A UPF1 variant drives conditional remodeling of nonsense-mediated mRNA decay

Sarah E. Fritz[1], Soumya Ranganathan[1], and J. Robert Hogg[1]*

[1]Biochemistry and Biophysics Center
National Heart, Lung, and Blood Institute
National Institutes of Health
Bethesda, MD 20892
USA

*Correspondence: j.hogg@nih.gov

Abstract

The nonsense-mediated mRNA decay (NMD) pathway monitors translation termination to degrade transcripts with premature stop codons and regulate thousands of human genes. Due to the major role of NMD in RNA quality control and gene expression regulation, it is important to understand how the pathway responds to changing cellular conditions. Here we show that an alternative mammalian-specific isoform of the core NMD factor UPF1, termed $UPF1_{LL}$, enables condition-dependent remodeling of NMD specificity. $UPF1_{LL}$ associates more stably with potential NMD target mRNAs than the major $UPF1_{SL}$ isoform, expanding the scope of NMD to include many transcripts normally immune to the pathway. Unexpectedly, the enhanced persistence of $UPF1_{LL}$ on mRNAs supports induction of NMD in response to rare translation termination events. Thus, while canonical NMD is abolished by translational repression, $UPF1_{LL}$ activity is enhanced, providing a mechanism to rapidly rewire NMD specificity in response to cellular stress.

Introduction

Nonsense-mediated mRNA decay (NMD) is an evolutionarily conserved mRNA quality-control pathway that degrades transcripts undergoing premature translation termination (Lavysh and Neu-Yilik, 2020; Smith and Baker, 2015). In addition, NMD performs a critical regulatory role by governing the turnover of ~5-10% of the transcriptome, including mRNAs with upstream open reading frames (uORFs), introns downstream of the stop codon, or long 3' untranslated regions (UTRs) (Kishor et al., 2019). Despite extensive studies of the large impact of NMD on the transcriptome, the mechanisms by which the pathway selects its regulatory and quality control targets are poorly understood.

A suite of conserved NMD factors acts in concert with general mRNA binding proteins and RNA decay enzymes to identify and degrade target mRNAs. The RNA helicase UPF1 is a central coordinator of the NMD pathway, as it directly binds mRNA and functions at multiple steps in the selection and degradation of target transcripts (Kim and Maquat, 2019). Additional core NMD factors UPF2 and UPF3 promote UPF1 activity and link UPF1 to the exon junction complex (EJC), which strongly stimulates decay (Chamieh et al., 2008; Le Hir et al., 2000, 2000b). In many eukaryotes, NMD execution also depends on the SMG1, 5, 6, and 7 proteins (Causier et al., 2017; Hodgkin et al., 1989; Page et al., 1999; Pulak and Anderson, 1993). Phosphorylation of UPF1 by the SMG1 kinase is a necessary and critical step in promoting the NMD pathway (Kashima et al., 2006), as phosphorylated UPF1 recruits the SMG6 endonuclease and/or general decapping and deadenylation enzymes through the SMG5/7 heterodimer (Eberle et al., 2009; Huntzinger et al., 2008; Loh et al., 2013).

In addition to the functions of specialized NMD proteins, substrate selection and degradation by the NMD pathway requires the translation termination machinery to detect in-frame stop codons (Karousis and Mühlemann, 2018). Although the exact molecular details remain to be elucidated, widely accepted models of NMD state that interactions between core NMD factors and a terminating ribosome are necessary for decay. Because of the strict

dependence of NMD on translation termination, decay efficiency of canonical NMD targets is expected to be tightly linked to translation efficiency. However, in contrast to this prevailing model, there is evidence that NMD efficiency for some targets is actually enhanced during conditions of impaired translation (Martinez-Nunez et al., 2017). These data warrant a more extensive investigation into the role of translation in shaping target specificity by the NMD pathway, particularly during changing physiological conditions.

The ability of UPF1 to bind and hydrolyze ATP is critical to the selection and degradation of potential NMD substrates (Franks et al., 2010; Lee et al., 2015). Specifically, an 11 amino acid regulatory loop in domain 1B of the helicase core functions to regulate the affinity of UPF1 for RNA in the presence of ATP (Chakrabarti et al., 2011; Cheng et al., 2007; Gowravaram et al., 2018). When UPF1 binds and hydrolyzes ATP, this regulatory loop protrudes into the RNA binding channel, causing UPF1 to release RNA. Intriguingly, mammals undergo an alternative splicing event to express two UPF1 isoforms that differ only in length of the regulatory loop (Figure 1A). Almost all NMD studies to date have focused on the more abundant UPF1 “short loop” isoform (designated herein $UPF1_{SL}$), which contains the 11 amino acid regulatory loop that weakens the affinity of UPF1 for RNA in the presence of ATP. Alternative 5’ splice site usage in exon 7 of UPF1 generates a second UPF1 isoform that extends the regulatory loop to 22 amino acids. This naturally occurring UPF1 “long loop” isoform (designated herein $UPF1_{LL}$), which represents ~15-25% of total UPF1 mRNA in diverse cell and tissue types (Figure 1B), has increased catalytic activity and a higher affinity for RNA in the presence of ATP than the $UPF1_{SL}$ isoform (Gowravaram et al., 2018). However, whether the differential biochemical properties of the UPF1 isoforms affect NMD specificity in cells is unknown.

Here, we show that the $UPF1_{LL}$ isoform gives the mammalian NMD pathway the latent ability to remodel NMD target specificity in response to changing physiological conditions. We show that changes to translation efficiency or UPF1 isoform expression conditionally alter which

mRNAs are selected and degraded by the NMD pathway, greatly expanding the scope of NMD in mammalian gene expression control.

Results

$UPF1_{LL}$ contributes to NMD under normal cellular conditions

As the endogenous $UPF1_{LL}$-encoding splice variant accounts for only ~15-25% of total UPF1 mRNA in most human and mouse cells (Figure 1B) (Gowravaram et al., 2018), we first sought to assess the functional significance of basal $UPF1_{LL}$ activity in regulating gene expression. RNA-seq of human HEK-293 cells depleted of the $UPF1_{LL}$ isoform using a short interfering RNA (siRNA) that specifically targeted $UPF1_{LL}$ mRNA levels without perturbing $UPF1_{SL}$ expression (Figure 1C and Figure 1-figure supplement 1A) uncovered a population (~1500 transcripts) that increased in abundance by at least 1.4-fold (Supplementary file 1). To test whether these changes in mRNA abundance were due to activity of $UPF1_{LL}$ in the NMD pathway, we compared our $UPF1_{LL}$-knockdown RNA-seq dataset with a previously published catalog of high-confidence UPF1 and/or SMG6/SMG7 targets (Colombo et al., 2017). Transcripts up-regulated by $UPF1_{LL}$ knockdown comprised a small but significant subpopulation of mRNAs identified as UPF1 and SMG6/SMG7 substrates, indicating that $UPF1_{LL}$ does contribute to the decay of NMD targets (Figure 1C). Analysis of canonical NMD substrates revealed that knockdown of endogenous $UPF1_{LL}$ did not systematically affect the abundance of mRNAs predicted to undergo decay stimulated by the EJC, as only a small number of such transcripts exhibited enhanced expression upon $UPF1_{LL}$ knockdown (Figure 1-figure supplement 1B). Quantitative PCR of select transcripts, including well validated NMD targets (Lareau et al., 2007; Ni et al., 2007), corroborated the transcriptome-wide RNA-seq results (Figure 1-figure supplement 1C).

Figure 1

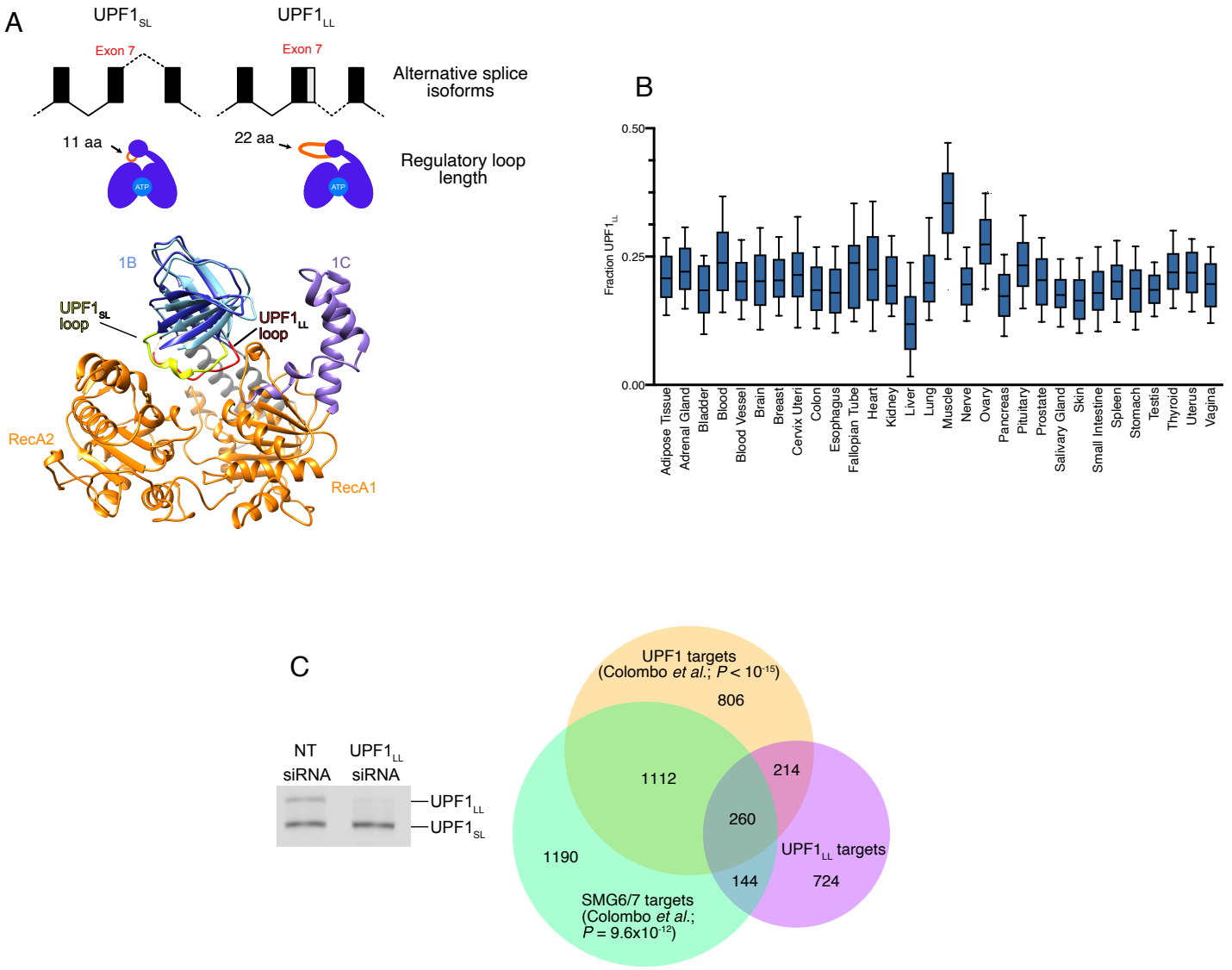

Figure 1: Alternative UPF1$_{LL}$ splice isoform contributes to NMD under normal cellular conditions. (A) (Top) Schematic representation of alternative 5' splice site usage in exon 7 of mammalian UPF1 that results in two UPF1 protein isoforms. **(Bottom)** Ribbon diagram of human UPF1$_{SL}$ helicase core overlaid with that of human UPF1$_{LL}$. The structure of the regulatory loop in domain 1B for each variant is indicated. This figure was generated using Protein Data Bank accessions 2XZP and 6EJ5 (Chakrabarti et al., 2011; Gowravaram et al., 2018). **(B)** Box plot of the fraction of UPF1$_{LL}$ mRNA expressed in indicated human tissues, as determined from the Genotype-Tissue Expression (GTEx) project. Boxes indicate interquartile ranges, and bars indicate 10-90% ranges. **(C) (Left)** Semiquantitative RT-PCR of UPF1$_{SL}$ or UPF1$_{LL}$ transcript levels following transfection of HEK-293 cells with a non-targeting (NT) siRNA or a siRNA that specifically targets the UPF1$_{LL}$ isoform. **(Right)** Venn diagram of overlapping targets from RNA-seq that showed significant stabilization with UPF1$_{LL}$ (this dataset), total UPF1, or SMG6/7 knockdown (Colombo et al., 2017).

$UPF1_{LL}$ overexpression causes downregulation of transcripts with long 3'UTRs

The observation that specific depletion of $UPF1_{LL}$ affected a select subpopulation of mRNAs indicated that it may perform distinct functions from those conducted by the major $UPF1_{SL}$ protein. To gain insight into how the biological activities of the two UPF1 isoforms differ, we engineered HEK-293 stable cell lines to express CLIP-tagged $UPF1_{LL}$ or $UPF1_{SL}$, with a GFP-expressing stable line as a control. We elected to express CLIP-tagged UPF1 constructs in order to enable downstream biochemical analysis of differential UPF1 mRNA binding (Kishor et al., 2020).

Previous work has suggested that SMG1-mediated UPF1 phosphorylation increases with increasing residence time of UPF1 on mRNPs (Durand et al., 2016). Because $UPF1_{LL}$ has been found to have higher affinity for RNA in the presence of ATP than $UPF1_{SL}$ *in vitro* (Gowravaram et al., 2018), we hypothesized that it would more stably associate with mRNPs in cells, in turn promoting its phosphorylation. Consistent with this hypothesis, immunoblotting with a phospho-(S/T)Q antibody revealed that CLIP-$UPF1_{LL}$ was more highly phosphorylated than CLIP-$UPF1_{SL}$ (Figure 2A). The two splice variants exhibited a similar capacity to be phosphorylated, as treatment of cells with puromycin increased the phosphorylation of both UPF1 proteins to equivalent levels, perhaps due to reduced displacement by elongating ribosomes (Dang et al., 2009; Zünd et al., 2013).

To analyze the target specificity of the UPF1 splice variants, we performed RNA-seq on CLIP-UPF1 and control cells (Supplementary file 2). Previous studies have indicated that overexpression of $UPF1_{SL}$ does not enhance overall NMD activity, potentially due to feedback control by the pathway (Huang et al., 2011). Consistent with these findings, analysis of mRNAs previously observed to be up-regulated upon total UPF1 knockdown (Baird et al., 2018) revealed that ~5 to 10-fold stable overexpression of CLIP-$UPF1_{SL}$ had a negligible impact on the expression of previously identified NMD targets (Figure 2-figure supplement 1A). Expression of CLIP-$UPF1_{LL}$ caused small increases in the levels of some previously identified NMD targets but

Figure 2

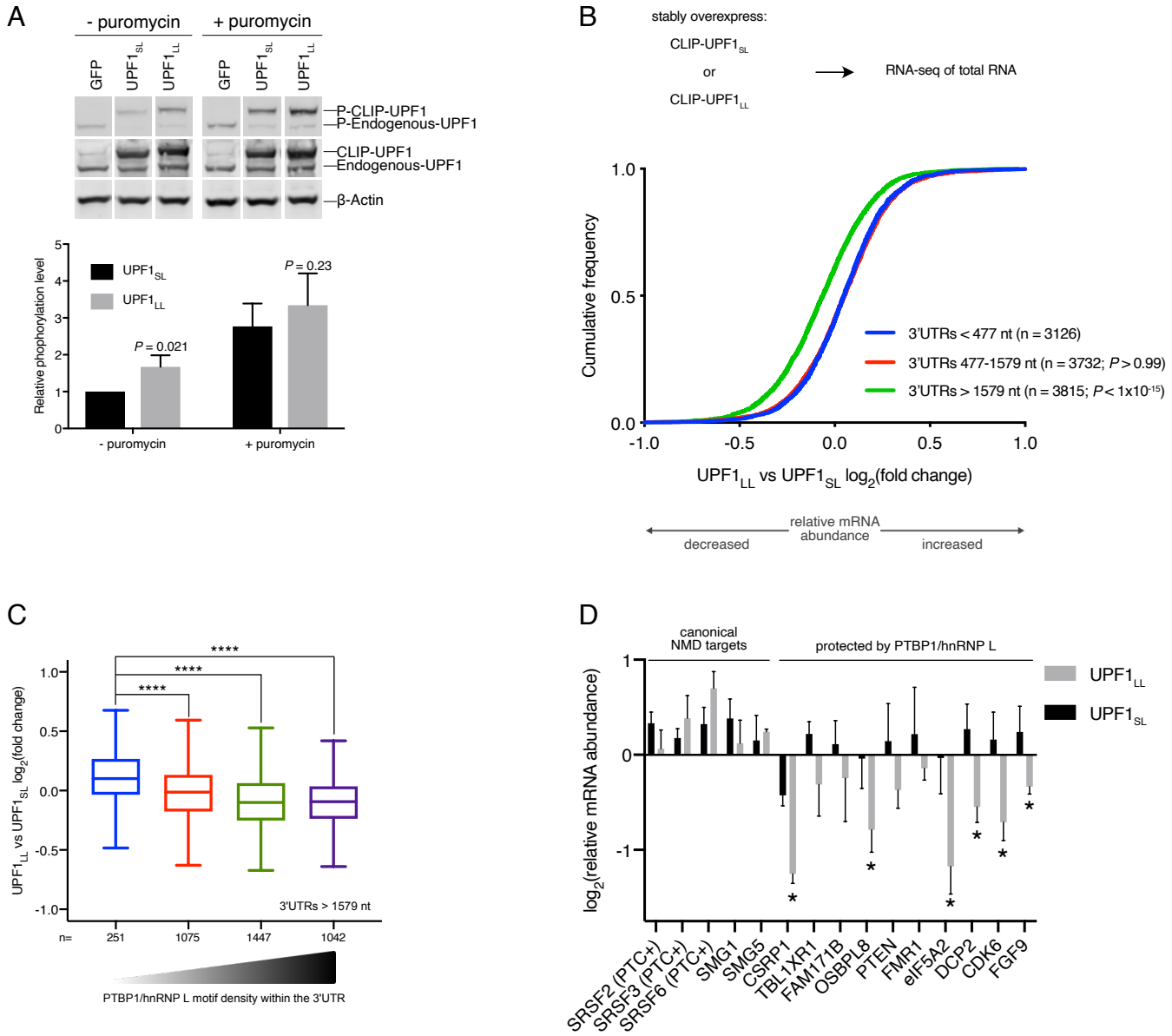

Figure 2: $UPF1_{LL}$ targets long 3'UTRs normally protected from NMD. **(A)** Western blots of CLIP-$UPF1_{SL}$ and CLIP-$UPF1_{LL}$ phosphorylation status at steady state (- puromycin) or following incubation with puromycin (+ puromycin). Graph below represents the relative phosphorylation level of each CLIP-UPF1 loop variant, as determined by dividing the phosphorylation signal for each protein (P-CLIP-UPF1) by that of its corresponding total amount (CLIP-UPF1) and normalizing to CLIP-$UPF1_{SL}$ at steady state. Error bars indicate SD (n = 3). Statistical significance was determined by unpaired Student's t-test. All immunoblots shown were generated from the same membrane. **(B)** CDF plot of relative mRNA abundance as determined from RNA-seq following CLIP-$UPF1_{LL}$ vs CLIP-$UPF1_{SL}$ overexpression. mRNAs were binned according to 3'UTR lengths (short, medium, long). Statistical significance was determined by K-W test, with Dunn's correction for multiple comparisons. **(C)** Box plot of relative mRNA abundance as determined from RNA-seq following CLIP-$UPF1_{LL}$ vs CLIP-$UPF1_{SL}$ overexpression. mRNAs with long 3'UTRs (> 1579 nt) were binned by PTBP1 and/or hnRNP L motif density within the 3'UTR. Statistical significance was determined by K-W test, with Dunn's correction for multiple comparisons (**** $P < 1x10^{-12}$). Boxes indicate interquartile ranges, and bars indicate Tukey whiskers. **(D)** qPCR analysis of indicated transcripts from CLIP-UPF1 overexpression RNA-seq experiments. Relative fold changes are in reference to the GFP expressing control line. Significance of CLIP-$UPF1_{LL}$ overexpression was compared to CLIP-$UPF1_{SL}$. Asterisk (*) indicates $P < 0.05$, as determined by unpaired Student's t-test. Error bars indicate SD (n = 3).

did not cause systematic changes in the abundance of putative EJC-stimulated NMD targets, although some mRNAs in this class were modestly perturbed (Figure 2-figure supplement 1B and 1C). These findings support that overexpression of CLIP-UPF1$_{LL}$ does not substantially impair cellular NMD functions but may cause some alteration of NMD target specificity.

Having determined that overexpression of UPF1$_{LL}$ had only minor effects on the abundance of established NMD targets, we next asked whether UPF1$_{LL}$ overexpression affected the abundance of mRNAs not normally regulated by NMD. Transcripts with long 3'UTRs represent a large population of potential NMD targets (Hurt et al., 2013; Yepiskoposyan et al., 2011), only some of which are degraded by the pathway under normal conditions (Toma et al., 2015). Indeed, subdivision of the transcriptome according to 3'UTR lengths (first tertile: < 477 nt; second tertile: 477-1578 nt; third tertile: > 1579 nt) revealed differential effects of the UPF1 isoforms on mRNAs containing long 3'UTRs. Specifically, mRNAs with the longest 3'UTRs were more likely to be reduced in abundance by UPF1$_{LL}$ overexpression than mRNAs with shorter 3'UTRs (Figure 2B). These results demonstrate that UPF1$_{LL}$ has an expanded reach that includes a population of long 3'UTR-containing transcripts that are not normally affected by the pathway.

UPF1$_{LL}$ downregulates mRNAs normally protected from NMD

We have identified hundreds to thousands of mammalian long 3'UTRs that evade UPF1 recognition and decay by association with the protective RNA-binding proteins (RBPs) polypyrimidine tract binding protein 1 (PTBP1) and heterogeneous nuclear ribonucleoprotein L (hnRNP L) (Ge et al., 2016; Kishor et al., 2018, 2020). The observation that UPF1$_{LL}$ has an expanded reach to target long 3'UTRs not normally affected by NMD led us to ask whether mRNAs protected by PTBP1 and/or hnRNP L undergo enhanced downregulation with UPF1$_{LL}$ overexpression. To explore this possibility, we further subdivided the population of long 3'UTR-containing mRNAs according to the density of PTBP1 and hnRNP L binding sites in their

3'UTRs. We previously showed that transcripts containing a low density of PTBP1 and/or hnRNP L binding sites were more likely to be bound and degraded by UPF1 than transcripts containing a high density of binding sites for the protective proteins (Kishor et al., 2018). Remarkably, mRNAs with a moderate or high density of 3'UTR PTBP1 and/or hnRNP L binding sites were more likely to be down-regulated by $UPF1_{LL}$ overexpression than those with a low density of binding sites (Figure 2C). Quantitative PCR of select transcripts confirmed these transcriptome-wide results of RNA-seq (Figure 2D), including CSRP1, which we have previously established as a long 3'UTR-containing mRNA that undergoes decay upon hnRNP L knockdown or mutation of hnRNP L binding sites in its 3'UTR (Kishor et al., 2018).

To investigate whether the observed changes in mRNA abundance with $UPF1_{LL}$ overexpression were due to decreased stability, we used REMBRANDTS software to infer changes in mRNA stability based upon differences in the relative abundance of exonic and intronic reads from each gene (Alkallas et al., 2017). The effects of $UPF1_{SL}$ overexpression on mRNA abundance were not correlated with changes to mRNA stability (Spearman's ρ = -0.45; $P = 10^{-5}$; Figure 2-figure supplement 1D), corroborating the previous observation that moderate $UPF1_{SL}$ overexpression does not enhance NMD activity (Huang et al., 2011). In contrast, the effects of $UPF1_{LL}$ overexpression on mRNA abundance and changes in stability were highly correlated (Spearman's ρ = 0.65; $P < 10^{-15}$; Figure 2-figure supplement 1E), with transcripts having long 3'UTRs and a high density of PTBP1 and/or hnRNP L binding motifs showing preferential destabilization with $UPF1_{LL}$ overexpression (Figure 2-figure supplement 1F and 1G). These data support the conclusion that the $UPF1_{LL}$ isoform is able to overcome the protective proteins to promote decay of mRNAs normally shielded from NMD.

Figure 3

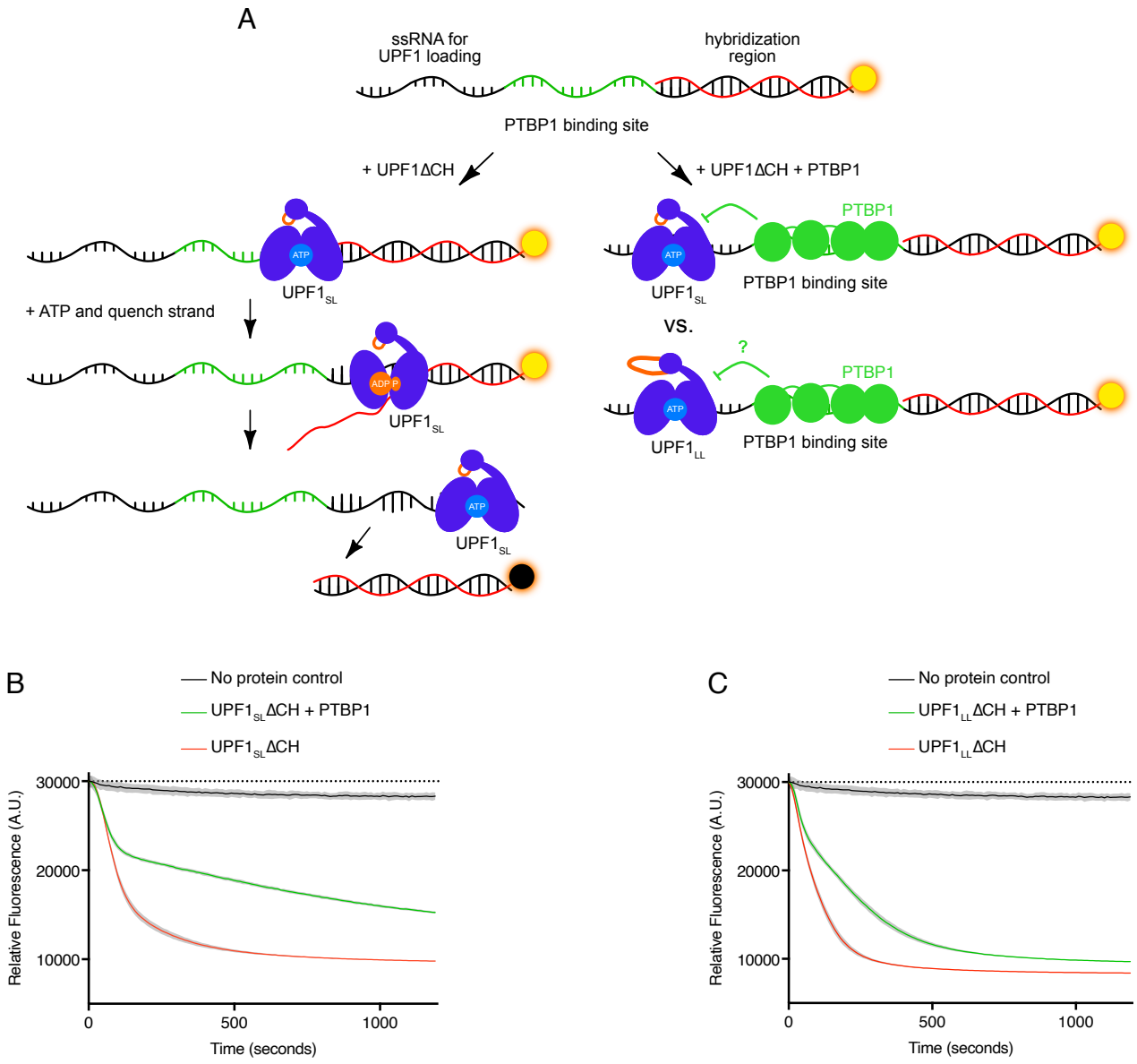

Figure 3: $UPF1_{LL}$ overcomes the translocation inhibition by PTBP1. (A) Scheme of the fluorescence-based unwinding assay to monitor UPF1 translocation in real time (Fritz et al., 2020). **(B)** $UPF1_{SL}$ΔCH translocation along an RNA substrate harboring a high-affinity PTBP1 binding site in the absence and presence of PTBP1. **(C)** Results as in (B) but with $UPF1_{LL}$ΔCH. Results of four technical replicates are shown for each dataset and represent at least three independent experiments. Shaded region indicates SD.

$UPF1_{LL}$ is less sensitive to PTBP1-mediated inhibition of translocation

We have proposed that the protective RBPs PTBP1 and hnRNP L promote UPF1 dissociation from potential NMD substrates prior to decay induction (Fritz et al., 2020). In support of this model, deletion of the regulatory loop from the helicase core of recombinant $UPF1_{SL}$ rendered UPF1 less sensitive to PTBP1 inhibition *in vitro* (Fritz et al., 2020). Because UPF1 variants with a truncated regulatory loop and the physiological $UPF1_{LL}$ isoform containing the extended 22 amino acid loop both exhibit a greater affinity for RNA in the presence of ATP than the PTBP1-sensitive $UPF1_{SL}$ isoform (Gowravaram et al., 2018), we hypothesized that $UPF1_{LL}$ could overcome the translocation inhibition by PTBP1.

We recently established a real-time assay to monitor UPF1 helicase activity (Fritz et al., 2020). In this assay, UPF1 translocation and duplex unwinding causes a fluorescently labeled oligonucleotide to be displaced from the assay substrate (Figure 3A, left). An excess of complementary oligonucleotide labeled with a dark quencher is provided, causing a decrease in fluorescence with increased UPF1 unwinding. Inhibition of UPF1 translocation results in sustained fluorescence over time, allowing for the determination of inhibitory effects on UPF1 unwinding activity. We leveraged this system to compare $UPF1_{LL}$ versus $UPF1_{SL}$ translocation on a duplexed RNA substrate harboring a high affinity PTBP1 binding site (Figure 3A, right) (Fritz et al., 2020). For these experiments, we compared the activity of UPF1 proteins containing the helicase core but lacking the autoinhibitory N-terminal cysteine-histidine domain (UPF1ΔCH) (Figure 3-figure supplement 1A and 1B) (Chakrabarti et al., 2011; Fiorini et al., 2012; Fritz et al., 2020). The ability of $UPF1_{SL}$ΔCH to unwind the duplexed RNA was impaired by the presence of PTBP1 (Figure 3B), as previously observed (Fritz et al., 2020). In contrast, inhibition of $UPF1_{LL}$ΔCH translocation by PTBP1 was reduced by extension of the regulatory loop (Figure 3C). These results indicate that $UPF1_{LL}$ can overcome the translocation inhibition by PTBP1, reinforcing the conclusion that PTBP1-mediated UPF1 dissociation depends on the steric clash between the UPF1 regulatory loop and RNA.

$UPF1_{LL}$ is able to bind NMD-resistant transcripts

The resistance of $UPF1_{LL}$ to inhibition by PTBP1 *in vitro* suggested that it may have altered RNA-binding specificity in cells. We therefore used the CLIP-UPF1 stable lines to ask whether the distribution of $UPF1_{LL}$ across the transcriptome differs from that of $UPF1_{SL}$. CLIP-UPF1 complexes were isolated from whole cell extracts by streptavidin affinity purification using a CLIP-biotin substrate (Figure 4A and Figure 4-figure supplement 1A), with GFP-expressing cell lines as a negative control for interaction specificity (Kishor et al., 2020). Bound mRNAs were then extracted and used for sequencing library preparation. Because recovery of RNA from GFP samples was at least 100-fold lower than from CLIP-UPF1 affinity purifications, only UPF1 samples were analyzed by RNA-seq. UPF1 occupancy was assessed by normalizing the abundance of transcripts in RIP-seq samples to their abundance in total RNA-seq (UPF1 RIP-seq efficiency; Supplementary file 2). The two UPF1 isoforms recovered thousands of mRNAs with similar efficiency, and both strongly enriched mRNAs up-regulated by total UPF1 depletion (Figure 4-figure supplement 1B), consistent with previous observations of preferential UPF1 association with NMD targets (Hurt et al., 2013; Kishor et al., 2018; Zünd et al., 2013). The efficiency of mRNA co-purification with both $UPF1_{SL}$ and $UPF1_{LL}$ increased with 3'UTR length (Figure 4B), reinforcing previous observations that UPF1 binding correlates with 3'UTR length (Baker and Hogg, 2017; Hogg and Goff, 2010; Hurt et al., 2013; Kurosaki et al., 2014). Distinctly, the preferential downregulation of long 3'UTR-containing mRNAs by $UPF1_{LL}$ was accompanied by a greater propensity to associate with the longest class of 3'UTRs than $UPF1_{SL}$ (Figure 4B). Indeed, mRNAs down-regulated upon $UPF1_{LL}$ overexpression were highly enriched in $UPF1_{LL}$ relative to $UPF1_{SL}$ RIP-seq (Figure 4C; compare solid and dotted blue lines), and we therefore designated mRNAs with $UPF1_{LL}$:$UPF1_{SL}$ RIP-seq efficiency greater than 1.5 as putative direct $UPF1_{LL}$ substrates for further analysis.

Based on the resistance of $UPF1_{LL}$ to PTBP1 in the *in vitro* unwinding assays (Figure 3), we hypothesized that $UPF1_{LL}$ would additionally occupy mRNAs with a high density of protective

Figure 4

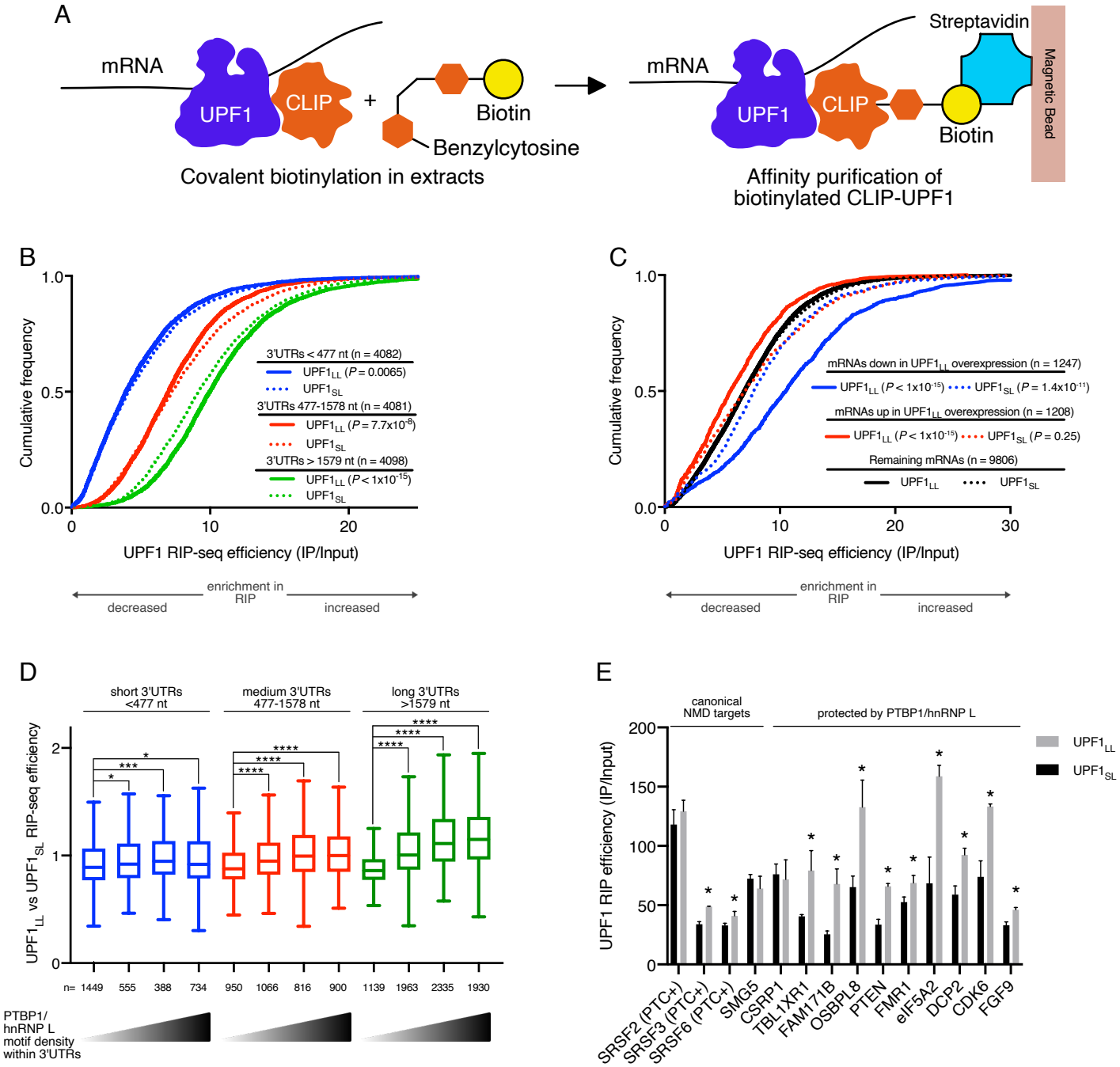

Figure 4: UPF1$_{LL}$ binds NMD-protected transcripts. (A) Scheme for the CLIP-UPF1 affinity purification (RIP) assay. **(B)** CDF plot of recovered mRNAs as determined from RIP-seq efficiency in CLIP-UPF1$_{LL}$ or CLIP-UPF1$_{SL}$ affinity purifications. mRNAs were binned according to 3'UTR length. Significance of mRNA co-purification with CLIP-UPF1$_{LL}$ was compared to that of CLIP-UPF1$_{SL}$, and was determined by K-S test. **(C)** CDF plot as in (B), with mRNAs binned by sensitivity (either up or down-regulated) to UPF1$_{LL}$ overexpression. Significance of indicated mRNA co-purification with CLIP-UPF1$_{LL}$ or CLIP-UPF1$_{SL}$ is relative to those unaffected by UPF1$_{LL}$ overexpression (remaining mRNAs), and was determined by K-W test with Dunn's correction for multiple comparisons. **(D)** Box plot of recovered mRNAs as determined from RIP-seq efficiency in CLIP-UPF1$_{LL}$ vs CLIP-UPF1$_{SL}$ affinity purifications. mRNAs were binned according to 3'UTR length (short, medium, long) and then further subdivided by PTBP1 and/or hnRNP L motif density within the 3'UTR. Statistical significance was determined by K-W test, with Dunn's correction for multiple comparisons (* $P < 0.03$, *** $P < 0.0002$, **** $P < 2\text{x}10^{-11}$). Boxes indicate interquartile ranges, and bars indicate Tukey whiskers. **(E)** qPCR analysis of indicated transcripts from UPF1 RIP-seq experiments. Relative fold enrichment was determined by dividing the recovered mRNA by its corresponding input amount. Significance of the CLIP-UPF1$_{LL}$ RIP was compared to that of CLIP-UPF1$_{SL}$. Asterisk (*) indicates $P < 0.05$, as determined by unpaired Student's t-test. Error bars indicate SD (n = 3).

protein binding sites. Consistent with this idea, $UPF1_{LL}$ and $UPF1_{SL}$ exhibited almost identical recovery of mRNAs with long 3'UTRs containing a low density of PTBP1 and hnRNP L motifs, while mRNAs with moderate or high protective protein motif density were preferentially recovered by $UPF1_{LL}$ (Figure 4D). Quantitative PCR of select transcripts confirmed these transcriptome-wide results of RIP-seq (Figure 4E). These data support that $UPF1_{LL}$ can circumvent the protective effects of PTBP1 and/or hnRNP L by remaining bound to otherwise protected mRNAs.

Coordinated downregulation of $UPF1_{LL}$ targets during ER stress and ISR induction

We next investigated whether the population of mRNAs bound and regulated by $UPF1_{LL}$ tended to encode proteins with particular properties or functions that might indicate physiological conditions in which $UPF1_{LL}$ influences gene expression. Based on previous observations that genes in functionally related pathways are often coordinately regulated at the posttranscriptional level (Keene, 2007), we performed a gene ontology (GO) enrichment analysis (Eden et al., 2009) to identify commonalities among mRNAs preferentially bound by $UPF1_{LL}$ in the RIP-seq experiments. This analysis revealed a high degree of enrichment for proteins that rely on the endoplasmic reticulum (ER) for biogenesis, including a large number of integral membrane proteins (204 of 686 $UPF1_{LL}$-enriched genes assigned to GO terms; Figure 5A; Supplementary file 3). We used a previous survey of ER-localized translation (Jan et al., 2014) to corroborate the results of the GO analysis, finding that many $UPF1_{LL}$ targets were indeed found to be preferentially translated at the ER (Figure 5-figure supplement 1A).

The ER is a major site of protein synthesis and folding, and thus serves as a nexus for signaling to prevent proteotoxic stress. In response to the accumulation of unfolded proteins, cells activate the integrated stress response (ISR), which restores homeostasis by repressing translation and inducing expression of a battery of stress response genes (Costa-Mattioli and Walter, 2020). A major effect of ISR-mediated translational repression is corresponding

Figure 5

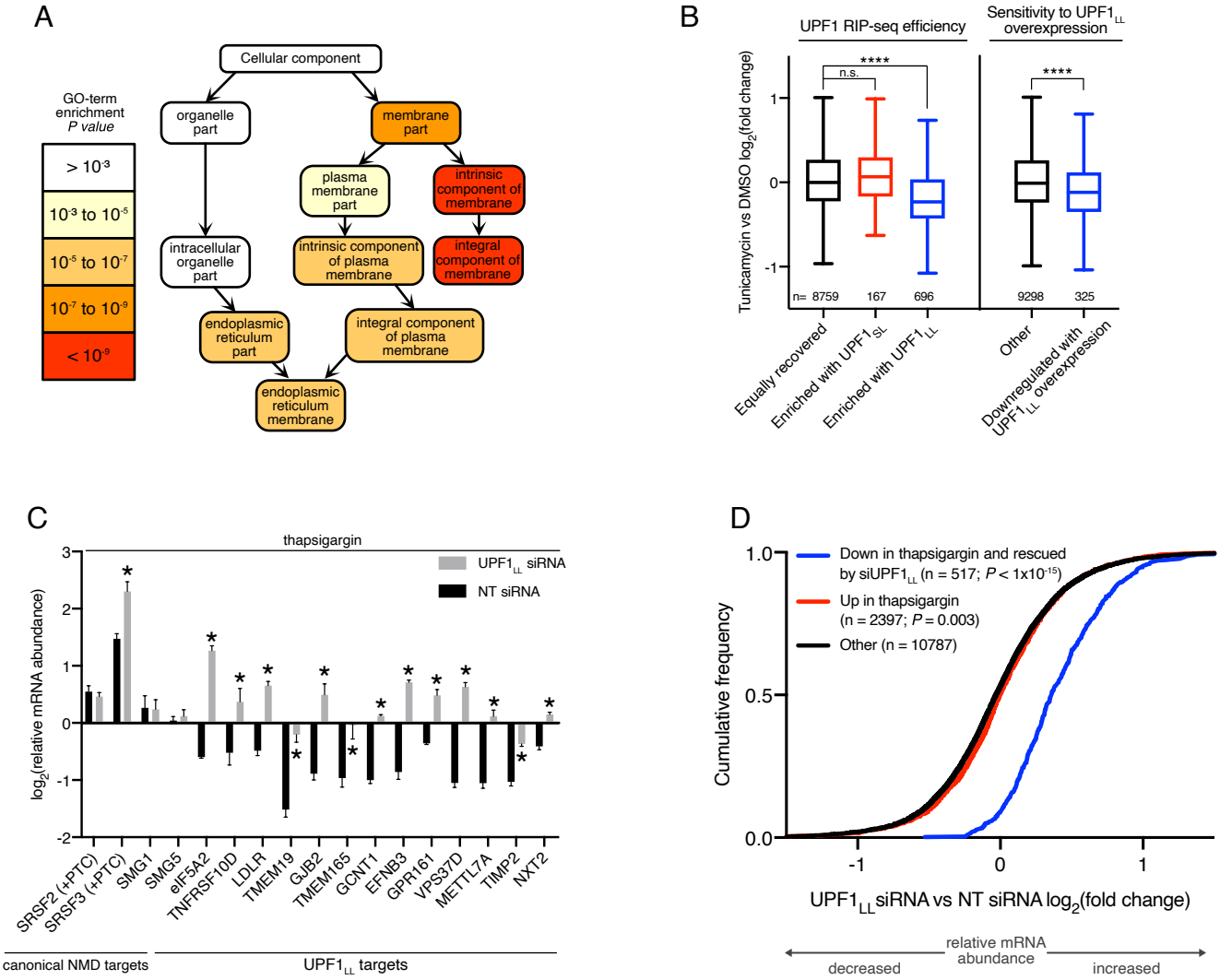

Figure 5: UPF1$_{LL}$ downregulates a population of mRNAs during ER stress and induction of the ISR. **(A)** Gene ontology analysis of UPF1$_{LL}$ target mRNAs. **(B)** Box plot of relative mRNA abundance as determined from RNA-seq following treatment of HEK-293 cells with 1 μM tunicamycin for 6hr (GSE97384). mRNAs were binned by RIP-seq efficiency in CLIP-UPF1$_{LL}$ or CLIP-UPF1$_{SL}$ affinity purifications (left) or sensitivity to CLIP-UPF1$_{LL}$ overexpression (right). Statistical significance was determined by K-W test, with Dunn's correction for multiple comparisons (left; **** $P < 1x10^{-15}$), or K-S test (right; **** $P < 1x10^{-6}$). Boxes indicate interquartile ranges, and bars indicate Tukey whiskers. **(C)** qPCR analysis of indicated transcripts following treatment of HEK-293 cells with 1 μM thapsigargin for 6hr. Relative fold changes are in reference to vehicle-treated, NT siRNA. Significance of UPF1$_{LL}$ knockdown was compared to the NT siRNA control in thapsigargin treatment. Asterisk (*) indicates $P < 0.05$, as determined by unpaired Student's t-test. Error bars indicate SD (n = 3). **(D)** CDF plot of relative mRNA abundance as determined from RNA-seq following UPF1$_{LL}$ knockdown in HEK-293 cells. mRNAs were binned by sensitivity to thapsigargin and rescue by UPF1$_{LL}$ knockdown. Statistical significance was determined by K-W test, with Dunn's correction for multiple comparisons.

stabilization of canonical NMD targets, including several mRNAs encoding factors integral to the activation and resolution of the stress response (Goetz and Wilkinson, 2017). Because of our finding that $UPF1_{LL}$ preferentially binds and downregulates mRNAs enriched for membrane and ER-associated gene products, we asked whether $UPF1_{LL}$ activity is regulated in response to ER stress and induction of the ISR. In sharp contrast to previously characterized NMD targets, analysis of publicly-available RNA-seq data of HEK-293 cells treated with the ER stress-inducing agent tunicamycin (Park et al., 2017) revealed that mRNAs preferentially bound and down-regulated by $UPF1_{LL}$ were systematically less abundant during ER stress (Figure 5B). In comparison, transcripts preferentially bound by $UPF1_{SL}$, and more likely to represent canonical NMD targets, were up-regulated. This effect was not limited to a specific cell type, as $UPF1_{LL}$ targets were similarly down-regulated in HeLa cells treated with tunicamycin (Figure 5-figure supplement 1B) (Boswell et al., 2017). Mouse NIH3T3 cells treated with the ER stress-inducing agent thapsigargin (Namkoong et al., 2018) also exhibited a coordinated decrease in the abundance of $UPF1_{LL}$ targets (Figure 5-figure supplement 1C), indicating a conserved response across mammals in response to distinct ISR-inducing conditions.

We next asked whether the downregulation of specific mRNAs during ISR induction occurs via $UPF1_{LL}$. To this end, we transfected HEK-293 cells with non-targeting (NT) or $UPF1_{LL}$-specific siRNAs and then treated cells with thapsigargin (Figure 5-figure supplement 1D). Changes in gene expression caused by $UPF1_{LL}$ depletion were not due to differential ISR induction, as previously established stress response genes were comparably up-regulated in response to thapsigargin treatment following NT and $UPF1_{LL}$-specific knockdown (Figure 5-figure supplement 1E) (Ashburner et al., 2000; The Gene Ontology Consortium, 2019). Strikingly, we identified a population of mRNAs (~500 transcripts) that decreased in abundance with thapsigargin treatment and were rescued upon $UPF1_{LL}$ knockdown (Supplementary file 4). Quantitative PCR of select transcripts confirmed these transcriptome-wide results of RNA-seq (Figure 5C). Many transcripts down-regulated in thapsigargin and rescued by $UPF1_{LL}$

knockdown were also significantly up-regulated with basal depletion of the $UPF1_{LL}$ isoform (Figure 5D). In contrast, mRNAs that were up-regulated in thapsigargin were only modestly perturbed by $UPF1_{LL}$ knockdown under basal conditions. Thapsigargin treatment alone did not alter the relative levels of $UPF1_{LL}$ and $UPF1_{SL}$ mRNAs (Figure 5-figure supplement 1F), indicating that $UPF1_{LL}$ activity in ER stress was likely due to activity of the existing population of protein rather than a consequence of altered UPF1 splicing. Taken together, these data suggest that the ISR enhances $UPF1_{LL}$-dependent NMD, while suppressing decay of canonical $UPF1_{SL}$ target mRNAs.

$UPF1_{LL}$ activity is enhanced by translational repression

We have identified that endogenous $UPF1_{LL}$ regulates the abundance of a population of transcripts that undergo enhanced downregulation during ER stress and induction of the ISR. We next sought to determine the underlying mechanism controlling $UPF1_{LL}$ activity under cellular stress conditions. Activation of the ISR induces phosphorylation of eIF2α (Figure 5-figure supplement 1G), driving global downregulation of protein synthesis due to impaired eIF2-GTP-Met-$tRNA_i$ ternary complex recycling and reduced delivery of the initiator Met-$tRNA_i$ to translating ribosomes (Baird and Wek, 2012; Wek, 2018; Young and Wek, 2016). We therefore asked whether translational repression would induce $UPF1_{LL}$ activity outside of the context of the ISR. To explore this possibility, we first assessed the response of $UPF1_{LL}$ target transcripts in previously published RNA-seq data from HEK-293 cells treated with the translation elongation inhibitor emetine (Martinez-Nunez et al., 2017). We observed that mRNAs preferentially bound and down-regulated by CLIP-$UPF1_{LL}$ were overwhelmingly decreased in abundance with emetine treatment (Figure 6A and Figure 6-figure supplement 1A). Notably, a similar effect was evident upon impairment of translation initiation, as we observed decreased expression of direct CLIP-$UPF1_{LL}$ target mRNAs with decreased translation initiation upon hippuristanol treatment (Figure 6B and Figure 6-figure supplement 1B) in a publicly available RNA-seq dataset

Figure 6

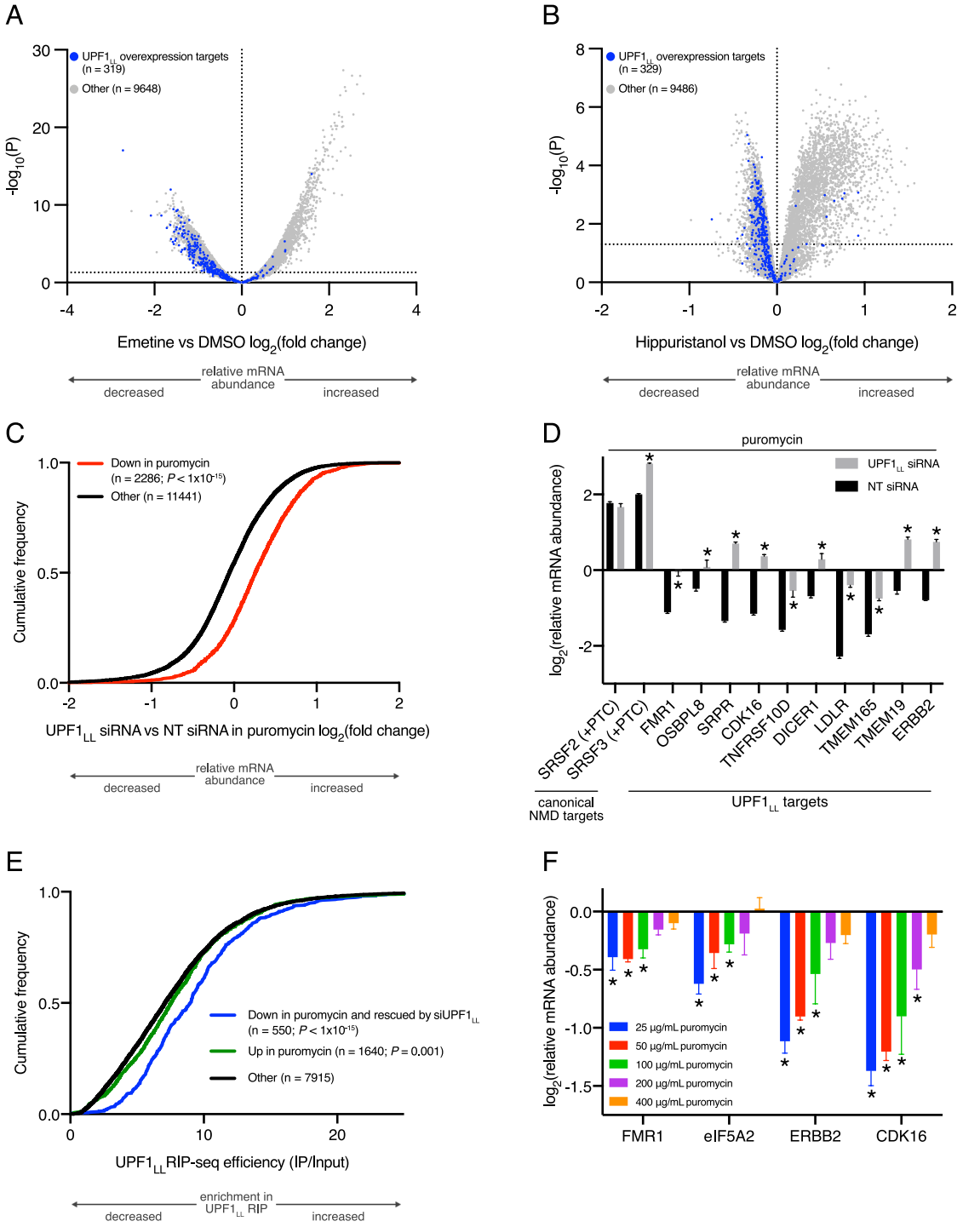
A
UPF1LL overexpression targets (n = 319)
Other (n = 9648)
-log10(P)
Emetine vs DMSO log2(fold change)
relative mRNA abundance
decreased
increased
B
UPF1LL overexpression targets (n = 329)
Other (n = 9486)
-log10(P)
Hippuristanol vs DMSO log2(fold change)
relative mRNA abundance
decreased
increased
C
Down in puromycin (n = 2286; P < 1x10-15)
Other (n = 11441)
Cumulative frequency
UPF1LL siRNA vs NT siRNA in puromycin log2(fold change)
relative mRNA abundance
decreased
increased
D
puromycin
UPF1LL siRNA
NT siRNA
log2(relative mRNA abundance)
SRSF2 (+PTC)
SRSF3 (+PTC)
FMR1
OSBPL8
SRPR
CDK16
TNFRSF10D
DICER1
LDLR
TMEM165
TMEM19
ERBB2
canonical NMD targets
UPF1LL targets
E
Down in puromycin and rescued by siUPF1LL (n = 550; P < 1x10-15)
Up in puromycin (n = 1640; P = 0.001)
Other (n = 7915)
Cumulative frequency
UPF1LL RIP-seq efficiency (IP/Input)
enrichment in UPF1LL RIP
decreased
increased
F
25 µg/mL puromycin
50 µg/mL puromycin
100 µg/mL puromycin
200 µg/mL puromycin
400 µg/mL puromycin
log2(relative mRNA abundance)
FMR1
eIF5A2
ERBB2
CDK16

Figure 6: $UPF1_{LL}$ activity is enhanced by translational repression. (A) Volcano plot of relative mRNA abundance as determined from RNA-seq following treatment of HEK-293 cells with 50 ng/mL emetine for 4hr (GSE89774). mRNAs were binned by sensitivity to CLIP-$UPF1_{LL}$ overexpression. Horizontal line indicates the significance threshold $P \leq 0.05$. **(B)** Volcano plot as in (A), following treatment of MCF7 cells with 150 nM hippuristanol for 1hr (GSE134888). **(C)** CDF plot of relative mRNA abundance as determined from RNA-seq following $UPF1_{LL}$-specific knockdown in HEK-293 cells and treatment with 50 μg/mL puromycin for 4hr. mRNAs were binned by sensitivity to puromycin in the absence of $UPF1_{LL}$ knockdown. Statistical significance was determined by K-S test. **(D)** qPCR analysis of indicated transcripts following treatment of HEK-293 cells with 50 μg/mL puromycin for 4hr. Relative fold changes are in reference to vehicle-treated, NT siRNA. Significance of $UPF1_{LL}$ knockdown was compared to the NT siRNA control in puromycin treatment. Asterisk (*) indicates $P < 0.05$, as determined by unpaired Student's t-test. Error bars indicate SD (n = 3). **(E)** CDF plot of $UPF1_{LL}$ RIP-seq efficiency. mRNAs were binned by sensitivity to puromycin and rescue by $UPF1_{LL}$ knockdown. Statistical significance was determined by K-W test, with Dunn's correction for multiple comparisons. **(F)** qPCR analysis of indicated transcripts following treatment of HEK-293 cells with indicated concentrations of puromycin for 4hr. Relative fold changes are in reference to vehicle-treated control. Significance of puromycin treatment on relative transcript abundance was compared to the vehicle-treated control. Asterisk (*) indicates $P < 0.05$, as determined by two-way ANOVA. Error bars indicate SD (n = 3).

(Waldron et al., 2019). Extending our analysis to independent datasets of emetine (Zhang et al., 2019) and hippuristanol (Karginov and Hannon, 2013), as well as RNA-seq data generated with the elongation inhibitor anisomycin (Hollerer et al., 2016), confirmed that the downregulation of $UPF1_{LL}$ target transcripts was not specific to a cell type or mechanism of translation inhibition (Figure 6-figure supplement 1C, 1D, and 1E).

Having established a relationship between translational repression and enhanced $UPF1_{LL}$ activity, we next sought to directly test the role of $UPF1_{LL}$ in modulating mRNA abundance in response to impaired translation. To this end, we transfected HEK-293 cells with non-targeting (NT) or $UPF1_{LL}$-specific siRNAs and then treated cells with the elongation inhibitor puromycin (50 µg/mL; Figure 6-figure supplement 1F). In RNA-seq analyses, we identified a population of mRNAs (~700 transcripts) that decreased in abundance by at least 1.4-fold with puromycin treatment and were rescued upon $UPF1_{LL}$ knockdown (Figure 6C and Supplementary file 4). Quantitative PCR of select transcripts confirmed these transcriptome-wide results (Figure 6D). Inferred mRNA stability changes using REMBRANDTS software indicated that the observed changes in mRNA abundance upon puromycin treatment and $UPF1_{LL}$ knockdown were due to changes in mRNA stability (Figure 6-figure supplement 1G). In support of this conclusion, $UPF1_{LL}$ target transcripts decayed with apparent first-order kinetics upon puromycin treatment, with significant loss of the most highly affected transcripts in as little as 15 minutes (Figure 6-figure supplement 1H). Similar to conditions of ER stress, puromycin treatment alone did not alter UPF1 splicing (Figure 6-figure supplement 1I), indicating that enhanced $UPF1_{LL}$ activity during conditions of impaired translation efficiency was likely due to the existing population of $UPF1_{LL}$ protein.

We next evaluated whether our $UPF1_{LL}$ overexpression RNA-seq and RIP-seq studies were predictive of the response of $UPF1_{LL}$ target transcripts to impaired translation efficiency. Indeed, mRNAs that were selectively down-regulated with $UPF1_{LL}$ overexpression showed a significant decrease in abundance with puromycin treatment, an effect that was likely due to

enhanced decay according to REMBRANDTS analysis (Figure 6-figure supplement 1J). Correspondingly, mRNAs that decreased in abundance with puromycin treatment in a $UPF1_{LL}$-dependent manner were enriched in the $UPF1_{LL}$ RIP-seq (Figure 6E). These results indicate that $UPF1_{LL}$ substrates are more efficiently targeted during conditions of impaired translation due to both translation repression and the intrinsic capacity of $UPF1_{LL}$ to stably associate with these transcripts.

Because of the well-established requirement for translation in NMD, we hypothesized that the $UPF1_{LL}$-dependent decay we observed under conditions of translational repression was due to rare translation termination events. To test this hypothesis, we treated cells with a titration of puromycin from 25 µg/mL to 400 µg/mL. If $UPF1_{LL}$ is capable of responding to rare translation termination events, its activity should be promoted at low concentrations of puromycin that permit a low level of termination events to persist, but inhibited by high concentrations of puromycin that more efficiently block translation. In line with these expectations, we observed a dose-dependent response, in which downregulation of representative $UPF1_{LL}$ target transcripts FMR1, eIF5A2, ERBB2, and CDK16 was most efficient at lower puromycin concentrations (Figure 6F). Treatment with high concentrations of puromycin did not have a significant effect on the levels of these $UPF1_{LL}$ target mRNAs, consistent with a requirement for translation termination events. Corroborating these findings, we observed globally more efficient downregulation of puromycin-sensitive $UPF1_{LL}$ targets with 25 µg/mL puromycin than 100 µg/mL puromycin in RNA-seq (Figure 6-figure supplement 1K). Based on these results, we conclude that translation termination is strictly required for all NMD events but that changes in translation efficiency can drive downregulation of a novel class of substrates by the $UPF1_{LL}$ isoform.

Discussion

The idea that the mammalian NMD pathway consists of multiple branches with distinct factor requirements and substrate specificities has been proposed by several groups, but the underlying mechanisms and regulatory roles of NMD branching are poorly understood (Chan et al., 2007; Huang et al., 2011; Yi et al., 2020). Our identification of differential activities of $UPF1_{SL}$ and $UPF1_{LL}$ is an unforeseen example of NMD pathway branching, which can be controlled at the cellular level by changes in the abundance of protective RBPs, translation, or UPF1 splicing. When cells have high translational capacity, both isoforms contribute to NMD, but the more abundant $UPF1_{SL}$ isoform predominates. Contrary to expectations, we find that the NMD pathway does not cease to function under conditions of acute translational repression. Instead, the pathway switches from $UPF1_{SL}$-dependent targeting of frequent translation termination events to $UPF1_{LL}$-dependent targeting of rare termination events. In this way, $UPF1_{LL}$ enables cellular NMD specificity to be dramatically remodeled in response to changing cellular conditions, including induction of the ISR. Because altered translation efficiency and induction of cellular stress pathways are pervasive in cancer, genetic disease, and infection, the expanded scope of $UPF1_{LL}$-dependent decay has far-reaching implications for the physiological roles of the mammalian NMD pathway in health and disease.

We propose a new model by which the alternative $UPF1_{LL}$ isoform expands the scope of NMD under conditions of translation inhibition (Figure 7). The differential RNA-binding properties of $UPF1_{SL}$ and $UPF1_{LL}$ in the presence of ATP and protective proteins provide a molecular mechanism by which NMD specificity is tuned in response to cellular conditions (Fritz et al., 2020; Gowravaram et al., 2018). We show that $UPF1_{LL}$ activity can be promoted either by overexpression of the $UPF1_{LL}$ protein or by translational repression. When $UPF1_{LL}$ expression is enhanced, it is able to overcome inhibition by the protective RBPs PTBP1 and hnRNP L, which normally induce dissociation of $UPF1_{SL}$ from potential NMD target mRNAs. Under conditions of translational repression, canonical NMD targets predominantly regulated by $UPF1_{SL}$ are

Figure 7

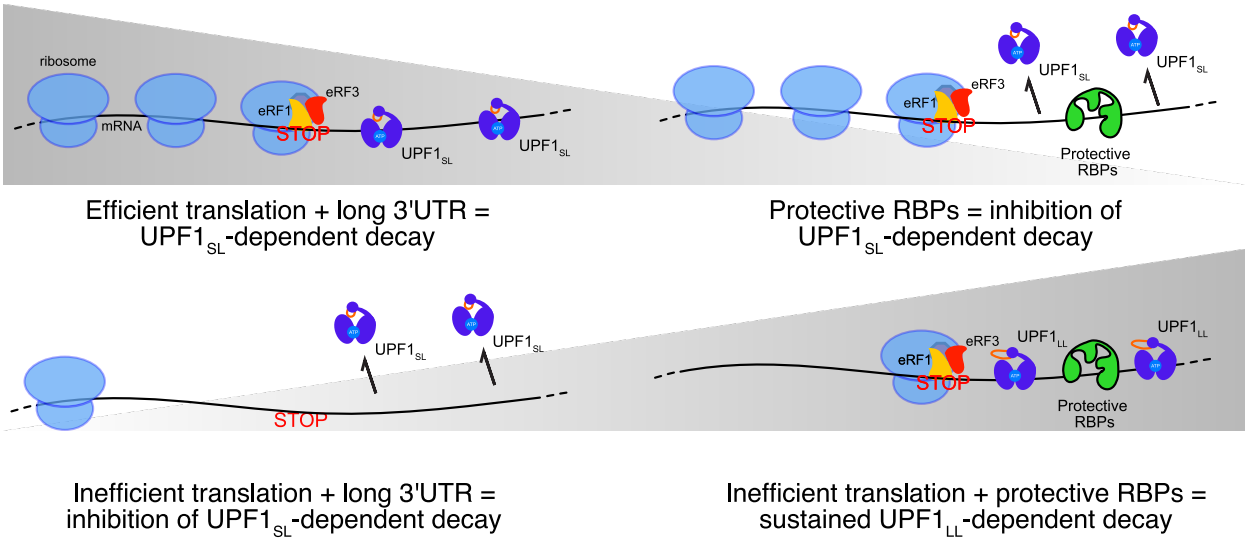

Figure 7: Translation efficiency and UPF1 isoform expression can conditionally alter target susceptibility by the NMD pathway. See Discussion for details.

stabilized, while $UPF1_{LL}$ targets are more efficiently degraded. Together, our data inform a model in which the increased residence time of $UPF1_{LL}$ on RNAs enables it to respond to rare translation termination events, while the faster dissociation of $UPF1_{SL}$ from potential NMD targets renders it unable to complete decay when translation is repressed. The rate-limiting step(s) downstream of UPF1 binding to potential NMD targets remain to be clearly defined, but an attractive possibility is that persistent binding of $UPF1_{LL}$ to mRNAs promotes its phosphorylation by SMG1, readying $UPF1_{LL}$ to scaffold assembly of productive decay complexes (Durand et al., 2016).

The majority of studies of mammalian UPF1 functions have been conducted in transformed, rapidly dividing, highly metabolically active cells under conditions that promote efficient translation. These experimental conditions have enabled the discovery that UPF1 and NMD as a whole have a major impact on transcriptome expression, affecting 5-10% of all human genes. However, our work strongly suggests that these studies have been blind to major physiological functions of UPF1. Notably, we find that $UPF1_{LL}$ is able to regulate several proteins of central importance in cancer and other diseases, including fragile X mental retardation 1 (FMR1), the low-density lipoprotein receptor (LDLR), and oncogenes PTEN, EIF5A2, and ERBB2, among others (Dockendorff and Labrador, 2019; Goldstein and Brown, 2009; Harbeck et al., 2019; Mathews and Hershey, 2015; Song et al., 2012). Along with EIF5A2 and FMR1, several additional translation-dependent $UPF1_{LL}$ targets are themselves important regulators of translation and mRNA decay, including the signal recognition particle receptor SRPR, the major mRNA decapping enzyme DCP2, and the miRNA-processing endonuclease DICER1 (Akopian et al., 2013; Michlewski and Cáceres, 2019; Mugridge et al., 2018).

The ability of $UPF1_{LL}$ to target specific mRNAs in response to translational repression positions NMD to function as a mechanism to reset the transcriptome upon cellular stress. Translational repression via eIF2α phosphorylation is a central feature of responses to diverse cellular insults, including nutrient deprivation, oncogene activation, viral infection, hypoxia, and

accumulation of unfolded proteins (Pakos-Zebrucka et al., 2016). We show that induction of ER stress with thapsigargin induces $UPF1_{LL}$-dependent downregulation of hundreds of transcripts, a population particularly enriched for mRNAs encoding proteins trafficked through the ER. The extent to which $UPF1_{LL}$ complements or collaborates with IRE1-mediated mRNA decay or the recently identified ER-associated NMD factor NBAS will require further study (Hollien and Weissman, 2006; Longman et al., 2020), but our data suggest that $UPF1_{LL}$ may help to relieve proteotoxic stress by reducing the abundance of transcripts encoding proteins with complex biosynthetic needs.

Finally, we note that, even in the absence of stress, most cells and tissues *in vivo* likely have lower basal rates of translation than those attained in exponentially growing transformed cell lines. Based on our findings, understanding how NMD shapes gene expression in diverse tissue types in health and disease will require not just analysis of NMD targets characterized in transformed cells, but also transcripts that are conditionally targeted in response to changing translational states.

Materials and methods

UPF1 isoform nomenclature

$UPF1_{SL}$ refers to the protein encoded by Ensembl transcript ID ENST00000262803.9 and $UPF1_{LL}$ by Ensembl transcript ID ENST00000599848.5. These isoforms are also referred to as UPF1 isoform 2 and UPF1 isoform 1, respectively (Gowravaram et al., 2018; Nicholson et al., 2014).

In vitro helicase assays

Unwinding assays were performed as previously described in detail (Fritz et al., 2020). Briefly 75 nM of the pre-assembled RNA duplex substrate (described below) was combined with 1x unwinding reaction buffer (10 mM MES pH 6.0, 50 mM KOAc, 0.1 mM EDTA), 2 mM MgOAc, and 1 unit RNasin in a well of a half-area, black, flat-bottom 96-well plate (Corning 3993). PTBP1 (80 nM) was then added, the sample mixed, and incubated at room temperature for 10 minutes. Then, UPF1 (80 nM) was added, the sample mixed, and incubated at room temperature for 10 minutes. Finally, BHQ1 quencher (0.56 μM final; GTGTGCGTACAACTAGCT / 3BHQ_1) and 2 mM ATP was added to initiate the unwinding reaction. An Infinite^{R} F200 Pro microplate reader and associated i-control^{TM} 1.9 software (Tecan) was used to monitor Alexa Fluor 488 fluorescence every 10 seconds for 20 minutes at 37°C. Measured fluorescence intensities were normalized to the zero timepoint for each condition to obtain relative fluorescence values. Four technical replicates for each condition were obtained at the same time.

To generate the RNA duplex, a DNA oligo template (5' TAATACGACTCACTATAGGGACACAAAACAAAAGACAAAAACACAAAACAAAAGACAAAAAC ACAAAACAAAAGACAAAAAGCCTCTCCTTCTCTCTGCTTCTCTCTCGCTGTGTGCGTACAAC TAGCT 3') was PCR-amplified and then *in vitro* transcribed using the MEGAshortscript^{TM} T7 Transcription kit (Invitrogen). An 11:7 ratio of the helicase RNA substrate to 5' Alex Fluor 488

fluorescent oligo strand (Alexa Fluor 488/ AGCTAGTTGTACGCACAC) was incubated with 2 mM MgOAC and 1x unwinding reaction buffer at 95°C for 3 minutes and 30 seconds and then slowly cooled to 30°C. All DNA oligos were obtained from Integrated DNA Technologies. The 5' fluorescent oligo probe and 3' BHQ1 quencher were acquired as RNase-free, HPLC-purified.

Cloning, expression, and purification of recombinant $UPF1_{LL}\Delta CH$, $UPF1_{SL}\Delta CH$, and PTBP1 were conducted as previously described (Fritz et al., 2020; Gowravaram et al., 2018). In this study, $UPF1_{LL}\Delta CH$ was purified in the same manner as $UPF1_{SL}\Delta CH$ (Fritz et al., 2020).

Mammalian cell lines and generation of CLIP-UPF1 expression lines

HEK-293 cells used in the endogenous $UPF1_{LL}$ knockdown and RNA-seq experiments were received from ATCC (CRL-3216) and maintained at 37°C and 5% CO_2 in DMEM with 10% FBS (Gibco) and 1% pen/strep. Stable integration of CLIP-$UPF1_{SL}$ into human Flp-In™ T-REx™-293 cells (Invitrogen) and subsequent maintenance of this stable line was previously described (Kishor et al., 2020). CLIP-$UPF1_{LL}$ expression lines were generated and maintained in an identical manner.

Western for phosphorylation status of UPF1 variants

HEK-293 FLP-In CLIP-UPF1 stable cell lines were seeded at a density of $1x10^6$ cells in 10 cm plates and then treated with 200 ng/mL doxycycline hyclate (Sigma) for 48 hours to induce CLIP-UPF1 expression. On the day of harvest, cells were treated with or without 100 µg/mL puromycin (Sigma) for 2 hours. Whole cell lysate was then generated by the freeze/thaw method as previously described (Fritz et al., 2018; Hogg and Collins, 2007a, 2007b; Hogg and Goff, 2010). A total of 10 µg protein was run on a NuPage™ 4-12% Bis-Tris Protein Gel (Invitrogen) using MOPS buffer according to manufacturer instructions and subsequently transferred to a nitrocellulose membrane according to the NuPage™ manufacturer's protocol (Invitrogen). Membranes were incubated with a blocking buffer for fluorescent Western blotting

(Rockland) for 1 hour at room temperature and then incubated overnight at 4°C with the indicated primary antibody. Primary antibodies used were: anti-Phospho-(Ser/Thr) ATM/ATR (rabbit polyclonal, Cell Signaling, #2851, 1:1000), anti-RENT1/UPF1 (goat polyclonal, Bethyl, A300-038A, 1:1000), or anti-β-actin (mouse monoclonal, Cell Signaling, #3700, 1:1000). Membranes were subsequently washed three times with 1xTBS supplemented with 0.1% Tween-20 and then incubated with the appropriate secondary antibody for 1 hour at room temperature. Secondary antibodies used were: anti-rabbit IgG (H&L) Antibody DyLight™ 680 Conjugated (Rockland, 1:10,000), anti-goat IgG (H&L) Antibody DyLight™ 800 Conjugated (Rockland, 1:10,000), or anti-mouse IgG (H&L) Antibody DyLight™ 680 Conjugated (Rockland, 1:10,000). Membranes were washed three times with 1xTBS supplemented with 0.1% Tween-20 and then two times with 1xTBS. Western blot images were obtained on an Amersham Typhoon imaging system (GE Healthcare Life Sciences) and quantified using ImageStudio software (LI-COR Biosciences).

CLIP-UPF1 overexpression RIP-seq and RNA-seq

CLIP-UPF1$_{SL}$ overexpression RIP-seq and RNA-seq sample preparation were previously reported (Kishor et al., 2020); CLIP-UPF1$_{LL}$ datasets were generated in parallel. Briefly, CLIP-UPF1 stable cell lines or a GFP-expressing control line were seeded in 6 x 15 cm plates and then treated with 200 ng/mL doxycycline hyclate (Sigma) for 48 hours to induce CLIP-UPF1 expression. Cells were harvested 48 hours post-induction at 80-85% confluency and whole cell lysate generated by the freeze/thaw method as previously described (Fritz et al., 2018; Hogg and Collins, 2007a, 2007b; Hogg and Goff, 2010). Equilibrated cell extracts, reserving 1/10th for downstream input analysis, were then combined with 10 μM CLIP-Biotin (New England Biolabs) and rotated (end-over-end) for 1 hour at 4°C to allow CLIP-UPF1 to react with the CLIP-Biotin substrate and yield covalently biotinylated protein. Unbound CLIP-Biotin was subsequently removed by passing the samples through Zeba™ Spin Desalting Columns, 40K MWCO, 2 mL

(Thermo Scientific) according to manufacturer instructions. Buffer exchange was performed with HLB-150 supplemented with 1 mM DTT and 0.1% NP-40. Pre-washed Dynabeads™ MyOne™ Streptavidin T1 (Invitrogen) were then added and the samples rotated (end-over-end) for 1 hour at 4°C to immobilize the biotin-bound CLIP-UPF1 complexes. The samples were then washed three times with 500 µL of HLB-150 supplemented with 1 mM DTT and 0.1% NP-40 and one one-hundredth reserved for downstream Western blot analysis of CLIP-UPF1 pull-down efficiency. The remainder was combined with TRIzol™ (Invitrogen) and RNA was isolated according to manufacturer instructions. DNase-treatment was subsequently performed using RQ1 RNase-Free DNase (Promega) and RNA isolated by acid phenol chloroform extraction according to standard protocol. This resulted in approximately 1 µg of total RNA that was then subjected to high-throughput sequencing. In parallel, the reserved input lysate (approximately 3 µg of total RNA) was also sent for high-throughput sequencing. A total of three biological replicates from each condition were processed. Sequencing libraries were prepared from input and bound RNA using the Illumina TruSeq Stranded Total RNA Human kit and sequenced on an Illumina HiSeq 3000 instrument.

To validate select transcripts by RT-qPCR, an equivalent volume of reserved RNA (3 µL) was used as a template for cDNA synthesis using the Maxima First Strand cDNA synthesis kit for RT-qPCR (Thermo Scientific) according to manufacturer instructions. The resulting cDNA was diluted with nuclease-free water and subsequently analyzed by qPCR using iTaq Universal SYBR Green Supermix (BioRad) on a Roche LightCycler 96 instrument (Roche). Sequences for gene-specific primers used for amplification are listed in supplementary file 5. For input samples, relative fold change was determined by calculating $2^{-\Delta\Delta CT}$ values using GAPDH for normalization. For pull-down samples, relative fold enrichment was determined by dividing the C_q value of the pull-down by its corresponding input, multiplying by 100 and then normalizing to the relative recovery of the SMG1 transcript.

For assessment of CLIP-UPF1 expression and pull-down efficiency by Western blot, reserved input and IP samples were run on a NuPage™ 4-12% Bis-Tris Protein Gel (Invitrogen) using MOPS buffer according to manufacturer instructions and subsequently transferred to a nitrocellulose membrane according to the NuPage™ manufacturer's protocol (Invitrogen). Membranes were incubated with a blocking buffer for fluorescent Western blotting (Rockland) for 1 hour at room temperature and then incubated overnight at 4°C with the indicated primary antibody. Primary antibodies used were: anti-RENT1 (goat polyclonal, Bethyl, A300-038A, 1:1000) and anti-β-actin (mouse monoclonal, Cell Signaling, #3700, 1:1000). Membranes were subsequently washed three times with 1xTBS supplemented with 0.1% Tween-20 and then incubated with the appropriate secondary antibody for 1 hour at room temperature. Secondary antibodies used were: anti-goat IgG (H&L) Antibody DyLight™ 680 Conjugated (Rockland, 605-744-002, 1:10,000) and anti-mouse IgG (H&L) Antibody DyLight™ 680 Conjugated (Rockland, 610-744-124, 1:10,000). Membranes were washed three times with 1xTBS supplemented with 0.1% Tween-20 and then two times with 1xTBS. Western blot images were obtained on an Amersham Typhoon imaging system (GE Healthcare Life Sciences) and quantified using ImageStudio software (LI-COR Biosciences).

Endogenous UPF1$_{LL}$ knockdown with puromycin or thapsigargin treatment and RNA-seq

$3x10^5$ HEK-293 cells were reverse transfected with 40 nM of a non-targeting or UPF1$_{LL}$-specific (Forward sequence: 5' GGUAAUGAGGAUUUAGUCA 3'; Reverse sequence: 5' UGACUAAAUCCUCAUUACC 3') siRNA using Lipofectamine RNAiMAX according to manufacturer instructions. Forty-eight hours post-siRNA transfection, cells were replated at a density of $5x10^5$ cells per well of a 6-well plate. This step was critical to achieve 70-75% confluency on the day of drug treatment and cell harvest. The next day, cells were treated with vehicle control, 25, 50, or 100 µg/mL puromycin (Sigma) for 4 hours, or 1 µM thapsigargin for 6 hours. A total of three replicates were generated for each condition. Total RNA was then

isolated using the RNeasy Plus Mini Kit (QIAGEN). Sequencing libraries were prepared from 2 µg total RNA using the Illumina TruSeq Stranded Total RNA Human kit and sequenced on an NovaSeq 6000 instrument. For RT-qPCR validation of changes in target gene expression, 500 ng of total RNA was used as input for cDNA synthesis using the Maxima First Strand cDNA synthesis kit for RT-qPCR (Thermo Scientific) according to manufacturer instructions. The resulting cDNA was diluted with nuclease-free water and subsequently analyzed by qPCR using iTaq Universal SYBR Green Supermix (BioRad) on a Roche LightCycler 96 instrument (Roche). Sequences for gene-specific primers used for amplification are listed in supplementary file 5. Relative fold changes were determined by calculating $2^{-\Delta\Delta CT}$ values using GAPDH for normalization.

Western for phospho-eIF2α

HEK-293 cells treated with 1 µM thapsigargin were lysed in 1X Passive Lysis Buffer (Promega) supplemented with Halt™ Protease and Phosphatase Inhibitor Cocktail (ThermoScientific) according to manufacturer instructions. A total of 5 µg protein was run on a NuPage™ 4-12% Bis-Tris Protein Gel (Invitrogen) using MOPS buffer according to manufacturer instructions and subsequently transferred to a nitrocellulose membrane according to the NuPage™ manufacturer's protocol (Invitrogen). Detection of phospho and total eIF2α was performed as previously described (Young-Baird et al., 2020).

Semiquantitative PCR to detect UPF1 isoform ratios

Generated cDNA (1/40th) from RNA-seq samples was used as input for PCR amplification with Phusion High-Fidelity DNA polymerase (New England Biolabs) and UPF1 specific primers that flank the regulatory loop sequence (Forward: 5’ AACAAGCTGGAGGAGCTGTGGA 3’; Reverse: 5’ ACTTCCACACAAAATCCACCTGGAAGTT 3’). The PCR cycling conditions used were: an

initial denaturation at 98°C for 30 seconds and then 22 cycles of 98°C for 10 seconds, 63°C for 30 seconds, 72°C for 15 seconds. PCR products were then run on a 8% Novex™ TBE gel (Invitrogen) according to manufacturer instructions and subsequently stained with SYBR® Gold Nucleic Acid Stain (Invitrogen). Images were obtained on an Amersham Typhoon imaging system (GE Healthcare Life Sciences) and quantified using ImageStudio software (LI-COR Biosciences).

Computational analyses

Raw fastq reads from CLIP-UPF1 overexpression RNA-seq and RIP-seq data were trimmed with Cutadapt using the following parameters: --times 2 -e 0 -O 5 --quality-cutoff 6 -m 18 -a AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC -A AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCATT -b AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA -b TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT (Martin, 2011). Trimmed reads were quantified against a custom HEK-293 transcriptome index, prepared as described (Kishor et al., 2018), using Kallisto software with parameters --bias -b 1000 -t 16 --single --rf-stranded -l 200 -s 20 (Bray et al., 2016). RIP-seq enrichment values were obtained by dividing TPM values from IP samples by TPM values from input samples. Differential gene expression analysis was performed using RUVSeq and edgeR (Risso et al., 2014; Robinson et al., 2009). Normalization and batch correction was performed with the RUVg function, based on transcripts with invariant expression among all samples. The edgeR TMM method was used to obtain normalized differential expression values and to calculate FDRs. The most abundant transcript isoform from each gene was used for further analysis. PTBP1 and hnRNP L binding motif position-specific scoring matrices were downloaded from the RBPmap database and used for motif finding in 3'UTRs derived from the custom HEK-293 transcriptome with HOMER, as described (Heinz et al., 2010; Kishor et al., 2018; Paz et al., 2014). IsoformSwitchAnalyzeR was

used to annotate PTCs in the custom HEK-293 transcriptome, as described (Kishor et al., 2018; Vitting-Seerup and Sandelin, 2019), and NMD-sensitive transcripts were classified on the basis of a $\log_2$ fold change of 0.5 or greater in previous RNA-seq studies of total UPF1 knockdown in HEK-293 cells (GSE105436) (Baird et al., 2018). For analysis of RNA-seq data from HEK-293 cells treated with non-targeting or anti-$UPF1_{LL}$ siRNAs, raw fastq reads from the NovaSeq 6000 platform were trimmed with fastp (Chen et al., 2018), with the parameters --detect_adapter_for_pe and --trim_poly_g. Trimmed reads were aligned with HISAT2 to the hg19/GRCh37 genome and transcriptome index provided by the authors (Kim et al., 2019). Reads mapping to Ensembl GRCh37 release 75 gene annotations were quantified with featureCounts (Liao et al., 2014), and differential gene expression was analyzed using limma/voom, as implemented by the Degust server (Powell, 2015; Ritchie et al., 2015). Alternative splicing was analyzed using rMATS 4.0.1 (Shen et al., 2014), and Sashimi plots were generated using Integrative Genomics Viewer 2.8.2 (Thorvaldsdóttir et al., 2013).

GTEx data analysis

GTEx data were downloaded from the GTEx Portal on 11/11/2020. To determine the relative representation of $UPF1_{LL}$ and $UPF1_{SL}$ mRNA isoforms, transcript TPM values for transcript ENST00000599848.5 ($UPF1_{LL}$) were divided by the total TPM values derived from transcripts ENST00000599848.5 ($UPF1_{LL}$) and ENST00000262803.9 ($UPF1_{SL}$). GTEx samples were assigned to the indicated tissue types using the sample attributes provided in GTEx Analysis v8.

Data availability

CLIP-UPF1 overexpression RIP-seq and RNA-seq data are available from the NCBI GEO database with accession number GSE134059. The endogenous $UPF1_{LL}$ knockdown RNA-seq data, including puromycin or thapsigargin treatment, are available with accession number

GSE162699. All publicly available datasets used in this study are directly identified in the text and figure legends.

Acknowledgements

We thank Sutapa Chakrabarti for the UPF1$_{LL}$ΔCH recombinant expression plasmid and members of the Hogg lab and Nicholas R. Guydosh for critical reading of the manuscript. We are grateful to Sandy Mattijssen and Richard J. Maraia for helpful discussion and to Sara K. Young-Baird for thapsigargin reagents, helpful discussion, and assistance with phospho-eIF2α immunoblotting. High-throughput sequencing was conducted by Yan Luo and Poching Liu in the NHLBI DNA Sequencing and Genomics Core. The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. This work was supported by the Intramural Research Program, National Institutes of Health, National Heart, Lung, and Blood Institute and utilized the computational resources of the NIH HPC Biowulf cluster (http://hpc.nih.gov).

Author contributions

S.E.F. and J.R.H. conceived and designed the study, analyzed and interpreted the data, and wrote the manuscript. S.E.F. acquired the UPF1 RIP-seq and RNA-seq data and performed associated cellular studies. S.R. acquired the unwinding data. All authors read and approved the manuscript.

Competing interests

The authors declare no competing interests.

References

Akopian, D., Shen, K., Zhang, X., and Shan, S.-O. (2013). Signal recognition particle: an essential protein-targeting machine. Annu. Rev. Biochem. *82*, 693–721.

Alkallas, R., Fish, L., Goodarzi, H., and Najafabadi, H.S. (2017). Inference of RNA decay rate from transcriptional profiling highlights the regulatory programs of Alzheimer's disease. Nat. Commun. *8*, 909.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat. Genet. *25*, 25–29.

Baird, T.D., and Wek, R.C. (2012). Eukaryotic initiation factor 2 phosphorylation and translational control in metabolism. Adv. Nutr. *3*, 307–321.

Baird, T.D., Cheng, K.C.-C., Chen, Y.-C., Buehler, E., Martin, S.E., Inglese, J., and Hogg, J.R. (2018). ICE1 promotes the link between splicing and nonsense-mediated mRNA decay. Elife *7*, e33178.

Baker, S.L., and Hogg, J.R. (2017). A system for coordinated analysis of translational readthrough and nonsense-mediated mRNA decay. PLoS One *12*, e0173980.

Boswell, S.A., Snavely, A., Landry, H.M., Churchman, L.S., Gray, J.M., and Springer, M. (2017). Total RNA-seq to identify pharmacological effects on specific stages of mRNA synthesis. Nat. Chem. Biol. *13*, 501–507.

Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. Nat. Biotechnol. *34*, 525–527.

Causier, B., Li, Z., De Smet, R., Lloyd, J.P.B., Van de Peer, Y., and Davies, B. (2017). Conservation of Nonsense-Mediated mRNA Decay Complex Components Throughout Eukaryotic Evolution. Sci. Rep. *7*, 16692.

Chakrabarti, S., Jayachandran, U., Bonneau, F., Fiorini, F., Basquin, C., Domcke, S., Le Hir, H., and Conti, E. (2011). Molecular mechanisms for the RNA-dependent ATPase activity of Upf1 and its regulation by Upf2. Mol. Cell *41*, 693–703.

Chamieh, H., Ballut, L., Bonneau, F., and Le Hir, H. (2008). NMD factors UPF2 and UPF3 bridge UPF1 to the exon junction complex and stimulate its RNA helicase activity. Nat. Struct. Mol. Biol. *15*, 85–93.

Chan, W.K., Huang, L., Gudikote, J.P., and Chang, Y.F. (2007). An alternative branch of the nonsense-mediated decay pathway. EMBO J.

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics *34*, i884–i890.

Cheng, Z., Muhlrad, D., Lim, M.K., Parker, R., and Song, H. (2007). Structural and functional insights into the human Upf1 helicase core. EMBO J. *26*, 253–264.

Colombo, M., Karousis, E.D., Bourquin, J., Bruggmann, R., and Mühlemann, O. (2017).

Transcriptome-wide identification of NMD-targeted human mRNAs reveals extensive redundancy between SMG6- and SMG7-mediated degradation pathways. RNA *23*, 189–201.

Costa-Mattioli, M., and Walter, P. (2020). The integrated stress response: From mechanism to disease. Science *368*.

Dang, Y., Low, W.-K., Xu, J., Gehring, N.H., Dietz, H.C., Romo, D., and Liu, J.O. (2009). Inhibition of nonsense-mediated mRNA decay by the natural product pateamine A through eukaryotic initiation factor 4AIII. J. Biol. Chem. *284*, 23613–23621.

Dockendorff, T.C., and Labrador, M. (2019). The Fragile X Protein and Genome Function. Mol. Neurobiol. *56*, 711–721.

Durand, S., Franks, T.M., and Lykke-Andersen, J. (2016). Hyperphosphorylation amplifies UPF1 activity to resolve stalls in nonsense-mediated mRNA decay. Nat. Commun. *7*, 12434.

Eberle, A.B., Lykke-Andersen, S., Mühlemann, O., and Jensen, T.H. (2009). SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. Nat. Struct. Mol. Biol. *16*, 49–55.

Eden, E., Navon, R., Steinfeld, I., Lipson, D., and Yakhini, Z. (2009). GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics *10*, 48.

Fiorini, F., Bonneau, F., and Le Hir, H. (2012). Biochemical characterization of the RNA helicase UPF1 involved in nonsense-mediated mRNA decay. Methods Enzymol. *511*, 255–274.

Franks, T.M., Singh, G., and Lykke-Andersen, J. (2010). Upf1 ATPase-dependent mRNP disassembly is required for completion of nonsense- mediated mRNA decay. Cell *143*, 938–950.

Fritz, S.E., Haque, N., and Hogg, J.R. (2018). Highly efficient in vitro translation of authentic affinity-purified messenger ribonucleoprotein complexes. RNA *24*, 982–989.

Fritz, S.E., Ranganathan, S., Wang, C.D., and Hogg, J.R. (2020). The RNA-binding protein PTBP1 promotes ATPase-dependent dissociation of the RNA helicase UPF1 to protect transcripts from nonsense-mediated mRNA decay. J. Biol. Chem. *295*, 11613–11625.

Ge, Z., Quek, B.L., Beemon, K.L., and Robert Hogg, J. (2016). Polypyrimidine tract binding protein 1 protects mRNAs from recognition by the nonsense-mediated mRNA decay pathway. eLife Sciences *5*, e11155.

Goetz, A.E., and Wilkinson, M. (2017). Stress and the nonsense-mediated RNA decay pathway. Cell. Mol. Life Sci. *74*, 3509–3531.

Goldstein, J.L., and Brown, M.S. (2009). The LDL receptor. Arterioscler. Thromb. Vasc. Biol. *29*, 431–438.

Gowravaram, M., Bonneau, F., Kanaan, J., Maciej, V.D., Fiorini, F., Raj, S., Croquette, V., Le Hir, H., and Chakrabarti, S. (2018). A conserved structural element in the RNA helicase UPF1 regulates its catalytic activity in an isoform-specific manner. Nucleic Acids Res. *46*, 2648–2659.

Harbeck, N., Penault-Llorca, F., Cortes, J., Gnant, M., Houssami, N., Poortmans, P., Ruddy, K., Tsang, J., and Cardoso, F. (2019). Breast cancer. Nat Rev Dis Primers *5*, 66.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

Hodgkin, J., Papp, A., Pulak, R., Ambros, V., and Anderson, P. (1989). A new kind of informational suppression in the nematode Caenorhabditis elegans. Genetics *123*, 301–313.

Hogg, J.R., and Collins, K. (2007a). RNA-based affinity purification reveals 7SK RNPs with distinct composition and regulation. RNA *13*, 868–880.

Hogg, J.R., and Collins, K. (2007b). Human Y5 RNA specializes a Ro ribonucleoprotein for 5S ribosomal RNA quality control. Genes Dev. *21*, 3067–3072.

Hogg, J.R., and Goff, S.P. (2010). Upf1 senses 3′ UTR length to potentiate mRNA decay. Cell *143*, 379–389.

Hollerer, I., Curk, T., Haase, B., Benes, V., Hauer, C., Neu-Yilik, G., Bhuvanagiri, M., Hentze, M.W., and Kulozik, A.E. (2016). The differential expression of alternatively polyadenylated transcripts is a common stress-induced response mechanism that modulates mammalian mRNA expression in a quantitative and qualitative fashion. RNA *22*, 1441–1453.

Hollien, J., and Weissman, J.S. (2006). Decay of endoplasmic reticulum-localized mRNAs during the unfolded protein response. Science *313*, 104–107.

Huang, L., Lou, C.-H., Chan, W., Shum, E.Y., Shao, A., Stone, E., Karam, R., Song, H.-W., and Wilkinson, M.F. (2011). RNA homeostasis governed by cell type-specific and branched feedback loops acting on NMD. Mol. Cell *43*, 950–961.

Huntzinger, E., Kashima, I., Fauser, M., Saulière, J., and Izaurralde, E. (2008). SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. RNA *14*, 2609–2617.

Hurt, J.A., Robertson, A.D., and Burge, C.B. (2013). Global analyses of UPF1 binding and function reveal expanded scope of nonsense-mediated mRNA decay. Genome Res. *23*, 1636–1650.

Jan, C.H., Williams, C.C., and Weissman, J.S. (2014). Principles of ER cotranslational translocation revealed by proximity-specific ribosome profiling. Science *346*.

Karginov, F.V., and Hannon, G.J. (2013). Remodeling of Ago2-mRNA interactions upon cellular stress reflects miRNA complementarity and correlates with altered translation rates. Genes & Development *27*, 1624–1632.

Karousis, E.D., and Mühlemann, O. (2018). Nonsense-Mediated mRNA Decay Begins Where Translation Ends. Cold Spring Harb. Perspect. Biol.

Kashima, I., Yamashita, A., Izumi, N., Kataoka, N., Morishita, R., Hoshino, S., Ohno, M., Dreyfuss, G., and Ohno, S. (2006). Binding of a novel SMG-1-Upf1-eRF1-eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. Genes Dev. *20*, 355–367.

Keene, J.D. (2007). RNA regulons: coordination of post-transcriptional events. Nat. Rev. Genet.

8, 533–543.

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

Kim, D., Paggi, J.M., Park, C., Bennett, C., and Salzberg, S.L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat. Biotechnol. *37*, 907–915.

Kishor, A., Ge, Z., and Hogg, J.R. (2018). hnRNP L-dependent protection of normal mRNAs from NMD subverts quality control in B cell lymphoma. EMBO J.

Kishor, A., Fritz, S.E., and Robert Hogg, J. (2019). Nonsense-mediated mRNA decay: The challenge of telling right from wrong in a complex transcriptome. Wiley Interdisciplinary Reviews: RNA *10*.

Kishor, A., Fritz, S.E., Haque, N., Ge, Z., Tunc, I., Yang, W., Zhu, J., and Hogg, J.R. (2020). Activation and inhibition of nonsense-mediated mRNA decay control the abundance of alternative polyadenylation products. Nucleic Acids Res.

Kurosaki, T., Li, W., Hoque, M., Popp, M.W.-L., Ermolenko, D.N., Tian, B., and Maquat, L.E. (2014). A post-translational regulatory switch on UPF1 controls targeted mRNA degradation. Genes Dev. *28*, 1900–1916.

Lareau, L.F., Inada, M., Green, R.E., Wengrod, J.C., and Brenner, S.E. (2007). Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. Nature *446*, 926–929.

Lavysh, D., and Neu-Yilik, G. (2020). UPF1-Mediated RNA Decay-Danse Macabre in a Cloud. Biomolecules *10*, 999.

Lee, S.R., Pratt, G.A., Martinez, F.J., Yeo, G.W., and Lykke-Andersen, J. (2015). Target Discrimination in Nonsense-Mediated mRNA Decay Requires Upf1 ATPase Activity. Mol. Cell *59*, 413–425.

Le Hir, H., Moore, M.J., and Maquat, L.E.M. (2000). Pre-mRNA splicing alters mRNP composition: evidence for stable association of proteins at exon-exon junctions. Genes Dev. *14*, 1098–1108.

Le Hir, H., Izaurralde, E., Maquat, L.E., and Moore, M.J. (2000). The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. EMBO J. *19*, 6860–6869.

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

Loh, B., Jonas, S., and Izaurralde, E. (2013). The SMG5-SMG7 heterodimer directly recruits the CCR4-NOT deadenylase complex to mRNAs containing nonsense codons via interaction with POP2. Genes Dev. *27*, 2125–2138.

Longman, D., Jackson-Jones, K.A., Maslon, M.M., Murphy, L.C., Young, R.S., Stoddart, J.J., Hug, N., Taylor, M.S., Papadopoulos, D.K., and Cáceres, J.F. (2020). Identification of a localized nonsense-mediated decay pathway at the endoplasmic reticulum. Genes Dev.

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

Martinez-Nunez, R.T., Wallace, A., Coyne, D., Jansson, L., Rush, M., Ennajdaoui, H., Katzman, S., Bailey, J., Deinhardt, K., Sanchez-Elsner, T., et al. (2017). Modulation of nonsense mediated decay by rapamycin. Nucleic Acids Res. *45*, 3448–3459.

Mathews, M.B., and Hershey, J.W.B. (2015). The translation factor eIF5A and human cancer. Biochim. Biophys. Acta *1849*, 836–844.

Michlewski, G., and Cáceres, J.F. (2019). Post-transcriptional control of miRNA biogenesis. RNA *25*, 1–16.

Mugridge, J.S., Coller, J., and Gross, J.D. (2018). Structural and molecular mechanisms for the control of eukaryotic 5'-3' mRNA decay. Nat. Struct. Mol. Biol. *25*, 1077–1085.

Namkoong, S., Ho, A., Woo, Y.M., Kwak, H., and Lee, J.H. (2018). Systematic Characterization of Stress-Induced RNA Granulation. Mol. Cell *70*, 175–187.e8.

Ni, J.Z., Grate, L., Donohue, J.P., Preston, C., Nobida, N., O'Brien, G., Shiue, L., Clark, T.A., Blume, J.E., and Ares, M., Jr (2007). Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. Genes Dev. *21*, 708–718.

Nicholson, P., Josi, C., Kurosawa, H., Yamashita, A., and Mühlemann, O. (2014). A novel phosphorylation-independent interaction between SMG6 and UPF1 is essential for human NMD. Nucleic Acids Res. *42*, 9217–9235.

Page, M.F.P., Carr, B.C., Anders, K.R.A., Grimson, A.G., and Anderson, P.A. (1999). SMG-2 is a phosphorylated protein required for mRNA surveillance in Caenorhabditis elegans and related to Upf1p of yeast. Mol. Cell. Biol. *19*, 5943–5951.

Pakos-Zebrucka, K., Koryga, I., Mnich, K., Ljujic, M., Samali, A., and Gorman, A.M. (2016). The integrated stress response. EMBO Rep. *17*, 1374–1395.

Park, Y., Reyna-Neyra, A., Philippe, L., and Thoreen, C.C. (2017). mTORC1 Balances Cellular Amino Acid Supply with Demand for Protein Synthesis through Post-transcriptional Control of ATF4. Cell Reports *19*, 1083–1090.

Paz, I., Kosti, I., Ares, M., Jr, Cline, M., and Mandel-Gutfreund, Y. (2014). RBPmap: a web server for mapping binding sites of RNA-binding proteins. Nucleic Acids Res. *42*, W361–W367.

Powell, D. (2015). Degust: Visualize, explore and appreciate RNA-seq differential gene-expression data. In COMBINE RNA-Seq Workshop,.

Pulak, R., and Anderson, P. (1993). mRNA surveillance by the Caenorhabditis elegans smg genes. Genes Dev. *7*, 1885–1897.

Risso, D., Ngai, J., Speed, T.P., and Dudoit, S. (2014). Normalization of RNA-seq data using factor analysis of control genes or samples. Nat. Biotechnol. *32*, 896–902.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic

Acids Res. *43*, e47.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2009). edgeR : a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140.

Shen, S., Park, J.W., Lu, Z.-X., Lin, L., Henry, M.D., Wu, Y.N., Zhou, Q., and Xing, Y. (2014). rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. Proc. Natl. Acad. Sci. U. S. A. *111*, E5593–E5601.

Smith, J.E., and Baker, K.E. (2015). Nonsense-mediated RNA decay - a switch and dial for regulating gene expression. Bioessays *37*, 612–623.

Song, M.S., Salmena, L., and Pandolfi, P.P. (2012). The functions and regulation of the PTEN tumour suppressor. Nat. Rev. Mol. Cell Biol. *13*, 283–296.

The Gene Ontology Consortium (2019). The Gene Ontology Resource: 20 years and still GOing strong. Nucleic Acids Res. *47*, D330–D338.

Thorvaldsdóttir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief. Bioinform. *14*, 178–192.

Toma, K.G., Rebbapragada, I., Durand, S., and Lykke-Andersen, J. (2015). Identification of elements in human long 3' UTRs that inhibit nonsense-mediated decay. RNA *21*, 887–897.

Vitting-Seerup, K., and Sandelin, A. (2019). IsoformSwitchAnalyzeR: Analysis of changes in genome-wide patterns of alternative splicing and its functional consequences. Bioinformatics btz247.

Waldron, J.A., Tack, D.C., Ritchey, L.E., Gillen, S.L., Wilczynska, A., Turro, E., Bevilacqua, P.C., Assmann, S.M., Bushell, M., and Le Quesne, J. (2019). mRNA structural elements immediately upstream of the start codon dictate dependence upon eIF4A helicase activity. Genome Biol. *20*, 300.

Wek, R.C. (2018). Role of eIF2α Kinases in Translational Control and Adaptation to Cellular Stress. Cold Spring Harb. Perspect. Biol. *10*.

Yepiskoposyan, H., Aeschimann, F., Nilsson, D., Okoniewski, M., and Mühlemann, O. (2011). Autoregulation of the nonsense-mediated mRNA decay pathway in human cells. RNA *17*, 2108–2118.

Yi, Z., Sanjeev, M., and Singh, G. (2020). The Branched Nature of the Nonsense-Mediated mRNA Decay Pathway. Trends Genet.

Young, S.K., and Wek, R.C. (2016). Upstream Open Reading Frames Differentially Regulate Gene-specific Translation in the Integrated Stress Response. J. Biol. Chem. *291*, 16927–16935.

Young-Baird, S.K., Lourenço, M.B., Elder, M.K., Klann, E., Liebau, S., and Dever, T.E. (2020). Suppression of MEHMO Syndrome Mutation in eIF2 by Small Molecule ISRIB. Mol. Cell *77*, 875–886.e7.

Zhang, J., Harvey, S.E., and Cheng, C. (2019). A high-throughput screen identifies small molecule modulators of alternative splicing by targeting RNA G-quadruplexes. Nucleic Acids

1 Res. *47*, 3667–3679.

2 Zünd, D., Gruber, A.R., Zavolan, M., and Mühlemann, O. (2013). Translation-dependent
3 displacement of UPF1 from coding sequences causes its enrichment in 3' UTRs. Nat. Struct.
4 Mol. Biol. *20*, 936–943.

Figure 1-figure supplement 1

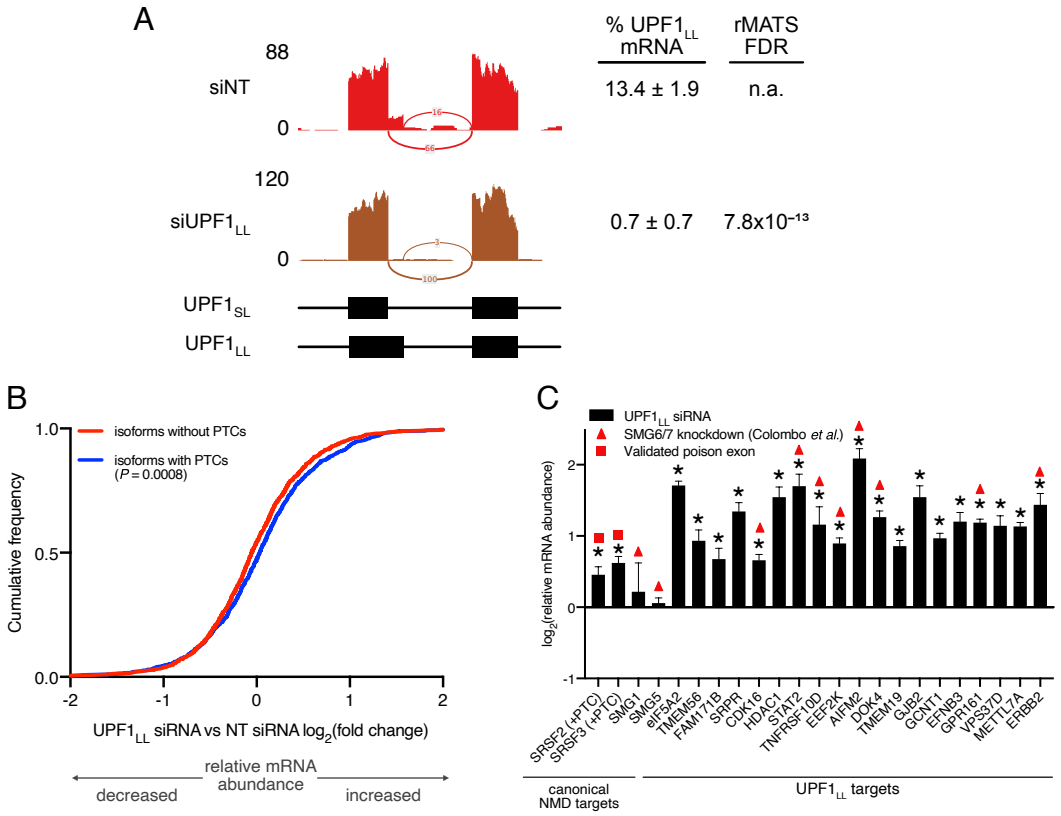

Figure 1-figure supplement 1: Alternative $UPF1_{LL}$ splice isoform contributes to NMD. (A) Sashimi plot from representative RNA-seq samples of siNT and $siUPF1_{LL}$ knockdown cells. Percent spliced in values and FDR were calculated with rMATS software. **(B)** CDF plot of relative mRNA abundance as determined from RNA-seq following $UPF1_{LL}$-specific knockdown in HEK-293 cells. Genes for which expression of an isoform containing a termination codon greater than 50 nt upstream of the final exon junction were analyzed, with mRNA isoforms classified as having or lacking a PTC. Statistical significance was determined by K-S test. **(C)** qPCR analysis of indicated transcripts from $UPF1_{LL}$ knockdown RNA-seq experiments. Relative fold changes are in reference to NT siRNA. Significance of $UPF1_{LL}$ knockdown was compared to the NT siRNA control. Asterisk (*) indicates $P < 0.05$, as determined by unpaired Student's t-test. Error bars indicate SD (n = 3). Transcripts that increased in abundance at least 1.4-fold in the SMG6/7 knockdown Colombo et al. dataset (2017) are indicated with a red triangle. mRNAs with validated poison exons (Lareau et al., 2007; Ni et al., 2007) are indicated with a red square.

Figure 2-figure supplement 1

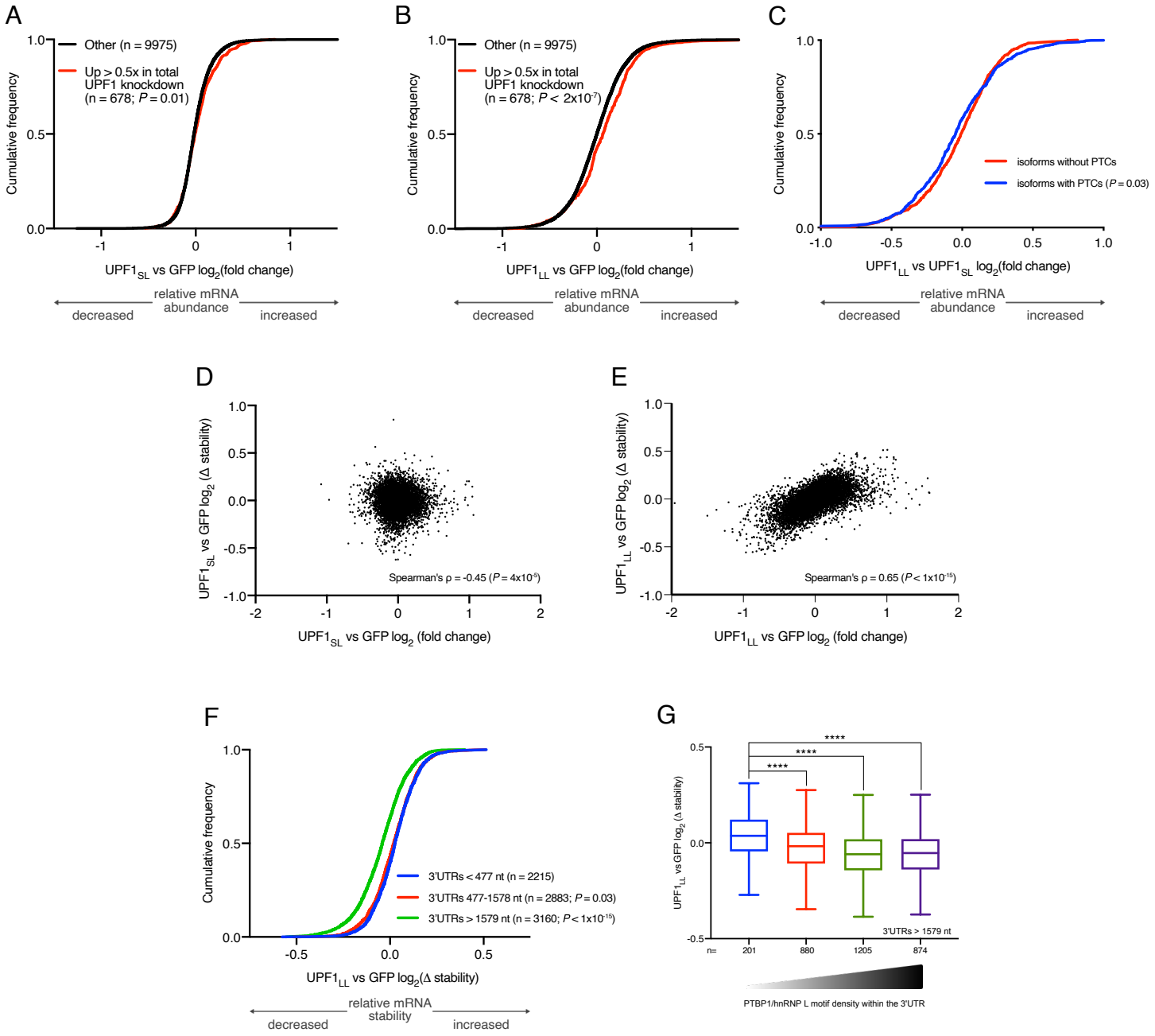

Figure 2-figure supplement 1: NMD protected transcripts are targeted by UPF1$_{LL}$. (A) CDF plot of relative mRNA abundance as determined from RNA-seq following CLIP-UPF1$_{SL}$ overexpression. mRNAs previously determined to be up-regulated upon total UPF1 knockdown (Baird et al., 2018) were analyzed. Statistical significance was determined by K-S test. **(B)** CDF plot as in (A), following CLIP-UPF1$_{LL}$ overexpression. **(C)** CDF plot of relative mRNA abundance as determined from RNA-seq following CLIP-UPF1$_{LL}$ vs CLIP-UPF1$_{SL}$ overexpression. Genes for which expression of an isoform containing a termination codon greater than 50 nt upstream of the final exon junction were analyzed, with mRNA isoforms classified as having or lacking a PTC. Statistical significance was determined by K-S test. **(D)** Scatterplot of changes in relative mRNA stability as determined by REMBRANDTS analysis of RNA-seq vs relative mRNA abundance as determined by kallisto, comparing CLIP-UPF1$_{SL}$ to GFP overexpression. **(E)** Scatterplot as in (D), comparing CLIP-UPF1$_{LL}$ to GFP overexpression. **(F)** CDF plot of relative mRNA stability as determined by REMBRANDTS analysis of RNA-seq following CLIP-UPF1$_{LL}$ or GFP overexpression. mRNAs were binned according to 3'UTR lengths (short, medium, long). Statistical significance was determined by K-W test, with Dunn's correction for multiple comparisons. **(G)** Box plot of relative mRNA stability as determined by REMBRANDTS analysis of RNA-seq for transcripts having long 3'UTRs (> 1579 nt) following CLIP-UPF1$_{LL}$ or GFP overexpression. mRNAs were binned by PTBP1 and/or hnRNP L motif density within the 3'UTR. Statistical significance was determined by K-W test, with Dunn's correction for multiple comparisons (**** $P < 7x10^{-8}$). Boxes indicate interquartile ranges, and bars indicate Tukey whiskers.

Figure 3-figure supplement 1

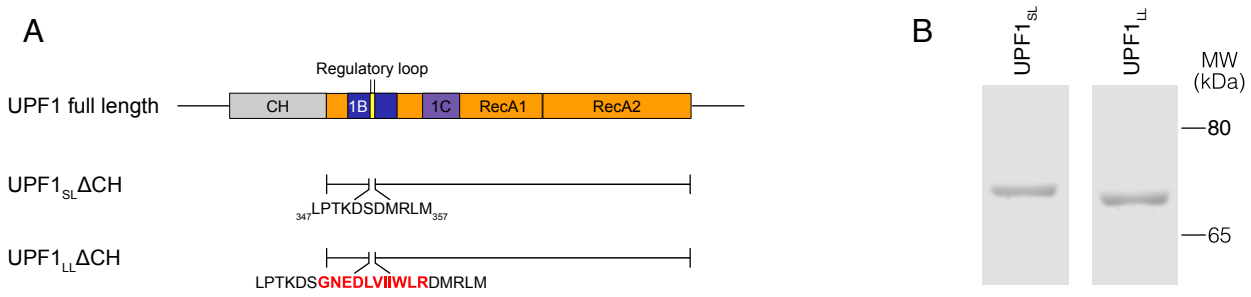

Figure 3-figure supplement 1: Reduced sensitivity of $UPF1_{LL}$ to translocation inhibition by PTBP1. **(A)** Schematic representation of the domain organization of full-length human UPF1, with the UPF1ΔCH constructs assessed in this study depicted below. The 11 amino acid extension in the regulatory loop of $UPF1_{LL}$ is indicated in red text. **(B)** SDS-PAGE of recombinant 6xHIS-CBP-$UPF1_{SL}$ΔCH (MW: 76 kDa) and 6xHIS-$UPF1_{LL}$ΔCH (MW: 73 kDa) proteins purified from *E. coli* and used in this study.

Figure 4-figure supplement 1

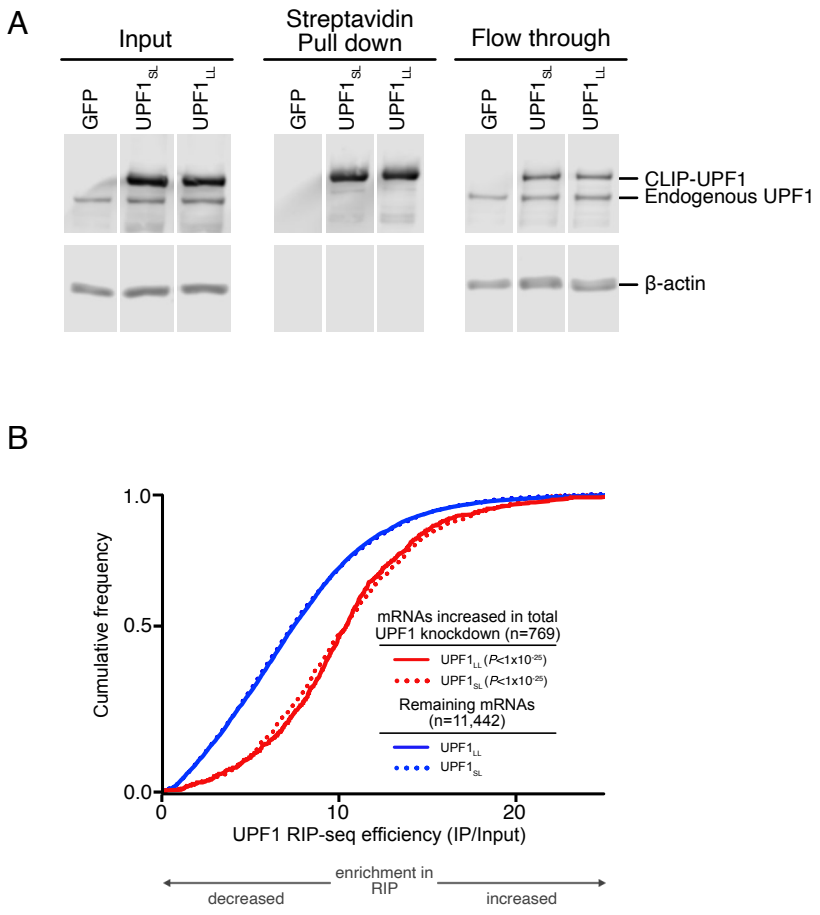

Figure 4-figure supplement 1: $UPF1_{LL}$ binds transcripts normally shielded by protective proteins. **(A)** Western blots of CLIP-$UPF1_{SL}$ and CLIP-$UPF1_{LL}$ cell lysates before (input) and after (pull down) biotin-streptavidin affinity purification. Input and flow through lanes represent 5% of total material. **(B)** CDF plot of recovered mRNAs as determined from RIP-seq efficiency in CLIP-$UPF1_{LL}$ or CLIP-$UPF1_{SL}$ affinity purifications. mRNAs were binned by sensitivity to total UPF1 knockdown (Baird et al., 2018). Statistical significance was determined by K-S test.

Figure 5-figure supplement 1

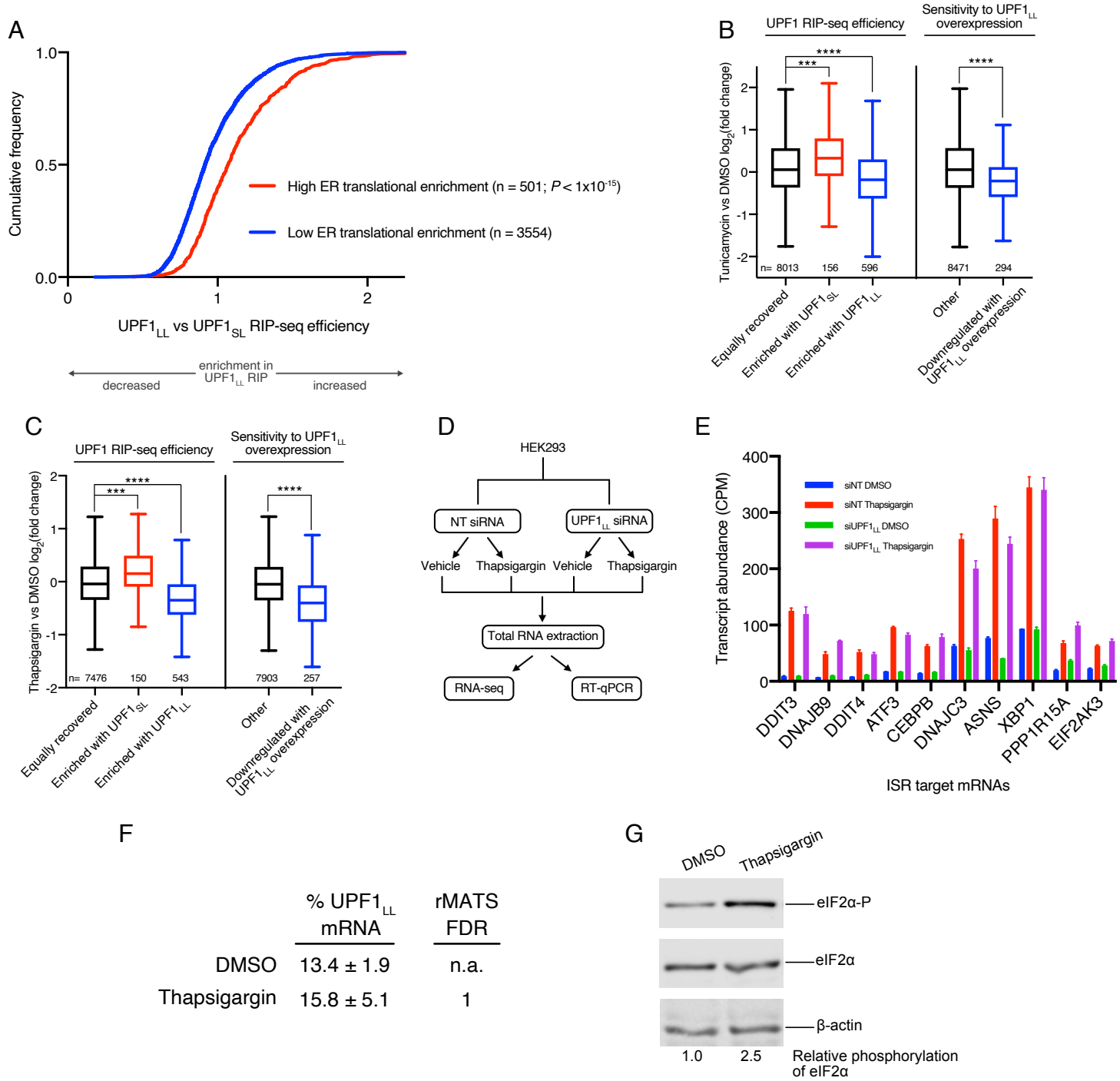

Figure 5-figure supplement 1: Transcripts targeted by $UPF1_{LL}$ are coordinately downregulated during ER stress and induction of the ISR. (A) CDF plot of recovered mRNAs as determined from RIP-seq efficiency in CLIP-$UPF1_{LL}$ or CLIP-$UPF1_{SL}$ affinity purifications. mRNAs were binned according to log_2 enrichment for translation at the ER (Jan et al., 2014). Statistical significance was determined by K-S test. **(B)** Box plot of relative mRNA abundance as determined from RNA-seq following treatment of HeLa cells with 30 μM tunicamycin for 6hr (GSE86857). mRNAs were binned by RIP-seq efficiency in CLIP-$UPF1_{LL}$ or CLIP-$UPF1_{SL}$ affinity purifications (left) or sensitivity to CLIP-$UPF1_{LL}$ overexpression (right). Statistical significance was determined by K-W test, with Dunn's correction for multiple comparisons (left; *** $P < 0.001$; **** $P < 1x10^{-14}$), or K-S test (right; **** $P < 1x10^{-12}$). Boxes indicate interquartile ranges, and bars indicate Tukey whiskers. **(C)** Box plot as in (B), following treatment of NIH3T3 cells with 1 μM thapsigargin for 1.5hr (GSE90869; **** $P < 1x10^{-6}$). **(D)** Schematic of the RNA-seq experimental workflow and conditions for $UPF1_{LL}$ knockdown and thapsigargin treatment. **(E)** Quantification of characterized ISR-target transcript abundance in RNA-seq of the indicated conditions. Error bars indicate SD (n = 3). **(F)** Quantification of $UPF1_{LL}$ isoform expression in control and thapsigargin-treated HEK-293 cells from rMATS analyses. **(G)** Western blot of eIF2α phosphorylation following treatment of HEK-293 cells with 1 μM thapsigargin for 6hr.

Figure 6-figure supplement 1

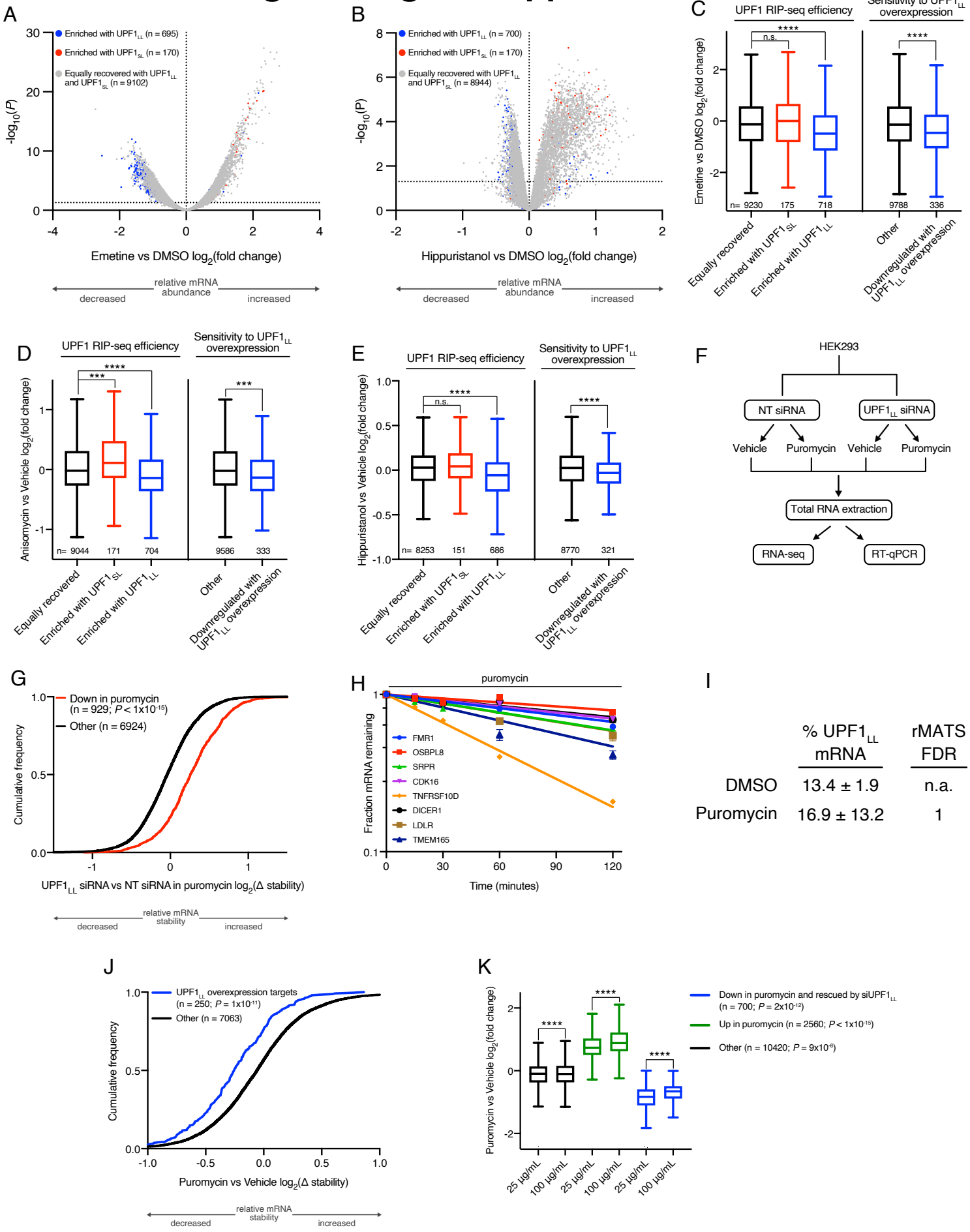

Figure 6-figure supplement 1

Figure 6-figure supplement 1: Decreased translation promotes UPF1$_{LL}$ activity. (A) Volcano plot of relative mRNA abundance as determined from RNA-seq following treatment of HEK-293 cells with 50 ng/mL emetine for 4hr (GSE89774). mRNAs were binned by RIP-seq efficiency in CLIP-UPF1$_{LL}$ or CLIP-UPF1$_{SL}$ affinity purifications. Horizontal line indicates the significance threshold $P \leq 0.05$. **(B)** Volcano plot as in (A), following treatment of MCF7 cells with 150 nM hippuristanol for 1hr (GSE134888). **(C)** Box plot of relative mRNA abundance as determined from RNA-seq following treatment of HMLE-TWIST cells with 10 μM emetine for 24hr (GSE113505). mRNAs were binned by RIP-seq efficiency in CLIP-UPF1$_{LL}$ or CLIP-UPF1$_{SL}$ affinity purifications (left) or sensitivity to CLIP-UPF1$_{LL}$ overexpression (right). Statistical significance was determined by K-W test, with Dunn's correction for multiple comparisons (left; **** $P < 1x10^{-15}$), or K-S test (right; **** $P < 1x10^{-7}$). Boxes indicate interquartile ranges, and bars indicate Tukey whiskers. **(D)** Box plot as in (C), following treatment of HEK-293 cells with 5 μg/mL anisomycin for 3hr (ERP010553; *** $P < 0.001$, **** $P < 1x10^{-11}$). **(E)** Box plot as in (C), following treatment of HEK-293 cells with 1 μM hippuristanol for 30min (GSE44378; **** $P < 1x10^{-15}$ (left), **** $P < 1x10^{-5}$ (right)). **(F)** Schematic of the RNA-seq experimental workflow and conditions for UPF1$_{LL}$ knockdown and puromycin treatment. **(G)** CDF plot of relative mRNA stability as determined by REMBRANDTS analysis of RNA-seq following UPF1$_{LL}$ knockdown in HEK-293 cells and treatment with 50 μg/mL puromycin for 4hr. mRNAs were binned by sensitivity to puromycin in the absence of UPF1$_{LL}$ knockdown. Statistical significance was determined by K-S test. **(H)** Fraction of mRNA remaining, as determined by RT-qPCR, for indicated transcripts following treatment of HEK-293 cells with 50 μg/mL puromycin over the course of 2hr. Error bars indicate SD (n = 3). **(I)** Quantification of UPF1$_{LL}$ isoform expression in control and puromycin-treated HEK-293 cells from rMATS analyses. **(J)** CDF plot of relative mRNA stability as determined by REMBRANDTS analysis of RNA-seq following treatment of HEK-293 cells with 50 μg/mL puromycin for 4hr. mRNAs were binned by sensitivity to CLIP-UPF1$_{LL}$ overexpression. Statistical significance was determined by K-S test. **(K)** Box plot of relative mRNA abundance as determined by RNA-seq following treatment of HEK-293 cells with 25 μg/mL or 100 μg/mL puromycin. mRNAs were binned according to sensitivity to 50 μg/mL puromycin and UPF1$_{LL}$ knockdown. Statistical significance was determined by K-S test. Boxes indicate interquartile ranges, and bars indicate Tukey whiskers.