure 3.10A**). For instance, knockdown of UBAP2L decreased ribosome association of the FMRP target, calcium voltage-gated channel subunit alpha1 A (CACNA1A), in which gain-of-function mutations are the cause of multiple neurological disorders. We also identified DNA methyltransferase 1 (DNMT1), as a transcript reciprocally regulated by FMRP and UBAP2L (**Figure 3.10B**). Together, our results demonstrate that UBAP2L inhibition counteracts translational de-repression of FMRP-target transcripts which are mis-regulated in FXS.

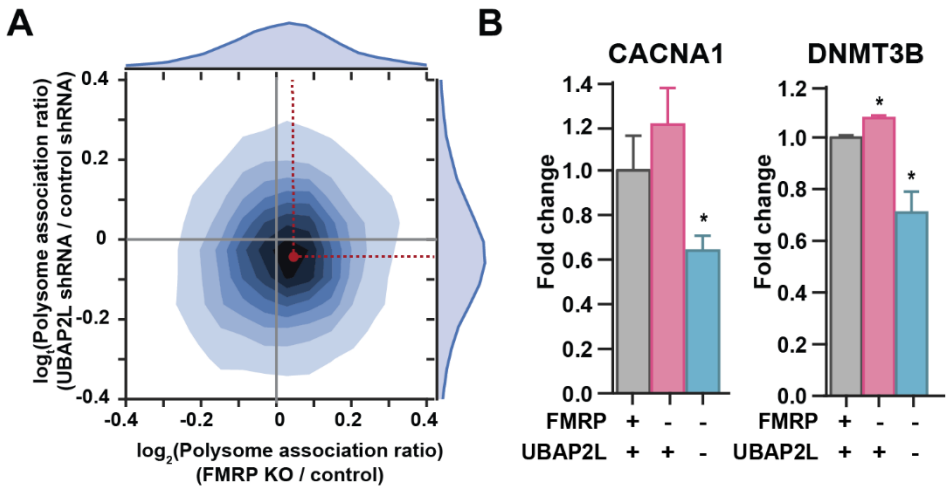


Figure 3. 10 UBAP2L acts as a translational activator on transcripts repressed by FMRP
(A) Density plot of differences of polysome association (input-normalized polysome transcript level ratios) for each transcript (RPKM ≥ 1 in inputs; n= 11,965) between FMRP knockout and control neurons (*x*-axis) and between UBAP2L knockdown and control neurons (*y*-axis).
(B) Input-normalized polysome transcript levels for *CACNA1* and *DNMT3B* transcripts between knockout FMRP vs control neurons (WT) or UBAP2L knockdown vs non-targeting shRNA control.

3.3.3 Depletion of UBAP2L rescues functional FXS defects in neurons

Exaggerated protein translation in FXS is thought to impact neuronal development and activity by perturbing synaptic maturation and function, eventually altering circuit-level processes in the developing adult. Our results thus raise the possibility that reduction of UBAP2L may rescue these deficits. To test this, we first examined spontaneous activity of UBAP2L-depleted FMRP knockout iPSC-derived cortical neurons on multi-electrode arrays. We disrupted FMR1 by CRISPR genome editing in the CV-B iPSC line, using the parental unedited line as an isogenic control, and generated cortical neurons using an established protocol (103). As expected, at day 35 of differentiation, the large majority of cells were positive for the cortical neuronal markers CTIP2 and TBR1 (**Figure 3.11A).**

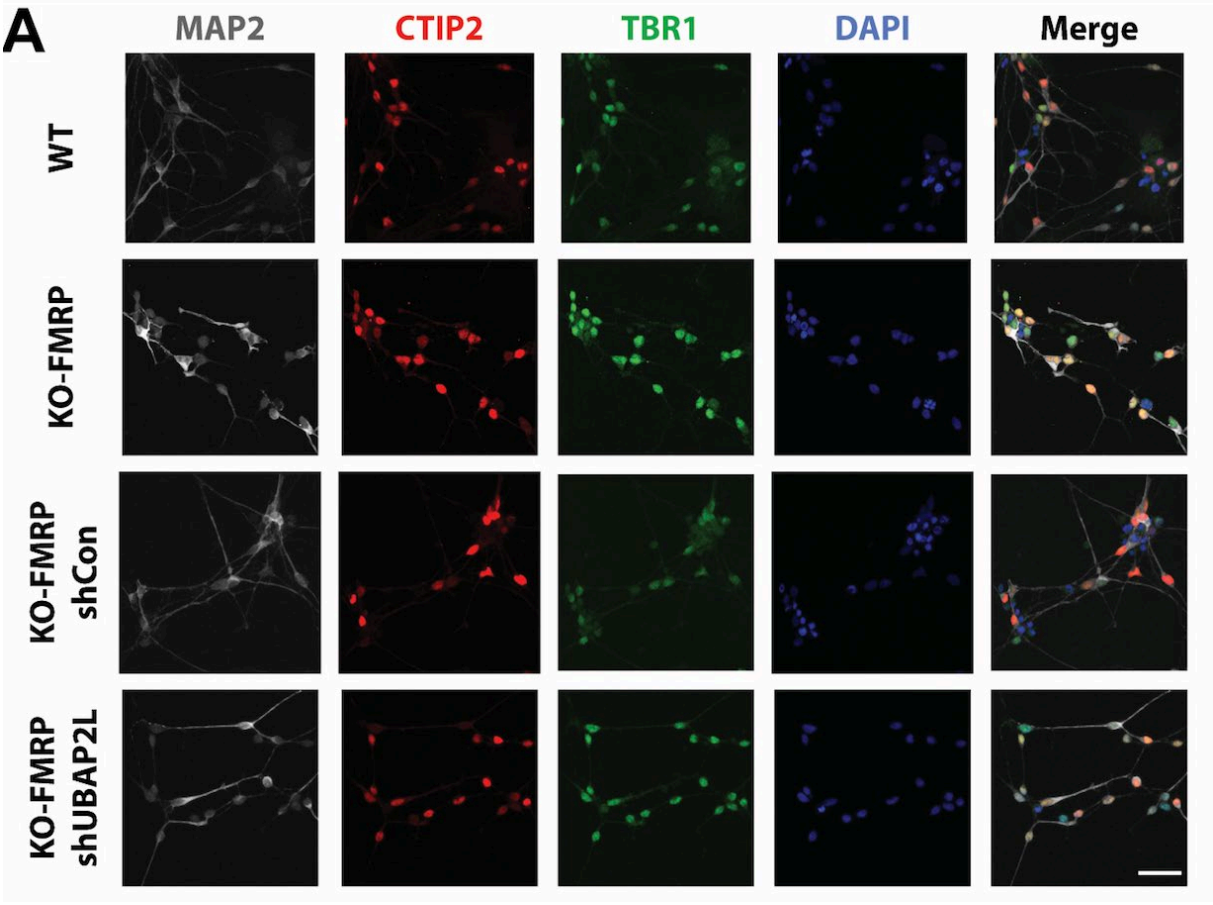


Figure 3. 11 Cells were positive for the cortical neuronal markers CTIP2 and TBR1 at Day 35
(A) Immunofluorescence images showing cortical neuron markers in control neurons (WT), knockout FMRP neurons (KO-FMRP), and KO-FMRP neurons treated with UBAP2L knockdown shRNAs (shUBAP2L). MAP2 (gray) is a pan-neuronal marker. CTIP2 (red), and TBR1 (green) are cortical neuron markers. Scale bar = 30 μm.

Previous studies established that FXS neurons show hyperexcitability, reflected in an increased spontaneous firing rate, which is thought to impair neural network function and memory formation in vivo (92). Consistent with this finding, we observed a two-fold increase in the firing rate of our FMRP knockout neurons compared to control. Strikingly, UBAP2L knockdown in FMRP knockout neurons using two shRNA constructs significantly reduced ($p < 0.05$, Student's t-test) neuronal firing rates, to levels similar to those observed in the control line, suggesting that the reduction of UBAP2L rescues hyperexcitability in FXS (**Figure 3.12A**).

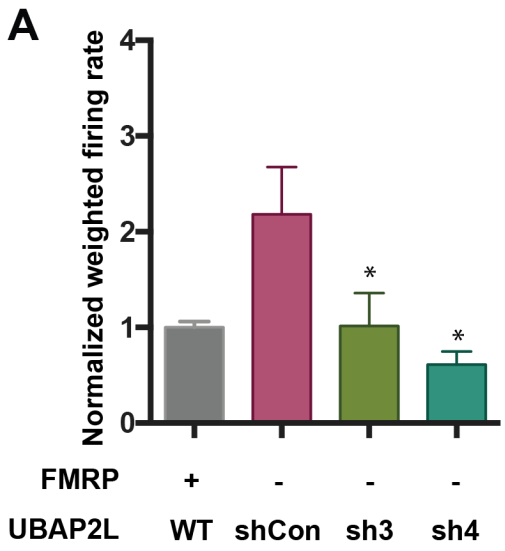
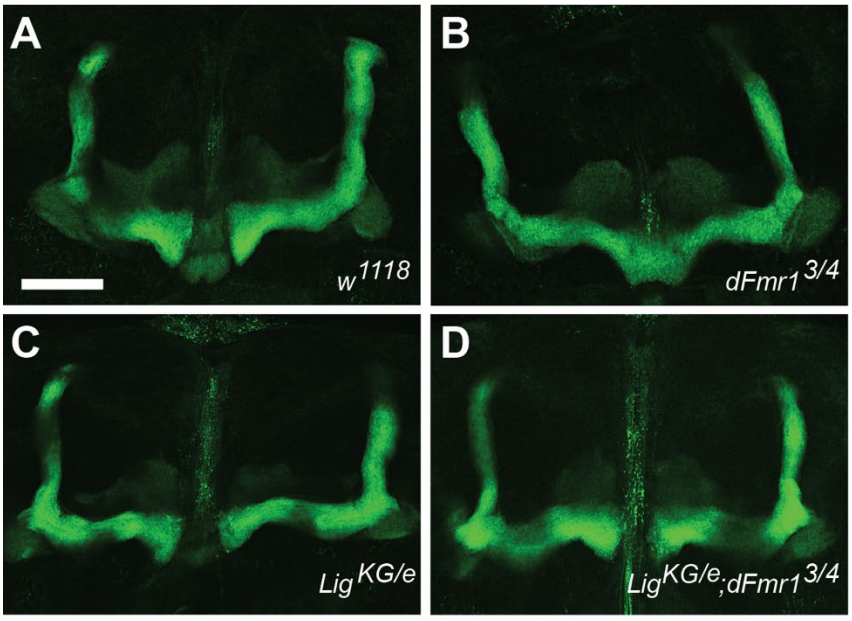


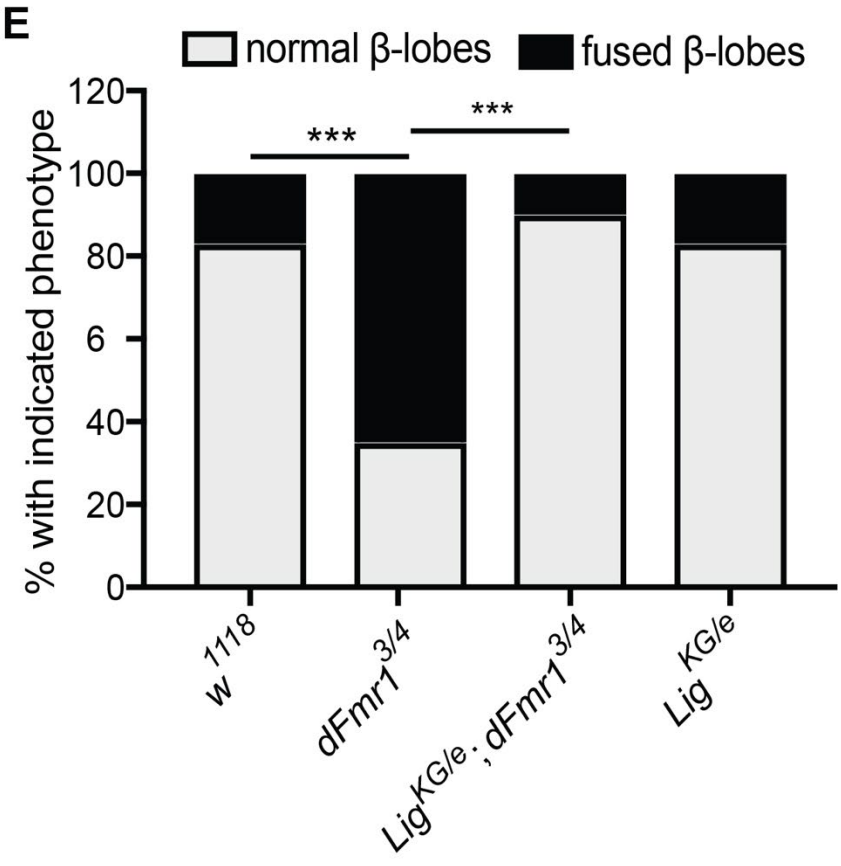
Figure 3. 12 The reduction of UBAP2L rescues hyperexcitability in FXS
(A) Bar plots showing weighted firing rates of knockout FMRP neurons treated with UBAP2L non-targeting shRNA control (shCon) or UBAP2L-targeting shRNAs (sh4 and sh4), normalized to control neurons (WT). Error bars denote mean ± SE, * $p < 0.05$ by Student's t-test.

3.3.4 Reduction of drosophila ortholog of UBAP2L rescues morphological FXS deficits

To see if depletion of UBAP2L can also counteract neurodevelopmental defects modeled by loss of FMRP, we turned to the Drosophila dfmr1 mutant model of FXS (91). In the brain of Drosophila, the mushroom body (MB) is a center for associative learning and memory. It is composed of axonal fibers of three specific neuronal classes that form paired α-, β-, and γ-lobes (104). In the wild-type fly brain, the β-lobes terminate at the midline. However, β-lobe fibers cross or fuse at the midline in dfmr1 mutant flies. To test the hypothesis that reduction of the Drosophila ortholog of UBAP2L, Lingerer (Lig), in the dfmr1 fly model rescues the neurodevelopmental defects due to loss of FMRP of dfmr1 and Lig, we first confirmed this neurodevelopmental defect in $dfmr1^{3}/dfmr1^{4}$ mutant flies. We stained whole adult brains with an anti-Fasciclin II antibody to visualize the MB β-lobe of the fly brain (**Figure 3.13A**). In contrast to control flies, significantly ($p < 10^{-4}$, χ^2 test) more $dfmr1^{3}/dfmr1^{4}$ mutant flies showed the β-lobe midline-crossing defect (**Figure 3.13B**), as reported previously (91). To examine the effect of loss of Lig activity in dfmr1 mutant flies, we generated two fly lines trans-heterozygous for different dfmr1 and lig mutant alleles. Strikingly, loss of Lig activity through the combination of $lig^{KG08209}/lig^{e04268}$ mutant alleles did not cause MB defects but rescued the β-lobe midline-crossing phenotype observed in $dfmr1^{3}/dfmr1^{4}$ mutant flies (**Figure 3.13C** to **E**). These results illustrate that reduction of Lig rescues neurodevelopmental defects caused by loss of FMRP. Taken together, the results from our cellular and in vivo FXS models show that depletion of UBAP2L rescues both functional deficits and neuroanatomical defects, the two major hallmarks of the disease.

Figure 3. 13 Rescue of neurodevelopmental phenotypes in fragile X models by UBAP2L depletion
(**A-D**) Representative immunofluorescence images showing mushroom bodies (MB) β-lobe fibers in fly brains revealed by anti-Fas II antibody immunostaining in (A) wildtype, (B) *dfmr1* mutant, , (C) *Lig* mutant and (D) *Lig* and *dFmr1* mutant flies. Arrows indicate the midline. Scale bar = 50 μm.
(**E**) Percentage of normal and fused β-lobes in brains analyzed for each genotype ($n \geq 20$, *** $p < 10^4$ by χ^2 test).


A
B
C
D
w1118
dFmr1 3/4
Lig KG/e
Lig KG/e;dFmr1 3/4


E
normal β-lobes
fused β-lobes
% with indicated phenotype
120
100
80
6
40
20
0
***
***
w 1118
dFmr1 3/4
Lig KG/e ; dFmr1 3/4
Lig KG/e

3.4 Discussion

Fragile X syndrome (FXS), an X-linked autism spectrum disorder and the leading genetic cause of intellectual disability, is caused by loss of function of FMRP (encoded by *FMR1*). A number of lines of investigation have established that FXS is a disease of persistently elevated translation (105). It is thought that FXS neurons are compromised in responding to synaptic activity: as they cannot further increase translation, protein-synthesis dependent synaptic remodeling processes, in particular long-term depression (LTD), are impaired. Mechanistically, FMRP is thought to repress translation by promoting ribosome stalling on its target mRNAs (85) and inhibiting translation initiation (106). Although at least 7 CLIP studies have reported FMRP targets for hundreds of RNAs each, the set of concordant FMRP targets from 5 of these (107) is small (n = 53). It remains unclear which RNA substrates are regulated by FMRP – and if so, by which mechanism – and which of those contribute to FXS pathophysiology. The diversity in FMRP mechanisms of action and the uncertainty with respect to the key FXS targets have hampered development of therapeutic strategies. For example, despite promising results from FXS mouse models, clinical trials using inhibitors of mGluR5 receptors, which promote protein synthesis dependent LTD in glutamatergic neurons, have failed to show efficacy (108, 109). Issues regarding patient stratification, treatment duration, clinical endpoints and drug tolerance notwithstanding (110), it stands to reason that targeting any single one of FMRP's multiple regulatory pathways is insufficient.

Our studies revealed UBAP2L as a novel translational enhancer and FMRP-interacting RBP with a role in opposing FMRP function. We present several lines of evidence pointing to UBAP2L as a ribosome-bound FMRP co-regulator and functional antagonist. Since our

previous studies have shown that UBAP2L deletion leads to a robust decrease in global cellular translation and more than 65% of UBAP2L eCLIP reads were rRNAs, suggesting that it functions as a translational enhancer and regulates translation through direct binding. First, our transcriptome-wide analyses reveal that FMRP and UBAP2L co-regulate a significant number of mRNA targets. Strikingly, of the 53 FMRP consensus targets from 5 CLIP datasets (107), 45 (85%) are bound by UBAP2L. Interestingly, several mRNAs targeted by both FMRP and UBAP2L are central regulators of RNA stability, translation, and protein degradation, suggesting a key role for UBAP2L in modulating FMRP-dependent global protein homeostasis. Secondly, in combination with our data suggesting that, like FMRP, UBAP2L binds translating ribosomes, we propose that both proteins are dynamically co-recruited to translating ribosome-mRNP complexes to control translation. Indeed, we find that UBAP2L depletion in FMRP knockout neurons reversed changes in polysome association of FMRP-regulated transcripts, including key autism-relevant genes.

Lastly, we find that depletion of UBAP2L corrects disease-relevant phenotypes in two distinct models of FXS. Increased neuronal and circuit-level excitability in multiple regions of the brain, including the sensory cortex, are key phenotypes in mouse models of FXR, and are considered to represent an endophenotype of sensory hypersensitivity in human patients. It is thought that loss of FMRP-dependent translational control of key synaptic proteins and receptors causes disruption of synaptic maturation and plasticity, ultimately leading to elevated excitability (111). We find that UBAP2L knock-down in FXR iPSC-derived cortical rescues elevated spontaneous firing, suggesting that normalization of exaggerated translation through UBAP2L depletion is sufficient to normalize neuronal hyperexcitability in this model. We also observe rescue of the neural architecture of the mushroom body, a complex structure of the fly

brain critically involved in learning and memory formation. As previously reported, we observe severe midline crossing of β-lobes in dFmr1 mutant flies (91). This phenotype is considered to be a correlate to impaired learning and memory in FXS patients and has been successfully used as a readout in small molecule screens for potential therapeutic targets and candidates in FXS (112). We find that genetic deletion of the UBAP2L fly ortholog normalizes the β-lobe midline crossing phenotype of dFmr1 mutant flies, demonstrating rescue of neuroanatomical deficits in an animal model of FXS.

In the aggregate, our findings indicating that UBAP2L antagonizes FMRP function point to inhibition of UBAP2L as an alternative therapeutic strategy in FXS. Additional investigation will be required to further dissect the molecular mechanisms by which UBAP2L promotes translation of its target mRNAs, and to what extent UBAP2L depletion in cell-based and animal models of FXS reverses clinically relevant cellular and functional phenotypes. In conclusion, our study presents unprecedented resources that demonstrate that rapid assignment of function to RBPs identifies new molecular roles that generalize transcriptome-wide, and facilitate new insights on the importance of post-transcriptional gene regulation in human biology and disease.

Future perspectives

Our results provide a new therapeutic strategy in FXS. However, many questions of UBAP2L regulations are still unclear. In the future study, we should focus on how UBAP2L regulates translation together with FMRP ? Do these two proteins both bind on the ribosome at the same time? Or they regulate translation separately? Since UBAP2L enhances the

translation, it would be interesting to understand how UBAP2L coordinates with translation repressor FMRP in the normal healthy cells. When and how the translation is up-regulated, down-regulated? What do the signaling pathways regulate the translation?

Furthermore, due to our rescue results were done on the iPSC cells and in drosophila, it is crucial to see if the depletion of UBAP2L rescue FXS defects in mammalian models such as FXS mice. The mice study could also provide an insight how translation regulates synapse plasticity, and affect neuron development. Lastly, If UBAP2L could be a new therapeutic strategy in FXS, it is important to discover the methods to reduce the expression level of UBAP2L or inhibit the interaction between UBAP2L and FMRP. It could be using ASO to degrade the UBAP2L gene or identify compounds to disrupt the functional interaction between UBAP2L and FMRP.

3.5 Materials and Methods

3.5.1 eCLIP library preparation and sequencing

eCLIP was performed essentially as described (*6*). Briefly, for each RBP, 1×10^7 HEK293T cells were UV-crosslinked (400 mJ/cm^2, 254 nm) and lysed. Lysates were sonicated and treated with RNase I to fragment RNA. 2% of each lysate sample was stored for preparation of a parallel size-matched input (SMInput) library. The remaining lysates were immunoprecipitated using RBP-specific antibodies. Bound RNA fragments in the IPs were dephosphorylated and 3'-end ligated to an RNA adapter. Protein-RNA complexes from SMInputs and immunoprecipitates (IPs) were run on an SDS-PAGE gel and transferred to nitrocellulose membrane. Membrane regions spanning from the exact RBP sizes to 75 kDa

above were excised and RNA released from the complexes with proteinase K. SMInput samples were dephosphorylated and 3'-end ligated to an RNA adapter. All RNA samples (IPs and SMInputs) were reverse transcribed with AffinityScript (Agilent). cDNAs were 5'-end ligated to a DNA adaptor. cDNA yields were quantified by qPCR and 100-500 fmol of libraries generated with Q5 PCR mix (NEB).

3.5.2 Computational analysis of eCLIP sequencing data

Reads were processed essentially as described (*6*). Briefly, reads were adapter-trimmed and mapped to human-specific repetitive elements from RepBase (version 18.05) by STAR (*76*). Repeat-mapping reads were removed and remaining reads mapped to the human genome assembly hg19 with STAR. PCR duplicate reads were removed using the unique molecular identifier (UMI) sequences in the 5' adapter and remaining reads retained as 'usable reads'. Peaks were called on the usable reads by CLIPper (*32*) and assigned to gene regions annotated in GENCODE (v19) with the following descending priority order: CDS, 5'UTR, 3'UTR, proximal intron, and distal intron. Proximal intron regions are defined as extending up to 500 bp from an exon-intron junction. Each peak was normalized to the size-matched input (SMInput) by calculating the fraction of the number of usable reads from immunoprecipitation to that of the usable reads from the SMInput. Peaks were deemed significant at ≥ 4-fold enrichment and $p \leq 10^{-5}$ (χ^2 test, or Fisher's exact test if the observed or expected read number in eCLIP or SMInput was below 5).

3.5.3 Repeat-family centric mapping

Binding to ribosomal rRNA was quantified using a family-aware repeat element mapping pipeline (*14*). Briefly, reads were mapped to a database of 7,419 multi-copy element transcripts, including the 5S, 5.8S, 18S, and 28S rRNA as well as tRNAs, retrotransposable elements, and numerous other RNAs. Reads mapping to multiple element families were not considered for further analysis. To summarize relative enrichment between IP and input, relative information was defined as the Kullback-Leibler divergence (relative entropy): $p_i \times log_2(\frac{p_i}{q_i})$, where p_i is the fraction of total reads in IP that map to a queried repetitive element i, and q_i is the fraction of total reads in input for the same element.

3.5.4 RNA-seq library preparation and analysis

RNA was extracted from cells with TRIzol (Invitrogen). Strand-specific RNA sequencing libraries were prepared from 0.5-3 μg total RNA using the Illumina TruSeq Stranded mRNA Sample Preparation kit (Illumina). Libraries were sequenced on the Illumina HiSeq 4000 platform at a depth of at least 12×10^6 reads per sample in SE50 mode.

RNA-sequencing reads were trimmed using *cutadapt* (v1.4.0) of adaptor sequences and mapped to repetitive elements (RepBase v18.04) using the STAR (v2.4.0i). Reads did not map to repetitive elements were then mapped to the human genome (hg19). Using GENCODE (v19) gene annotations and featureCounts (v.1.5.0) to create read count matrices. Differential expression was calculated using DESeq2 version 1.10.1 (*77*), individually pairing each knockdown or over-expression experiments with their respective controls. Genes with RPKM values lower than 1 were not used.

3.5.5 Polysome profiling

For lysate preparation, 40×10^6 neurons cells were prepared. Before collection, cells were treated with cycloheximide (CHX) at 100 µg/ml for 5 min at 37°C. Culture media was removed, cells were washed twice with cold PBS-CHX (100 µg/ml), resuspended in PBS-CHX by centrifugation at 200 × *g* at 4 °C for 5 min, collected in PBS-CHX and snap frozen in liquid nitrogen. Cells were lysed by trituration through a 27 gauge needle in 400 µl polysome lysis buffer (20 mM Tris-HCl pH 7.4, 150 mM NaCl, 5mM $MgCl_2$) with 1× protease inhibitor cocktail (EMD Millipore), CHX 100 µg/ml, 1 mM DTT, 25 U/ml DNase (TurboDNase; Thermo Fisher), 20 U/ml RNase inhibitor (RNaseOUT, Thermo Fisher) and incubation on ice for 15 min. Lysates were clarified by centrifugation at 17,500 × *g* at 4 °C for 5 min. 50 µl was reserved for inputs and the remainder used for fractionation.

For fractionation, a 14 ml 10-50% (w/v) sucrose gradient was prepared in polysome buffer. Samples were loaded on the sucrose gradient and centrifuged in a swinging bucket rotor at 35,000 × *g* at 4 °C for 3 h. Fractions were collected from the top and UV absorbance monitored using a Gradient Station (BioCamp) equipped with ECONO UV monitor (BioRad). Fractions (500 µl each) were collected using a FC203B (Gilson) fraction collector. Fractions containing polysomes were pooled. Total RNA from the inputs and polysome pools were extracted in TRIzol-LS (Thermo Fisher) and purified with Direct-zol RNA kits (Zymo). RNA sequencing libraries were generated sequenced and reads processed as described above.

For analysis of fractions by western blotting, trichloroacetic acid was added to samples to 20% (v/v). Protein was precipitated overnight at 4 °C and collected by centrifugation at 15,000 × *g* for 20 min at 4 °C. Protein pellets were washed twice with 1 ml ice-cold acetone,

centrifuged at 15,000 × *g* for 15 min at 4 °C, dried at room temperature, neutralized, re-suspended, and denatured by incubation in 50 mM tris-HCl pH 7.4 at 65 °C for 30 min and 98 °C for 15 min. Western blotting was performed as described above.

3.5.6 Polysome association analysis

The transcript RPKMs of input and polysome fractions were calculated from the read count matrices. Only genes with RPKM ≥ 1 were considered. Polysome association was measured by calculating the RPKM ratio of transcript levels in polysomes over input. Polysome association ratios between depletion samples and their respective controls were calculated, $\log_2$-transformed, sorted, and used to calculate cumulative probabilities. *p*-values were calculated using a two-sided Kolmogorov-Smirnov test.

3.5.7 FMRP-UBAP2L co-immunoprecipitation and in vitro interaction assay

Immunoprecipitations and western blot were performed as described by us (*29*), using 15×10^6 HEK293T cells lysed in 450 µl lysis buffer. Lysates were treated with 1000 U, 40 U, or 1 U RNase I (Thermo Scientific) at 37 °C for 5 min with shaking, followed by immunoprecipitation and western blot. For the in vitro interaction assay, affinity-purified recombinant FLAG-tagged human FMRP and UBAP2L (Origene) were incubated at 150 nM each in 400 µl binding buffer (20 mM tris-HCl, 10% glycerol, 100 mM NaCl, 1 mM EDTA, 5 mM $MgCl_2$, 0.1% NP-40; pH 7.4) at 4 °C for 8 h. The sample was split into 4 aliquots, and 10 µg FMRP (MBL), UBAP2L (Bethyl), FLAG M2 (Sigma), or control rabbit IgG (Invitrogen) antibody was added. Immunoprecipitation was at 4 °C for 8 h. Immune-complexes were

collected with protein A/G beads (Dynabeads, Thermo Fisher) and washed 3 × in binding buffer containing 500 mM NaCl, followed by western blotting.

3.5.8 Western blot

Cells were washed with PBS and lysed in lysis buffer (50 mM tris-HCl, 100 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate; pH 7.4) with protease inhibitor cocktail III (EMD Millipore). Lysates were sonicated in a water bath sonicator (Diagenode) at 4 °C for 5 min with 30 s on/off pulses at the 'low' setting. Protein extracts were denatured at 75 °C for 20 min and run at 150V for 1.5 h on 4%–12% NuPAGE Bis-Tris gels in NuPAGE MOPS running buffer (Thermo Fisher). Proteins were transferred to Polyvinylidene Difluoride (PVDF) membrane using NuPAGE transfer buffer (Thermo Fisher) with 10% methanol. Membranes were blocked in blocking buffer [tris-buffered saline containing 5% (w/v) dry milk powder] for 30 min, probed with primary antibodies in blocking buffer overnight at 4 °C. Membranes were washed 3 times with in tris-buffered saline and probed with secondary HRP-conjugated antibodies in blocking buffer for 1 hours at room temperature. Signal was detected by Pierce ECL Substrate (Thermo Fisher) and exposure to film.

3.5.9 Lentiviral shRNA treatment and cortical differentiation of human pluripotent stem cells

H1 human pluripotent stem cells (hPSCs) were edited by CRISPR-Cas9 to generate KO and WT *FMR1* clones. Details are described in a separate study (Kosmaczewski et al., *in preparation*). All hPSCs were maintained in mTeSR medium (Stem Cell Technologies) supplemented with 5 μM XAV939 (Stemgent), grown on Geltrex-coated plates (Life

Technologies) and passaged with TrypLE Express (Life Technologies) plus Y27632 ROCK inhibitor (Stemgent). For knockdown experiments, hPSCs were transduced in suspension with lentivirus for UPBAP2L-shRNA3, UPBAP2L-shRNA4, or non-targeting shRNA. Two days post-transduction, cells were selected and then maintained with 5 µg/ml puromycin. Excitatory cortical neurons were generated as described (99). Briefly, on day 1, hPSCs expressing doxycycline-inducible Ngn2 were seeded at 40,000 cells/cm^2 and 8 h later, Ngn2 expression was induced with doxycycline in KSR media supplemented with patterning molecules. On day two, media was changed to a 50:50 mix of KSR and N2 with Zeocin at 5 µg/ml (Thermo Fisher) to select for cells expressing Ngn2. On day three, media was changed to N2 plus B27 with Zeocin. On day 4, cells were re-plated and then maintained in NBM plus B27 and neurotrophic factors. Doxycycline was maintained in the media throughout the differentiation process. On day 14, neurons were treated with 100 µg/ml cycloheximide (Sigma) for 3 minutes, washed with ice-cold PBS plus cycloheximide, spun down, aspirated and flash frozen.

3.5.10 Lentiviral shRNA treatment and cortical differentiation of human iPSC

The CV-B iPSC line was edited by CRISPR-Cas9 to generated FMRP knockout clones. The iPSC lines were maintained in mTeSR medium (Stem Cell Technologies) and grown on Matrigel coated plates (Corning). iPSCs were first transduced in suspension with lentivirus for UBAP2L-shRNA3, UBAP2L-shRNA4, or non-targeting shRNA for 24 hours, and recovered for 2 days. Then, the iPSCs were subjected to a second transduction for 24 hours. After 2 days recovery, cells were selected and maintained with 2 µg/ml puromycin for 3 days. Cortical neurons were generated by using an established protocol (103). Briefly, on day 0 iPSCs were

plated on double Matrigel-coated plates. On day 1 cells were differentiated in N2B27 medium with 2 μM XAV939, 100 nM LDN-193189, and 10 μM SB431542 for 2 days. Concentrations of small molecules were gradually reduced until day 14. On day 15 cells were re-plated on Matrigel-coated plates and maintained in N2B27 media with 10 ng/ml brain-derived neurotrophic factor (BDNF) for 15 days. On day 30 cortical neurons were plated on Matrigel coated plates or glass slides and maintained in N2B27 medium for maturation.

3.5.11 *Drosophila* stocks and genetic crosses

Flies were raised at 25 °C on a standard food medium. The $dFmr1^{4}$ and d$Fmr1^{3}$ mutant fly lines were generated previously(*78*, *79*). Other fly lines used in this study were w^{1118}, $Lig^{KG08209}$ mutant allele (BL14943), and Lig^{e04268} mutant allele (BL18242) lines and were obtained from the Bloomington *Drosophila* Stock Center. To study genetic interaction between *Lig* and *dFmr1*, we performed genetic crosses and generated $Lig^{KG08209}$/*CyO*; $dFmr1^{3}$/*TM6B, Tb* and Lig^{e04268}/*CyO*; $dFmr1^{4}$/*TM6B, Tb* fly lines. These lines were crossed to each other and mushroom body morphology of adult male and female flies was examined at the age of 1-2 days old.

3.5.12 Analysis of *Drosophila* mushroom body morphology

Pharate adult flies of 1-2 days old were collected and fixed with 4% paraformaldehyde, 0.1% Triton X-100 in PBS for 3 hours at room temperature. After fixation, fly heads were washed two times and brains were dissected in wash buffer (0.5% Triton X-100 in PBS). The whole brain samples were incubated with blocking solution (5% NGS, 2% BSA, 0.5% Triton

X-100) for 3 hours at room temperature. The whole brains were then incubated with anti-Fasciclin II antibody (1D4-supernatant; Developmental Studies Hybridoma Bank) at a 1:10 dilution in blocking solution for 48 hours at 4°C. The whole brains were washed three times for 15 min each at room temperature with wash buffer and then incubated with secondary antibody (goat anti-mouse Alexa 488; Invitrogen) at a 1:1,000 dilution in blocking solution for 24 hours at 4 °C. The samples were rinsed three times for 15 min each at room temperature and mounted with Vectashield containing 4',6-diamidino-2-phenylindole (DAPI, Vector laboratories) for imaging. Statistical test of the MB β-lobe morphology was performed using Prism 7 (GraphPad software, Inc. 2007) and the results were analyzed by a χ^2 test.

3.5.13 Immunocytochemistry

Cortical neurons were fixed with 4% paraformaldehyde (PFA) for 10 min at room temperature. Neurons were permeabilized with PBST (1 x PBS solution with 0.1% Triton X-100) and blocked with blocking buffer (5% goat serum in PBST for 1 hour at room temperature. Cells were then incubated with appropriately diluted primary antibodies in blocking buffer for overnight at 4 °C and washed with PBST 3 times for 5 min each at room temperature and then incubated with desired secondary antibodies in blocking buffer for 1 hour. After staining cells were washed again with PBST 3 times for 5 min each at room temperature. Staining of the nuclei with DAPI was performed with mounting solution. Protein was detected with following antibodies: anti-CTIP2 antibodies (MABE1045, Millipore), and anti-TBR1 antibodies (ab31940, Abcam). Images were captured on a Zeiss LSM 710 confocal microscope and processed with ImageJ/Fiji.

3.5.14 Multi-electrode array recording

Cortical neurons at day 30 of differentiation were dissociated and 2×10^5 cells were plated on poly-L-ornithine and laminin-coated multi-electrode arrays (Axion Biosystems). Spontaneous activities were recorded during a 5 min period on the day 50 of differentiation. Four replicates were performed for each sample of neurons.

3.6 Acknowledgements

We thank members of the Yeo lab, in particular Emily Wheeler and Florian Krach, for helpful discussions. We thank Joel Richter, and Ravi Muddashetty for valuable comments on the project and sharing reagents. E.-C.L. was supported by a study-abroad fellowship from the Taiwanese government. This work was supported by grants from the NIH (R01NS075449, R01HG004659, U19MH107367, and U54HG007005) to G.W.Y., and R21MH109761 to L.E.B.

Chapter 3 is currently preparing as a manuscript by **Luo EC**, Lee S, Kosmaczewski SG, Shankar A, Barrett LE, Gao FB, Aigner S, Yeo GW. "UBAP2L is a functional antagonist to fragile X mental retardation protein in fragile X syndrome". The dissertation author is the primary author of this manuscript.

References

1. Van Nostrand EL, Pratt GA, Shishkin AA, Gelboin-Burkhart C, Fang MY, Sundararaman B, Blue SM, Nguyen TB, Surka C, Elkins K, Stanton R, Rigo F, Guttman M, Yeo GW. Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). Nat Methods. 2016;13(6):508-14. doi: 10.1038/nmeth.3810. PubMed PMID: 27018577; PMCID: PMC4887338.

2. Wheeler EC, Van Nostrand EL, Yeo GW. Advances and challenges in the detection of transcriptome-wide protein-RNA interactions. Wiley Interdiscip Rev RNA. 2018;9(1). Epub 2017/08/31. doi: 10.1002/wrna.1436. PubMed PMID: 28853213; PMCID: PMC5739989.

3. Licatalosi DD, Mele A, Fak JJ, Ule J, Kayikci M, Chi SW, Clark TA, Schweitzer AC, Blume JE, Wang X, Darnell JC, Darnell RB. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. Nature. 2008;456(7221):464-9. Epub 2008/11/04. doi: 10.1038/nature07488. PubMed PMID: 18978773; PMCID: PMC2597294.

4. Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano M, Jr., Jungkamp AC, Munschauer M, Ulrich A, Wardle GS, Dewell S, Zavolan M, Tuschl T. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell. 2010;141(1):129-41. Epub 2010/04/08. doi: 10.1016/j.cell.2010.03.009. PubMed PMID: 20371350; PMCID: PMC2861495.

5. Konig J, Zarnack K, Rot G, Curk T, Kayikci M, Zupan B, Turner DJ, Luscombe NM, Ule J. iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. Nat Struct Mol Biol. 2010;17(7):909-15. Epub 2010/07/06. doi: 10.1038/nsmb.1838. PubMed PMID: 20601959; PMCID: PMC3000544.

6. Lukong KE, Chang KW, Khandjian EW, Richard S. RNA-binding proteins in human genetic disease. Trends Genet. 2008;24(8):416-25. doi: 10.1016/j.tig.2008.05.004. PubMed PMID: 18597886.

7. Baltz AG, Munschauer M, Schwanhausser B, Vasile A, Murakawa Y, Schueler M, Youngs N, Penfold-Brown D, Drew K, Milek M, Wyler E, Bonneau R, Selbach M, Dieterich C, Landthaler M. The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. Mol Cell. 2012;46(5):674-90. doi: 10.1016/j.molcel.2012.05.021. PubMed PMID: 22681889.

8. Castello A, Fischer B, Eichelbaum K, Horos R, Beckmann BM, Strein C, Davey NE, Humphreys DT, Preiss T, Steinmetz LM, Krijgsveld J, Hentze MW. Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell. 2012;149(6):1393-406. doi: 10.1016/j.cell.2012.04.031. PubMed PMID: 22658674.

9. Gerstberger S, Hafner M, Ascano M, Tuschl T. Evolutionary conservation and expression of human RNA-binding proteins and their role in human genetic disease. Adv Exp Med Biol. 2014;825:1-55. doi: 10.1007/978-1-4939-1221-6_1. PubMed PMID: 25201102; PMCID: PMC4180674.

10. Gerstberger S, Hafner M, Tuschl T. A census of human RNA-binding proteins. Nat Rev Genet. 2014;15(12):829-45. Epub 2014/11/05. doi: 10.1038/nrg3813. PubMed PMID: 25365966.

11. Trendel J, Schwarzl T, Horos R, Prakash A, Bateman A, Hentze MW, Krijgsveld J. The Human RNA-Binding Proteome and Its Dynamics during Translational Arrest. Cell. 2019;176(1-2):391-403 e19. Epub 2018/12/12. doi: 10.1016/j.cell.2018.11.004. PubMed PMID: 30528433.

12. Queiroz RML, Smith T, Villanueva E, Marti-Solano M, Monti M, Pizzinga M, Mirea DM, Ramakrishna M, Harvey RF, Dezi V, Thomas GH, Willis AE, Lilley KS. Author Correction: Comprehensive identification of RNA-protein interactions in any organism using orthogonal organic phase separation (OOPS). Nat Biotechnol. 2019;37(6):692. Epub 2019/05/28. doi: 10.1038/s41587-019-0161-8. PubMed PMID: 31123356.

13. Urdaneta EC, Vieira-Vieira CH, Hick T, Wessels HH, Figini D, Moschall R, Medenbach J, Ohler U, Granneman S, Selbach M, Beckmann BM. Purification of cross-linked RNA-protein complexes by phenol-toluol extraction. Nat Commun. 2019;10(1):990. Epub 2019/03/03. doi: 10.1038/s41467-019-08942-3. PubMed PMID: 30824702; PMCID: PMC6397201.

14. Graindorge A, Pinheiro I, Nawrocka A, Mallory AC, Tsvetkov P, Gil N, Carolis C, Buchholz F, Ulitsky I, Heard E, Taipale M, Shkumatava A. In-cell identification and measurement of RNA-protein interactions. Nat Commun. 2019;10(1):5317. Epub 2019/11/24. doi: 10.1038/s41467-019-13235-w. PubMed PMID: 31757954; PMCID: PMC6876571.

15. Beckmann BM, Horos R, Fischer B, Castello A, Eichelbaum K, Alleaume AM, Schwarzl T, Curk T, Foehr S, Huber W, Krijgsveld J, Hentze MW. The RNA-binding proteomes from yeast to man harbour conserved enigmRBPs. Nat Commun. 2015;6:10127. Epub 2015/12/04. doi: 10.1038/ncomms10127. PubMed PMID: 26632259; PMCID: PMC4686815.

16. Kapeli K, Pratt GA, Vu AQ, Hutt KR, Martinez FJ, Sundararaman B, Batra R, Freese P, Lambert NJ, Huelga SC, Chun SJ, Liang TY, Chang J, Donohue JP, Shiue L, Zhang J, Zhu H, Cambi F, Kasarskis E, Hoon S, Ares M, Jr., Burge CB, Ravits J, Rigo F, Yeo GW. Distinct and shared functions of ALS-associated proteins TDP-43, FUS and TAF15 revealed by multisystem analyses. Nat Commun. 2016;7:12143. Epub 2016/07/06. doi: 10.1038/ncomms12143. PubMed PMID: 27378374; PMCID: PMC4935974.

17. Martinez FJ, Pratt GA, Van Nostrand EL, Batra R, Huelga SC, Kapeli K, Freese P, Chun SJ, Ling K, Gelboin-Burkhart C, Fijany L, Wang HC, Nussbacher JK, Broski SM, Kim HJ, Lardelli R, Sundararaman B, Donohue JP, Javaherian A, Lykke-Andersen J, Finkbeiner S, Bennett CF, Ares M, Jr., Burge CB, Taylor JP, Rigo F, Yeo GW. Protein-RNA Networks Regulated by Normal and ALS-Associated Mutant HNRNPA2B1 in the Nervous System. Neuron. 2016;92(4):780-95. Epub 2016/10/25. doi: 10.1016/j.neuron.2016.09.050. PubMed PMID: 27773581; PMCID: PMC5123850.

18. Van Nostrand EL, Huelga SC, Yeo GW. Experimental and Computational Considerations in the Study of RNA-Binding Protein-RNA Interactions. Adv Exp Med Biol. 2016;907:1-28. Epub 2016/06/04. doi: 10.1007/978-3-319-29073-7_1. PubMed PMID: 27256380; PMCID: PMC5522005.

19. Van Nostrand EL, Freese P, Pratt GA, Wang X, Wei X, Xiao R, Blue SM, Dominguez D, Cody NA, Olson S, Sundararaman B, Zhan L, Bazile C, Bouvrette LPB, Chen J, Duff MO, Garcia KE, Gelboin-Burkhart C, Hochman A, Lambert NJ, Li H, Nguyen TB, Palden T, Rabano I, Sathe S, Stanton R, Bergalet J, Zhou B, Su A, Wang R, Yee BA, Louie AL, Aigner S, Fu XD, Lecuyer E, Burge CB, Graveley BR, Yeo GW. A Large-Scale Binding and Functional Map of Human RNA Binding Proteins. Nature. 2019;in press. doi: http://www.biorxiv.org/lookup/doi/10.1101/179648.

20. Ule J, Jensen KB, Ruggiu M, Mele A, Ule A, Darnell RB. CLIP identifies Nova-regulated RNA networks in the brain. Science. 2003;302(5648):1212-5. doi: 10.1126/science.1090095. PubMed PMID: 14615540.

21. Ule J, Stefani G, Mele A, Ruggiu M, Wang X, Taneri B, Gaasterland T, Blencowe BJ, Darnell RB. An RNA map predicting Nova-dependent splicing regulation. Nature. 2006;444(7119):580-6. Epub 2006/10/27. doi: 10.1038/nature05304. PubMed PMID: 17065982.

22. Coller JM, Gray NK, Wickens MP. mRNA stabilization by poly(A) binding protein is independent of poly(A) and requires translation. Genes Dev. 1998;12(20):3226-35. PubMed PMID: 9784497; PMCID: PMC317214.

23. Coller J, Wickens M. Tethered function assays: an adaptable approach to study RNA regulatory proteins. Methods Enzymol. 2007;429:299-321. doi: 10.1016/S0076-6879(07)29014-7. PubMed PMID: 17913629.

24. Bos TJ, Nussbacher JK, Aigner S, Yeo GW. Tethered Function Assays as Tools to Elucidate the Molecular Roles of RNA-Binding Proteins. Adv Exp Med Biol. 2016;907:61-88. Epub 2016/06/04. doi: 10.1007/978-3-319-29073-7_3. PubMed PMID: 27256382; PMCID: PMC5569382.

25. Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J, Heger A, Holm L, Sonnhammer EL, Eddy SR, Bateman A, Finn RD.

The Pfam protein families database. Nucleic Acids Res. 2012;40(Database issue):D290-301. Epub 2011/12/01. doi: 10.1093/nar/gkr1065. PubMed PMID: 22127870; PMCID: PMC3245129.

26. Attwood TK, Bradley P, Flower DR, Gaulton A, Maudling N, Mitchell AL, Moulton G, Nordle A, Paine K, Taylor P, Uddin A, Zygouri C. PRINTS and its automatic supplement, prePRINTS. Nucleic Acids Res. 2003;31(1):400-2. Epub 2003/01/10. PubMed PMID: 12520033; PMCID: PMC165477.

27. Lykke-Andersen J, Wagner E. Recruitment and activation of mRNA decay enzymes by two ARE-mediated decay activation domains in the proteins TTP and BRF-1. Genes Dev. 2005;19(3):351-61. doi: 10.1101/gad.1282305. PubMed PMID: 15687258; PMCID: PMC546513.

28. Benjamini Y, Krieger AM, Yekutieli D. Adaptive linear step-up procedures that control the false discovery rate. Biometrika. 2006;93(3):491-507. doi: DOI 10.1093/biomet/93.3.491. PubMed PMID: WOS:000241271700001.

29. Nishimura T, Padamsi Z, Fakim H, Milette S, Dunham WH, Gingras AC, Fabian MR. The eIF4E-Binding Protein 4E-T Is a Component of the mRNA Decay Machinery that Bridges the 5' and 3' Termini of Target mRNAs. Cell Rep. 2015;11(9):1425-36. Epub 2015/06/02. doi: 10.1016/j.celrep.2015.04.065. PubMed PMID: 26027925.

30. Wang X, Lu Z, Gomez A, Hon GC, Yue Y, Han D, Fu Y, Parisien M, Dai Q, Jia G, Ren B, Pan T, He C. N6-methyladenosine-dependent regulation of messenger RNA stability. Nature. 2014;505(7481):117-20. doi: 10.1038/nature12730. PubMed PMID: 24284625; PMCID: PMC3877715.

31. Garneau NL, Wilusz J, Wilusz CJ. The highways and byways of mRNA decay. Nat Rev Mol Cell Biol. 2007;8(2):113-26. doi: 10.1038/nrm2104. PubMed PMID: 17245413.

32. Hu W, Yuan B, Lodish HF. Cpeb4-mediated translational regulatory circuitry controls terminal erythroid differentiation. Dev Cell. 2014;30(6):660-72. Epub 2014/09/16. doi: 10.1016/j.devcel.2014.07.008. PubMed PMID: 25220394; PMCID: PMC4182162.

33. Vicens Q, Kieft JS, Rissland OS. Revisiting the Closed-Loop Model and the Nature of mRNA 5'-3' Communication. Mol Cell. 2018;72(5):805-12. Epub 2018/12/12. doi: 10.1016/j.molcel.2018.10.047. PubMed PMID: 30526871; PMCID: PMC6294470.

34. Rissland OS. The organization and regulation of mRNA-protein complexes. Wiley Interdiscip Rev RNA. 2017;8(1). Epub 2016/06/22. doi: 10.1002/wrna.1369. PubMed PMID: 27324829; PMCID: PMC5213448.

35. Bicknell AA, Ricci EP. When mRNA translation meets decay. Biochem Soc Trans. 2017;45(2):339-51. Epub 2017/04/15. doi: 10.1042/BST20160243. PubMed PMID: 28408474.

36. Radhakrishnan A, Green R. Connections Underlying Translation and mRNA Stability. J Mol Biol. 2016;428(18):3558-64. Epub 2016/06/05. doi: 10.1016/j.jmb.2016.05.025. PubMed PMID: 27261255.

37. Roy B, Jacobson A. The intimate relationships of mRNA decay and translation. Trends Genet. 2013;29(12):691-9. Epub 2013/10/05. doi: 10.1016/j.tig.2013.09.002. PubMed PMID: 24091060; PMCID: PMC3854950.

38. Fu XF, Cheng SF, Wang LQ, Yin S, De Felici M, Shen W. DAZ Family Proteins, Key Players for Germ Cell Development. Int J Biol Sci. 2015;11(10):1226-35. Epub 2015/09/04. doi: 10.7150/ijbs.11536. PubMed PMID: 26327816; PMCID: PMC4551758.

39. Rosario R, Childs AJ, Anderson RA. RNA-binding proteins in human oogenesis: Balancing differentiation and self-renewal in the female fetal germline. Stem Cell Res. 2017;21:193-201. Epub 2017/04/25. doi: 10.1016/j.scr.2017.04.008. PubMed PMID: 28434825; PMCID: PMC5446320.

40. Julaton VT, Reijo Pera RA. NANOS3 function in human germ cell development. Hum Mol Genet. 2011;20(11):2238-50. Epub 2011/03/23. doi: 10.1093/hmg/ddr114. PubMed PMID: 21421998; PMCID: PMC3090199.

41. Youn JY, Dunham WH, Hong SJ, Knight JDR, Bashkurov M, Chen GI, Bagci H, Rathod B, MacLeod G, Eng SWM, Angers S, Morris Q, Fabian M, Cote JF, Gingras AC. High-Density Proximity Mapping Reveals the Subcellular Organization of mRNA-Associated Granules and Bodies. Mol Cell. 2018;69(3):517-32 e11. Epub 2018/02/06. doi: 10.1016/j.molcel.2017.12.020. PubMed PMID: 29395067.

42. Markmiller S, Soltanieh S, Server KL, Mak R, Jin W, Fang MY, Luo EC, Krach F, Yang D, Sen A, Fulzele A, Wozniak JM, Gonzalez DJ, Kankel MW, Gao FB, Bennett EJ, Lecuyer E, Yeo GW. Context-Dependent and Disease-Specific Diversity in Protein Interactions within Stress Granules. Cell. 2018;172(3):590-604 e13. Epub 2018/01/27. doi: 10.1016/j.cell.2017.12.032. PubMed PMID: 29373831; PMCID: PMC5969999.

43. Huang C, Chen Y, Dai H, Zhang H, Xie M, Zhang H, Chen F, Kang X, Bai X, Chen Z. UBAP2L arginine methylation by PRMT1 modulates stress granule assembly. Cell Death Differ. 2020;27(1):227-41. Epub 2019/05/23. doi: 10.1038/s41418-019-0350-5. PubMed PMID: 31114027.

44. Deragon JM, Bousquet-Antonelli C. The role of LARP1 in translation and beyond. Wiley Interdiscip Rev RNA. 2015;6(4):399-417. Epub 2015/04/22. doi: 10.1002/wrna.1282. PubMed PMID: 25892282.

45. Gray GA, Gray NK. A tail of translational regulation. Elife. 2017;6. Epub 2017/06/28. doi: 10.7554/eLife.29104. PubMed PMID: 28653621; PMCID: PMC5487164.

46. Andreev DE, Dmitriev SE, Loughran G, Terenin IM, Baranov PV, Shatsky IN. Translation control of mRNAs encoding mammalian translation initiation factors. Gene. 2018;651:174-82. Epub 2018/02/08. doi: 10.1016/j.gene.2018.02.013. PubMed PMID: 29414693.

47. Hortsch M, Griffiths G, Meyer DI. Restriction of docking protein to the rough endoplasmic reticulum: immunocytochemical localization in rat liver. Eur J Cell Biol. 1985;38(2):271-9. Epub 1985/09/01. PubMed PMID: 4043092.

48. Gehman LT, Stoilov P, Maguire J, Damianov A, Lin CH, Shiue L, Ares M, Jr., Mody I, Black DL. The splicing regulator Rbfox1 (A2BP1) controls neuronal excitation in the mammalian brain. Nat Genet. 2011;43(7):706-11. Epub 2011/05/31. doi: 10.1038/ng.841. PubMed PMID: 21623373; PMCID: PMC3125461.

49. Duncan PI, Stojdl DF, Marius RM, Scheit KH, Bell JC. The Clk2 and Clk3 dual-specificity protein kinases regulate the intranuclear distribution of SR proteins and influence pre-mRNA splicing. Exp Cell Res. 1998;241(2):300-8. Epub 1998/06/25. doi: 10.1006/excr.1998.4083. PubMed PMID: 9637771.

50. Scotti MM, Swanson MS. RNA mis-splicing in disease. Nat Rev Genet. 2016;17(1):19-32. Epub 2015/11/26. doi: 10.1038/nrg.2015.3. PubMed PMID: 26593421; PMCID: PMC5993438.

51. Chen CY, Xu N, Zhu W, Shyu AB. Functional dissection of hnRNP D suggests that nuclear import is required before hnRNP D can modulate mRNA turnover in the cytoplasm. RNA. 2004;10(4):669-80. Epub 2004/03/24. doi: 10.1261/rna.5269304. PubMed PMID: 15037776; PMCID: PMC1370557.

52. Luna R, Rondon AG, Aguilera A. New clues to understand the role of THO and other functionally related factors in mRNP biogenesis. Biochim Biophys Acta. 2012;1819(6):514-20. Epub 2011/12/31. doi: 10.1016/j.bbagrm.2011.11.012. PubMed PMID: 22207203.

53. Sundararaman B, Zhan L, Blue SM, Stanton R, Elkins K, Olson S, Wei X, Van Nostrand EL, Pratt GA, Huelga SC, Smalec BM, Wang X, Hong EL, Davidson JM, Lecuyer E, Graveley BR, Yeo GW. Resources for the Comprehensive Discovery of Functional RNA Elements. Mol Cell. 2016;61(6):903-13. Epub 2016/03/19. doi: 10.1016/j.molcel.2016.02.012. PubMed PMID: 26990993; PMCID: PMC4839293.

54. Miyasaka T, Morita M, Ito K, Suzuki T, Fukuda H, Takeda S, Inoue J, Semba K, Yamamoto T. Interaction of antiproliferative protein Tob with the CCR4-NOT deadenylase complex. Cancer Sci. 2008;99(4):755-61. doi: 10.1111/j.1349-7006.2008.00746.x. PubMed PMID: 18377426.

55. Lovci MT, Ghanem D, Marr H, Arnold J, Gee S, Parra M, Liang TY, Stark TJ, Gehman LT, Hoon S, Massirer KB, Pratt GA, Black DL, Gray JW, Conboy JG, Yeo GW. Rbfox proteins

regulate alternative mRNA splicing through evolutionarily conserved RNA bridges. Nat Struct Mol Biol. 2013;20(12):1434-42. doi: 10.1038/nsmb.2699. PubMed PMID: 24213538; PMCID: PMC3918504.

56. Lee DSM, Ghanem LR, Barash Y. Integrative analysis reveals RNA G-quadruplexes in UTRs are selectively constrained and enriched for functional associations. Nat Commun. 2020;11(1):527. Epub 2020/01/29. doi: 10.1038/s41467-020-14404-y. PubMed PMID: 31988292; PMCID: PMC6985247.

57. Urano J, Fox MS, Reijo Pera RA. Interaction of the conserved meiotic regulators, BOULE (BOL) and PUMILIO-2 (PUM2). Mol Reprod Dev. 2005;71(3):290-8. Epub 2005/04/05. doi: 10.1002/mrd.20270. PubMed PMID: 15806553.

58. Moore FL, Jaruzelska J, Fox MS, Urano J, Firpo MT, Turek PJ, Dorfman DM, Pera RA. Human Pumilio-2 is expressed in embryonic stem cells and germ cells and interacts with DAZ (Deleted in AZoospermia) and DAZ-like proteins. Proc Natl Acad Sci U S A. 2003;100(2):538-43. Epub 2003/01/04. doi: 10.1073/pnas.0234478100. PubMed PMID: 12511597; PMCID: PMC141031.

59. Schmidt EK, Clavarino G, Ceppi M, Pierre P. SUnSET, a nonradioactive method to monitor protein synthesis. Nat Methods. 2009;6(4):275-7. Epub 2009/03/24. doi: 10.1038/nmeth.1314. PubMed PMID: 19305406.

60. Liu B, Qian SB. Characterizing inactive ribosomes in translational profiling. Translation (Austin). 2016;4(1):e1138018. Epub 2016/06/24. doi: 10.1080/21690731.2015.1138018. PubMed PMID: 27335722; PMCID: PMC4909419.

61. Batra R, Nelles DA, Pirie E, Blue SM, Marina RJ, Wang H, Chaim IA, Thomas JD, Zhang N, Nguyen V, Aigner S, Markmiller S, Xia G, Corbett KD, Swanson MS, Yeo GW. Elimination of Toxic Microsatellite Repeat Expansion RNA by RNA-Targeting Cas9. Cell. 2017;170(5):899-912 e10. Epub 2017/08/15. doi: 10.1016/j.cell.2017.07.010. PubMed PMID: 28803727; PMCID: PMC5873302.

62. Nelles DA, Fang MY, O'Connell MR, Xu JL, Markmiller SJ, Doudna JA, Yeo GW. Programmable RNA Tracking in Live Cells with CRISPR/Cas9. Cell. 2016;165(2):488-96. doi: 10.1016/j.cell.2016.02.054. PubMed PMID: 26997482; PMCID: PMC4826288.

63. Fujii K, Susanto TT, Saurabh S, Barna M. Decoding the Function of Expansion Segments in Ribosomes. Mol Cell. 2018;72(6):1013-20 e6. Epub 2018/12/24. doi: 10.1016/j.molcel.2018.11.023. PubMed PMID: 30576652; PMCID: PMC6407129.

64. Van Nostrand EL, Freese P, Pratt GA, Wang X, Wei X, Xiao R, Blue SM, Dominguez D, Cody NA, Olson S, Sundararaman B, Zhan L, Bazile C, Bouvrette LPB, Chen J, Duff MO, Garcia KE, Gelboin-Burkhart C, Hochman A, Lambert NJ, Li H, Nguyen TB, Palden T, Rabano I, Sathe S, Stanton R, Bergalet J, Zhou B, Su A, Wang R, Yee BA, Louie AL, Aigner S, Fu XD,

Lecuyer E, Burge CB, Graveley BR, Yeo GW. A Large-Scale Binding and Functional Map of Human RNA Binding Proteins. Nature. in press. doi: http://www.biorxiv.org/lookup/doi/10.1101/179648.

65. Maeda M, Hasegawa H, Sugiyama M, Hyodo T, Ito S, Chen D, Asano E, Masuda A, Hasegawa Y, Hamaguchi M, Senga T. Arginine methylation of ubiquitin-associated protein 2-like is required for the accurate distribution of chromosomes. FASEB J. 2016;30(1):312-23. doi: 10.1096/fj.14-268987. PubMed PMID: 26381755.

66. Natchiar SK, Myasnikov AG, Kratzat H, Hazemann I, Klaholz BP. Visualization of chemical modifications in the human 80S ribosome structure. Nature. 2017;551(7681):472-7. Epub 2017/11/17. doi: 10.1038/nature24482. PubMed PMID: 29143818.

67. Lutz CS, Cooke C, O'Connor JP, Kobayashi R, Alwine JC. The snRNP-free U1A (SF-A) complex(es): identification of the largest subunit as PSF, the polypyrimidine-tract binding protein-associated splicing factor. RNA. 1998;4(12):1493-9. Epub 1998/12/16. PubMed PMID: 9848648; PMCID: PMC1369720.

68. Boelens WC, Jansen EJ, van Venrooij WJ, Stripecke R, Mattaj IW, Gunderson SI. The human U1 snRNP-specific U1A protein inhibits polyadenylation of its own pre-mRNA. Cell. 1993;72(6):881-92. Epub 1993/03/26. PubMed PMID: 8458082.

69. Fensterl V, Sen GC. Interferon-induced Ifit proteins: their role in viral pathogenesis. J Virol. 2015;89(5):2462-8. doi: 10.1128/JVI.02744-14. PubMed PMID: 25428874; PMCID: PMC4325746.

70. Liang D, Halpert MM, Konduri V, Decker WK. Stepping Out of the Cytosol: AIMp1/p43 Potentiates the Link Between Innate and Adaptive Immunity. Int Rev Immunol. 2015;34(5):367-81. Epub 2015/09/02. doi: 10.3109/08830185.2015.1077829. PubMed PMID: 26325028.

71. Meng X, Thiel KW, Leslie KK. Drug resistance mediated by AEG-1/MTDH/LYRIC. Adv Cancer Res. 2013;120:135-57. Epub 2013/07/31. doi: 10.1016/B978-0-12-401676-7.00005-X. PubMed PMID: 23889990; PMCID: PMC3967868.

72. Meng X, Zhu D, Yang S, Wang X, Xiong Z, Zhang Y, Brachova P, Leslie KK. Cytoplasmic Metadherin (MTDH) provides survival advantage under conditions of stress by acting as RNA-binding protein. J Biol Chem. 2012;287(7):4485-91. Epub 2011/12/27. doi: 10.1074/jbc.C111.291518. PubMed PMID: 22199357; PMCID: PMC3281628.

73. Cox DBT, Gootenberg JS, Abudayyeh OO, Franklin B, Kellner MJ, Joung J, Zhang F. RNA editing with CRISPR-Cas13. Science. 2017;358(6366):1019-27. doi: 10.1126/science.aaq0180. PubMed PMID: 29070703; PMCID: PMC5793859.

74. Coller J, Wickens M. Tethered function assays using 3' untranslated regions. Methods. 2002;26(2):142-50. doi: 10.1016/S1046-2023(02)00016-6. PubMed PMID: 12054890.
75. Brannan KW, Jin W, Huelga SC, Banks CA, Gilmore JM, Florens L, Washburn MP, Van Nostrand EL, Pratt GA, Schwinn MK, Daniels DL, Yeo GW. SONAR Discovers RNA-Binding Proteins from Analysis of Large-Scale Protein-Protein Interactomes. Mol Cell. 2016;64(2):282-93. doi: 10.1016/j.molcel.2016.09.003. PubMed PMID: 27720645; PMCID: PMC5074894.

76. Castello A, Fischer B, Frese CK, Horos R, Alleaume AM, Foehr S, Curk T, Krijgsveld J, Hentze MW. Comprehensive Identification of RNA-Binding Domains in Human Cells. Mol Cell. 2016;63(4):696-710. Epub 2016/07/28. doi: 10.1016/j.molcel.2016.06.029. PubMed PMID: 27453046; PMCID: PMC5003815.

77. Rual JF, Hirozane-Kishikawa T, Hao T, Bertin N, Li S, Dricot A, Li N, Rosenberg J, Lamesch P, Vidalain PO, Clingingsmith TR, Hartley JL, Esposito D, Cheo D, Moore T, Simmons B, Sequerra R, Bosak S, Doucette-Stamm L, Le Peuch C, Vandenhaute J, Cusick ME, Albala JS, Hill DE, Vidal M. Human ORFeome version 1.1: a platform for reverse proteomics. Genome Res. 2004;14(10B):2128-35. Epub 2004/10/19. doi: 10.1101/gr.2973604. PubMed PMID: 15489335; PMCID: PMC528929.

78. Consortium" TGO. Expansion of the Gene Ontology knowledgebase and resources. Nucleic Acids Res. 2017;45(D1):D331-D8. doi: 10.1093/nar/gkw1108. PubMed PMID: 27899567; PMCID: PMC5210579.

79. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 2000;25(1):25-9. doi: 10.1038/75556. PubMed PMID: 10802651; PMCID: PMC3037419.

80. Clement SL, Lykke-Andersen J. A tethering approach to study proteins that activate mRNA turnover in human cells. Methods Mol Biol. 2008;419:121-33. doi: 10.1007/978-1-59745-033-1_8. PubMed PMID: 18369979.

81. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15-21. doi: 10.1093/bioinformatics/bts635. PubMed PMID: 23104886; PMCID: PMC3530905.

82. Conway AE, Van Nostrand EL, Pratt GA, Aigner S, Wilbert ML, Sundararaman B, Freese P, Lambert NJ, Sathe S, Liang TY, Essex A, Landais S, Burge CB, Jones DL, Yeo GW. Enhanced CLIP Uncovers IMP Protein-RNA Targets in Human Pluripotent Stem Cells Important for Cell Adhesion and Survival. Cell Rep. 2016;15(3):666-79. Epub 2016/04/14. doi: 10.1016/j.celrep.2016.03.052. PubMed PMID: 27068461; PMCID: PMC4839292.

83. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550. doi: 10.1186/s13059-014-0550-8. PubMed PMID: 25516281; PMCID: PMC4302049.

84. Ascano M, Jr., Mukherjee N, Bandaru P, Miller JB, Nusbaum JD, Corcoran DL, Langlois C, Munschauer M, Dewell S, Hafner M, Williams Z, Ohler U, Tuschl T. FMRP targets distinct mRNA sequence elements to regulate protein expression. Nature. 2012;492(7429):382-6. Epub 2012/12/14. doi: 10.1038/nature11737. PubMed PMID: 23235829; PMCID: PMC3528815.

85. Darnell JC, Van Driesche SJ, Zhang C, Hung KY, Mele A, Fraser CE, Stone EF, Chen C, Fak JJ, Chi SW, Licatalosi DD, Richter JD, Darnell RB. FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. Cell. 2011;146(2):247-61. Epub 2011/07/26. doi: 10.1016/j.cell.2011.06.013. PubMed PMID: 21784246; PMCID: PMC3232425.

86. Pasciuto E, Bagni C. SnapShot: FMRP mRNA targets and diseases. Cell. 2014;158(6):1446- e1. Epub 2014/09/13. doi: 10.1016/j.cell.2014.08.035. PubMed PMID: 25215498.

87. Corbin F, Bouillon M, Fortin A, Morin S, Rousseau F, Khandjian EW. The fragile X mental retardation protein is associated with poly(A)+ mRNA in actively translating polyribosomes. Hum Mol Genet. 1997;6(9):1465-72. Epub 1997/09/01. doi: 10.1093/hmg/6.9.1465. PubMed PMID: 9285783.

88. Stefani G, Fraser CE, Darnell JC, Darnell RB. Fragile X mental retardation protein is associated with translating polyribosomes in neuronal cells. J Neurosci. 2004;24(33):7272-6. Epub 2004/08/20. doi: 10.1523/JNEUROSCI.2306-04.2004. PubMed PMID: 15317853; PMCID: PMC6729764.

89. Osterweil EK, Krueger DD, Reinhold K, Bear MF. Hypersensitivity to mGluR5 and ERK1/2 leads to excessive protein synthesis in the hippocampus of a mouse model of fragile X syndrome. J Neurosci. 2010;30(46):15616-27. Epub 2010/11/19. doi: 10.1523/JNEUROSCI.3888-10.2010. PubMed PMID: 21084617; PMCID: PMC3400430.

90. Bhat A, Andersen PL, Qin Z, Xiao W. Rev3, the catalytic subunit of Polzeta, is required for maintaining fragile site stability in human cells. Nucleic Acids Res. 2013;41(4):2328-39. Epub 2013/01/11. doi: 10.1093/nar/gks1442. PubMed PMID: 23303771; PMCID: PMC3575803.

91. Michel CI, Kraft R, Restifo LL. Defective neuronal development in the mushroom bodies of Drosophila fragile X mental retardation 1 mutants. J Neurosci. 2004;24(25):5798-809. Epub 2004/06/25. doi: 10.1523/JNEUROSCI.1102-04.2004. PubMed PMID: 15215302.

92. Liu XS, Wu H, Krzisch M, Wu X, Graef J, Muffat J, Hnisz D, Li CH, Yuan B, Xu C, Li Y, Vershkov D, Cacace A, Young RA, Jaenisch R. Rescue of Fragile X Syndrome Neurons by DNA Methylation Editing of the FMR1 Gene. Cell. 2018;172(5):979-92 e6. Epub 2018/02/20. doi: 10.1016/j.cell.2018.01.012. PubMed PMID: 29456084; PMCID: PMC6375087.

93. Pasciuto E, Bagni C. SnapShot: FMRP interacting proteins. Cell. 2014;159(1):218- e1. Epub 2014/09/27. doi: 10.1016/j.cell.2014.08.036. PubMed PMID: 25259928.
94. Chen E, Sharma MR, Shi X, Agrawal RK, Joseph S. Fragile X mental retardation protein regulates translation by binding directly to the ribosome. Mol Cell. 2014;54(3):407-17. Epub 2014/04/22. doi: 10.1016/j.molcel.2014.03.023. PubMed PMID: 24746697; PMCID: PMC4019695.

95. Tran SS, Jun HI, Bahn JH, Azghadi A, Ramaswami G, Van Nostrand EL, Nguyen TB, Hsiao YE, Lee C, Pratt GA, Martinez-Cerdeno V, Hagerman RJ, Yeo GW, Geschwind DH, Xiao X. Widespread RNA editing dysregulation in brains from autistic individuals. Nat Neurosci. 2019;22(1):25-36. Epub 2018/12/19. doi: 10.1038/s41593-018-0287-x. PubMed PMID: 30559470.

96. Qin M, Schmidt KC, Zametkin AJ, Bishu S, Horowitz LM, Burlin TV, Xia Z, Huang T, Quezado ZM, Smith CB. Altered cerebral protein synthesis in fragile X syndrome: studies in human subjects and knockout mice. J Cereb Blood Flow Metab. 2013;33(4):499-507. Epub 2013/01/10. doi: 10.1038/jcbfm.2012.205. PubMed PMID: 23299245; PMCID: PMC3618394.
97. Merkle FT, Ghosh S, Kamitaki N, Mitchell J, Avior Y, Mello C, Kashin S, Mekhoubad S, Ilic D, Charlton M, Saphier G, Handsaker RE, Genovese G, Bar S, Benvenisty N, McCarroll SA, Eggan K. Human pluripotent stem cells recurrently acquire and expand dominant negative P53 mutations. Nature. 2017;545(7653):229-33. Epub 2017/04/27. doi: 10.1038/nature22312. PubMed PMID: 28445466; PMCID: PMC5427175.

98. Hazelbaker DZ, Beccard A, Bara AM, Dabkowski N, Messana A, Mazzucato P, Lam D, Manning D, Eggan K, Barrett LE. A Scaled Framework for CRISPR Editing of Human Pluripotent Stem Cells to Study Psychiatric Disease. Stem Cell Reports. 2017;9(4):1315-27. Epub 2017/10/12. doi: 10.1016/j.stemcr.2017.09.006. PubMed PMID: 29020615; PMCID: PMC5639480.

99. Nehme R, Zuccaro E, Ghosh SD, Li C, Sherwood JL, Pietilainen O, Barrett LE, Limone F, Worringer KA, Kommineni S, Zang Y, Cacchiarelli D, Meissner A, Adolfsson R, Haggarty S, Madison J, Muller M, Arlotta P, Fu Z, Feng G, Eggan K. Combining NGN2 Programming with Developmental Patterning Generates Human Excitatory Neurons with NMDAR-Mediated Synaptic Transmission. Cell Rep. 2018;23(8):2509-23. Epub 2018/05/24. doi: 10.1016/j.celrep.2018.04.066. PubMed PMID: 29791859; PMCID: PMC6003669.

100. Zhang Y, Pak C, Han Y, Ahlenius H, Zhang Z, Chanda S, Marro S, Patzke C, Acuna C, Covy J, Xu W, Yang N, Danko T, Chen L, Wernig M, Sudhof TC. Rapid single-step induction of functional neurons from human pluripotent stem cells. Neuron. 2013;78(5):785-98. Epub 2013/06/15. doi: 10.1016/j.neuron.2013.05.029. PubMed PMID: 23764284; PMCID: PMC3751803.

101. Pak C, Danko T, Zhang Y, Aoto J, Anderson G, Maxeiner S, Yi F, Wernig M, Sudhof TC. Human Neuropsychiatric Disease Modeling using Conditional Deletion Reveals Synaptic

Transmission Defects Caused by Heterozygous Mutations in NRXN1. Cell Stem Cell. 2015;17(3):316-28. Epub 2015/08/19. doi: 10.1016/j.stem.2015.07.017. PubMed PMID: 26279266; PMCID: PMC4560990.

102. Yi F, Danko T, Botelho SC, Patzke C, Pak C, Wernig M, Sudhof TC. Autism-associated SHANK3 haploinsufficiency causes Ih channelopathy in human neurons. Science. 2016;352(6286):aaf2669. Epub 2016/03/12. doi: 10.1126/science.aaf2669. PubMed PMID: 26966193; PMCID: PMC4901875.

103. Kirwan P, Jura M, Merkle FT. Generation and Characterization of Functional Human Hypothalamic Neurons. Curr Protoc Neurosci. 2017;81:3 33 1-3 24. Epub 2017/10/25. doi: 10.1002/cpns.40. PubMed PMID: 29064566.

104. Heisenberg M. Mushroom body memoir: from maps to models. Nat Rev Neurosci. 2003;4(4):266-75. Epub 2003/04/03. doi: 10.1038/nrn1074. PubMed PMID: 12671643.

105. Richter JD, Bassell GJ, Klann E. Dysregulation and restoration of translational homeostasis in fragile X syndrome. Nat Rev Neurosci. 2015;16(10):595-605. Epub 2015/09/10. doi: 10.1038/nrn4001. PubMed PMID: 26350240; PMCID: PMC4688896.

106. Napoli I, Mercaldo V, Boyl PP, Eleuteri B, Zalfa F, De Rubeis S, Di Marino D, Mohr E, Massimi M, Falconi M, Witke W, Costa-Mattioli M, Sonenberg N, Achsel T, Bagni C. The fragile X syndrome protein represses activity-dependent translation through CYFIP1, a new 4E-BP. Cell. 2008;134(6):1042-54. Epub 2008/09/23. doi: 10.1016/j.cell.2008.07.031. PubMed PMID: 18805096.

107. Suhl JA, Chopra P, Anderson BR, Bassell GJ, Warren ST. Analysis of FMRP mRNA target datasets reveals highly associated mRNAs mediated by G-quadruplex structures formed via clustered WGGA sequences. Hum Mol Genet. 2014;23(20):5479-91. Epub 2014/05/31. doi: 10.1093/hmg/ddu272. PubMed PMID: 24876161; PMCID: PMC4168832.

108. Berry-Kravis E, Des Portes V, Hagerman R, Jacquemont S, Charles P, Visootsak J, Brinkman M, Rerat K, Koumaras B, Zhu L, Barth GM, Jaecklin T, Apostol G, von Raison F. Mavoglurant in fragile X syndrome: Results of two randomized, double-blind, placebo-controlled trials. Sci Transl Med. 2016;8(321):321ra5. Epub 2016/01/15. doi: 10.1126/scitranslmed.aab4109. PubMed PMID: 26764156.

109. Youssef EA, Berry-Kravis E, Czech C, Hagerman RJ, Hessl D, Wong CY, Rabbia M, Deptula D, John A, Kinch R, Drewitt P, Lindemann L, Marcinowski M, Langland R, Horn C, Fontoura P, Santarelli L, Quiroz JA, FragXis Study G. Effect of the mGluR5-NAM Basimglurant on Behavior in Adolescents and Adults with Fragile X Syndrome in a Randomized, Double-Blind, Placebo-Controlled Trial: FragXis Phase 2 Results. Neuropsychopharmacology. 2018;43(3):503-12. Epub 2017/08/18. doi: 10.1038/npp.2017.177. PubMed PMID: 28816242; PMCID: PMC5770759.

110. Berry-Kravis EM, Lindemann L, Jonch AE, Apostol G, Bear MF, Carpenter RL, Crawley JN, Curie A, Des Portes V, Hossain F, Gasparini F, Gomez-Mancilla B, Hessl D, Loth E, Scharf SH, Wang PP, Von Raison F, Hagerman R, Spooren W, Jacquemont S. Drug development for neurodevelopmental disorders: lessons learned from fragile X syndrome. Nat Rev Drug Discov. 2018;17(4):280-99. Epub 2017/12/09. doi: 10.1038/nrd.2017.221. PubMed PMID: 29217836.

111. Contractor A, Klyachko VA, Portera-Cailliau C. Altered Neuronal and Circuit Excitability in Fragile X Syndrome. Neuron. 2015;87(4):699-715. Epub 2015/08/21. doi: 10.1016/j.neuron.2015.06.017. PubMed PMID: 26291156; PMCID: PMC4545495.

112. Chang S, Bray SM, Li Z, Zarnescu DC, He C, Jin P, Warren ST. Identification of small molecules rescuing fragile X syndrome phenotypes in Drosophila. Nat Chem Biol. 2008;4(4):256-63. Epub 2008/03/11. doi: 10.1038/nchembio.78. PubMed PMID: 18327252.