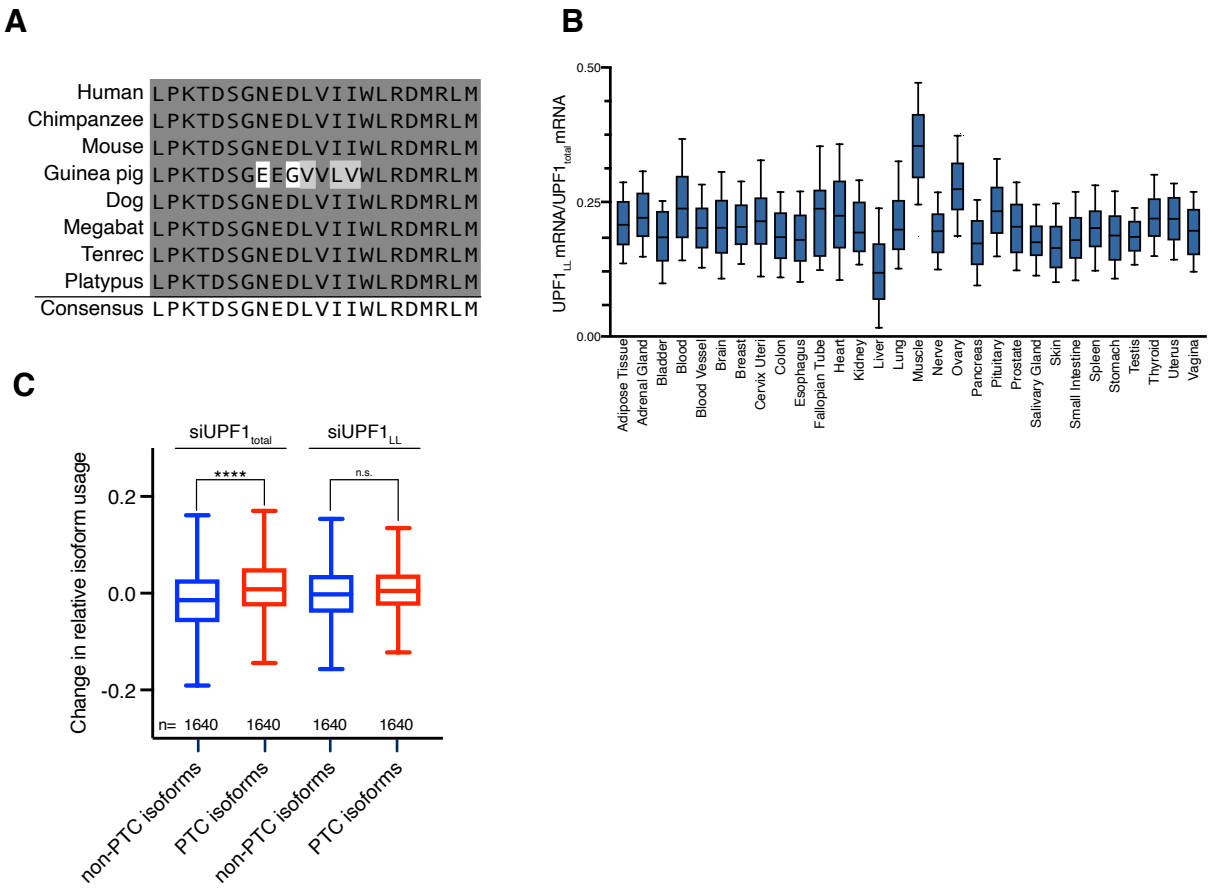
Appendix

Table of Contents

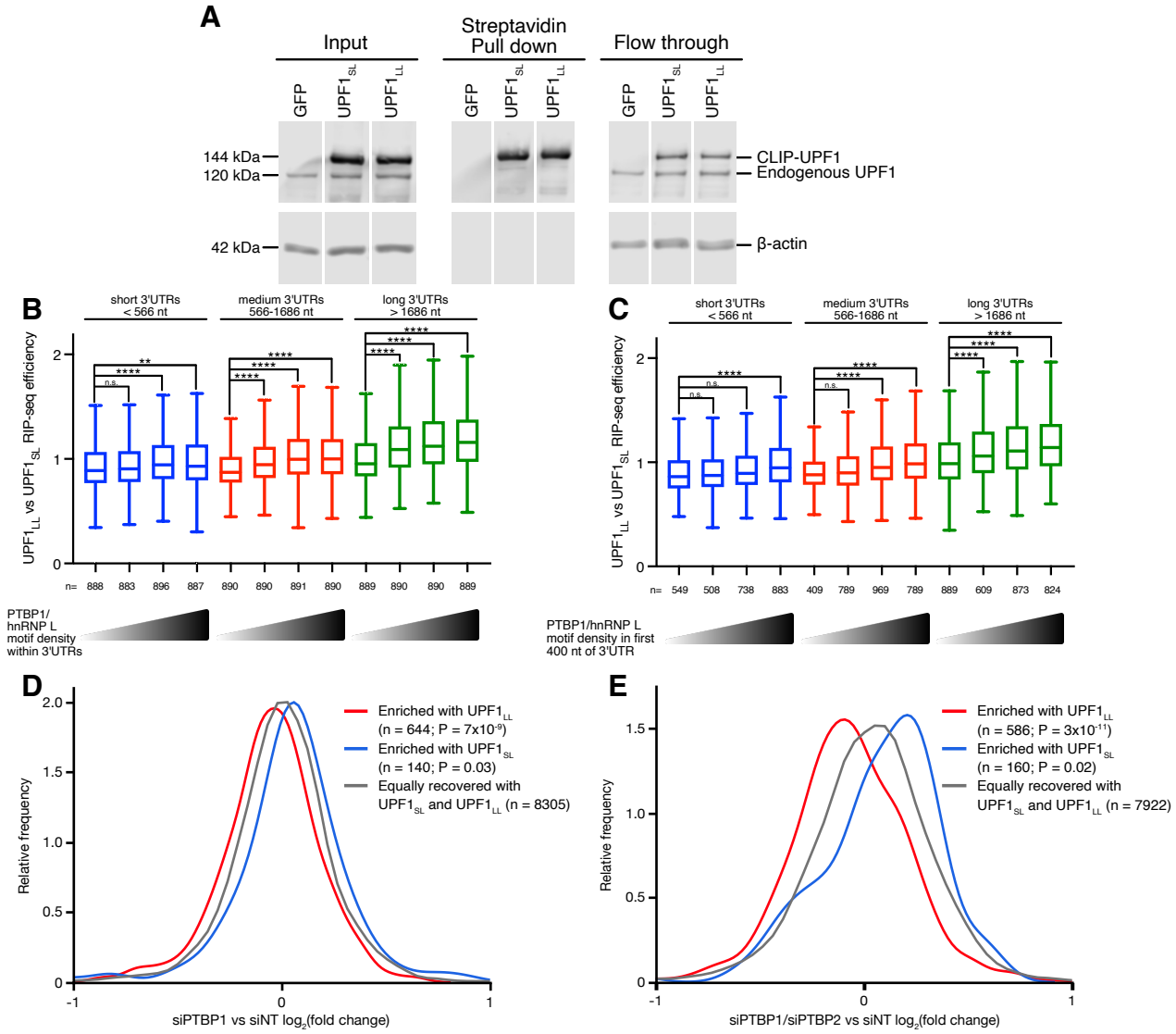


Appendix Figure S1. Conserved UPF1$_{LL}$ does not systematically regulate EJC-enhanced NMD.

A. UPF1$_{LL}$ regulatory loop sequence alignment across mammals. Indicated UPF1$_{LL}$ regulatory loop amino acid sequences were retrieved from the UCSC Genome Browser and aligned using T-Coffee. Conserved residues are shaded grey.

B. Box plot of the fraction of UPF1$_{LL}$/UPF1$_{total}$ 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. Box plot of the change in relative isoform usage of PTC-containing transcripts as determined from RNA-seq following total UPF1 (siUPF1$_{total}$) or 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 premature termination codon (PTC). Statistical significance was determined by one-way ANOVA (**** $P < 1x10^{-15}$). Boxes indicate interquartile ranges, horizontal line indicates the statistical median, and bars indicate Tukey whiskers.



Appendix Figure S2. UPF1$_{LL}$ can bind transcripts normally shielded by protective proteins.

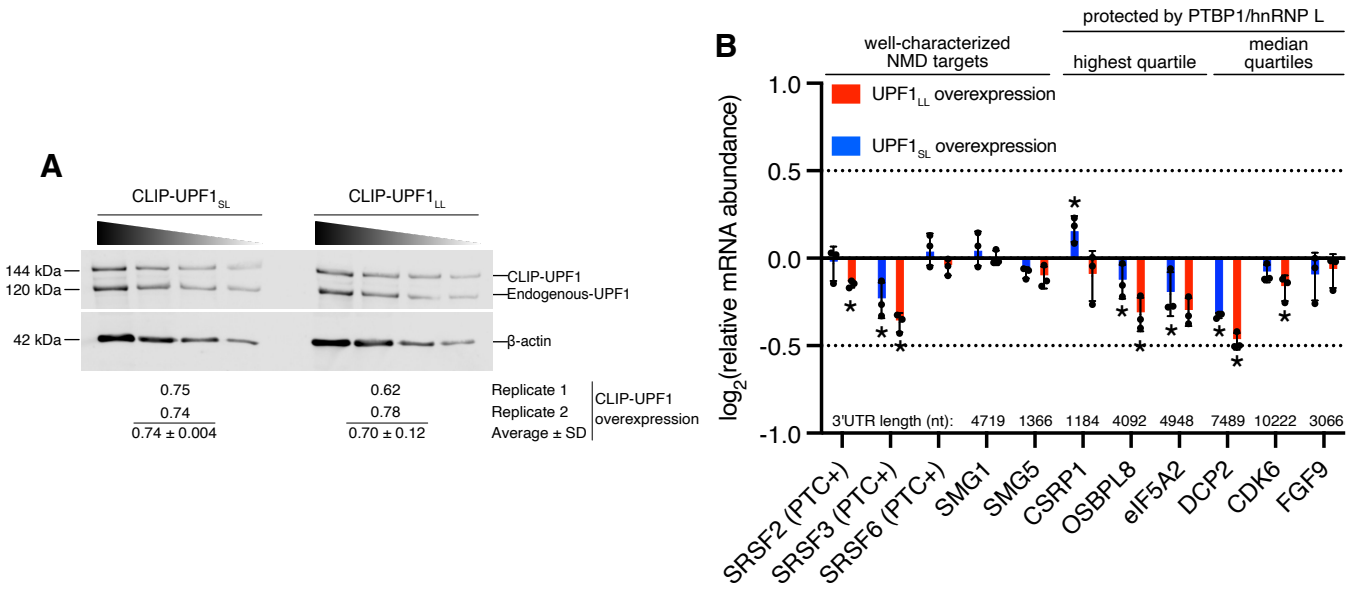
A. Western blots of CLIP-UPF1$_{SL}$ and CLIP-UPF1$_{LL}$ cell lysates before (input) and after (pull down) biotin-streptavidin affinity purification. Input and flow through lanes represent 5% of total material.

B. 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 equally subdivided by PTBP1 and/or hnRNP L motif density within the 3'UTR, as indicated by the gradient triangles. Statistical significance was determined by K-W test, with Dunn's correction for multiple comparisons (** $P < 0.002$, **** $P < 5x10^{-6}$). Boxes indicate interquartile ranges, horizontal line indicates the statistical median, and bars indicate Tukey whiskers.

C. Box plot as in (B), except that 3'UTR length bins were subdivided by PTBP1 and/or hnRNP L motif density with in the first 400 nt of the 3'UTR. Statistical significance was determined by K-W test, with Dunn's correction for multiple comparisons (**** $P < 6x10^{-5}$). Boxes indicate interquartile ranges, horizontal line indicates the statistical median, and bars indicate Tukey whiskers.

D. Density plot of changes in relative mRNA abundance as determined by RNA-seq following PTBP1 knockdown in HEK-293 cells (Data ref: Ge *et al*, 2016). mRNAs were binned by RIP-seq efficiency in CLIP-UPF1$_{LL}$ or CLIP-UPF1$_{SL}$ affinity purifications. Statistical significance was determined by K-W test, with Dunn's correction for multiple comparisons.

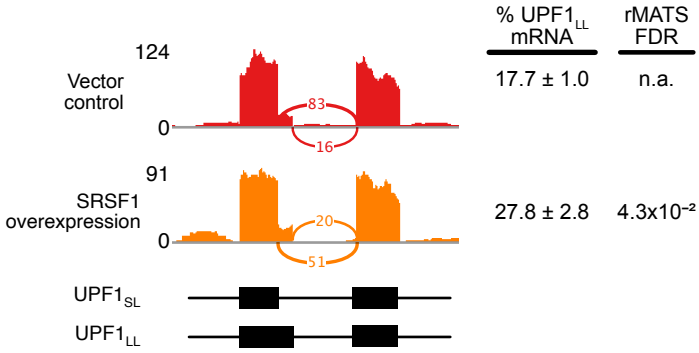
E. Density plot as in (D), following PTBP1 and PTBP2 knockdown in mouse neuronal progenitor cells (Data ref: Linares *et al*, 2015).



Appendix Figure S3. UPF1$_{LL}$ overexpression down-regulates NMD-protected mRNAs in a dose-dependent manner.

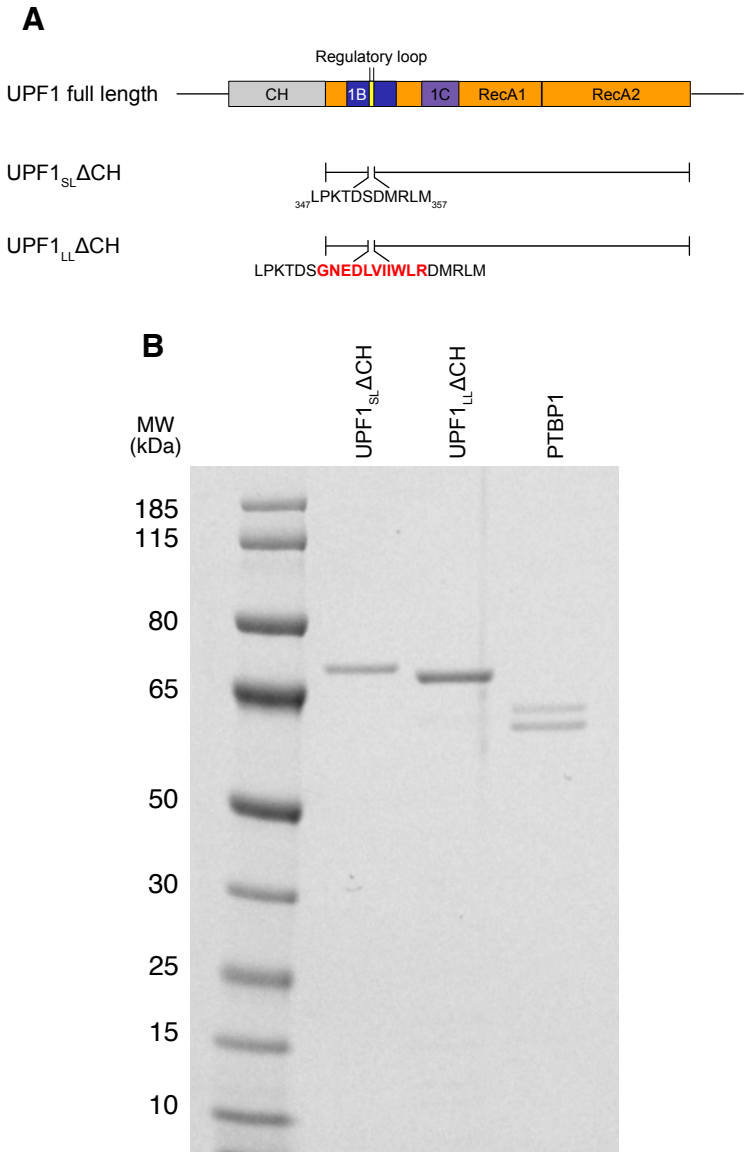
A. Western blots of reduced CLIP-UPF1$_{SL}$ and CLIP-UPF1$_{LL}$ overexpression. Membranes were probed with an anti-UPF1 antibody that detects both endogenous and CLIP-tagged UPF1. Wedge indicates serial two-fold dilutions of lysate. Mean (± standard deviation) of CLIP-UPF1 overexpression was determined from two replicate membranes.

B. RT-qPCR analysis of indicated transcripts from reduced CLIP-UPF1 overexpression experiments. Relative fold changes are in reference to a parental control line. Significance of CLIP-UPF1$_{SL}$ or CLIP-UPF1$_{LL}$ overexpression was compared to the parental control line. Asterisk (*) indicates $P < 0.05$, as determined by two-way ANOVA. Black dots represent individual data points and error bars indicate mean±SD (n = 3 biological replicates). Dashed lines indicate $\log_2$(fold change) of ±0.5. For protected mRNAs, the motif density of PTBP1/hnRNP L within the 3'UTR is indicated. PTC+ indicates the use of primers specific to transcript isoforms with validated poison exons (Lareau et al., 2007; Ni et al., 2007). See also Dataset EV3 for *P* values associated with each statistical comparison.



Appendix Figure S4. $UPF1_{LL}$ expression is regulated by SRSF1.

Sashimi plot from representative RNA-seq samples of SRSF1 overexpression (Data ref: Caputi *et al*, 2019). Percent spliced in values and FDR were calculated with rMATS software (n = 3 biological replicates).



Appendix Figure S5. Reduced sensitivity of UPF1$_{LL}$ to translocation inhibition by PTBP1.

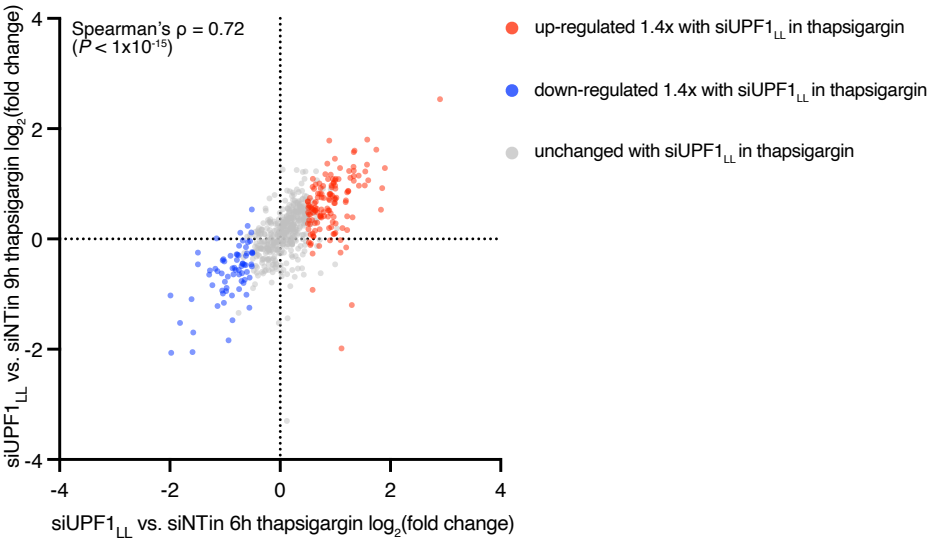
A. Schematic representation of the domain organization of full-length human UPF1, with the UPF1ΔCH constructs assessed in this study depicted below. The 11 amino acid extension in the regulatory loop of UPF1$_{LL}$ is indicated in red text.

B. SDS-PAGE of recombinant 6xHIS-CBP-UPF1$_{SL}$ΔCH (MW: 76 kDa), 6xHIS-UPF1$_{LL}$ΔCH (MW: 73 kDa), and 6xHIS-PTBP1 (MW: 62 kDa) proteins purified from *E. coli* and used in this study. PTBP1 doublet consists of differentially migrating conformers of identical MW, as determined by mass spectrometry of bands excised from SDS-PAGE gels (Fritz et al., 2020).

A

	606 genes down 1.4x (FDR < 0.05) in thapsigargin		
	Up 1.4x (FDR < 0.05)	Down 1.4x (FDR < 0.05)	Unchanged
siUPF1$_{LL}$ vs. siNT in 6h thapsigargin	135	70	401
siUPF1$_{LL}$ vs. siNT in 9h thapsigargin	143	62	401

B



Appendix Figure S6. mRNAs selectively down-regulated during ER stress and induction of the ISR are systematically dependent upon UPF1$_{LL}$ for their regulation.

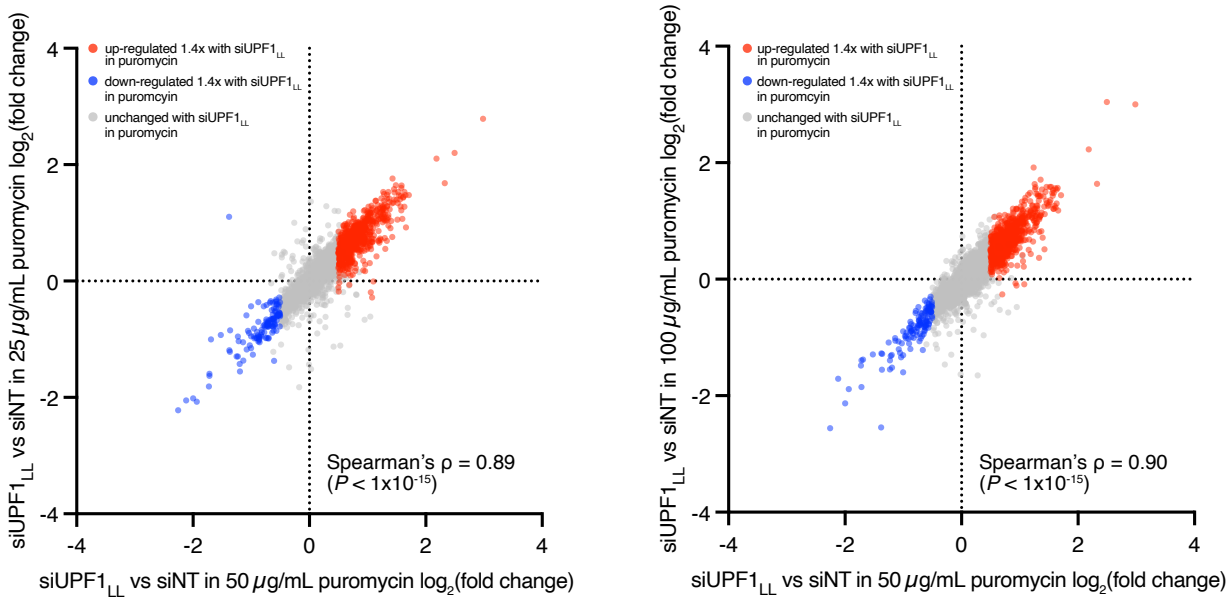
A. RNA-seq analysis of HEK-293 cells treated with 1 μM thapsigargin identified a population of 606 genes that decreased in abundance, of which a 135 (6h) or 143 (9h) were rescued by UPF1$_{LL}$-specific knockdown. Indicated are the number of genes from this population of 606 that increased or decreased in abundance at least 1.4-fold (FDR < 0.05) with UPF1$_{LL}$-specific knockdown following 6 or 9hr treatment in thapsigargin. At least 2-fold more genes were selectively up-regulated with siUPF1$_{LL}$ compared to those that were down-regulated with UPF1$_{LL}$-specific depletion in thapsigargin.

B. Scatterplot of changes in abundance of the 606 mRNAs that decreased in abundance upon 1 μM thapsigargin treatment (6h) as determined by RNA-seq following knockdown of UPF1$_{LL}$ in HEK-293 cells and treatment with 1 μM thapsigargin, comparing 6hr vs 9hr of drug treatment. Genes are categorized as being up or down-regulated 1.4-fold (FDR < 0.05) in response to UPF1$_{LL}$ knockdown in thapsigargin treatment. Dashed lines indicate the center-point of zero on both axes.

A

	2279 genes down 1.4x (FDR < 0.05) in puromycin		
	Up 1.4x (FDR < 0.05)	Down 1.4x (FDR < 0.05)	Unchanged
siUPF1$_{LL}$ vs. siNT in 25 μg/mL puromycin	626	167	1486
siUPF1$_{LL}$ vs. siNT in 50 μg/mL puromycin	700	124	1455
siUPF1$_{LL}$ vs. siNT in 100 μg/mL puromycin	585	192	1502

B



Appendix Figure S7. Downregulation of select genes during partial translational repression is dependent upon UPF1$_{LL}$ activity.

A. RNA-seq analysis of HEK-293 cells treated with 50 μg/mL puromycin identified a population of 2279 genes that decreased in abundance, of which 700 were rescued by UPF1$_{LL}$-specific knockdown. Indicated are the number of genes from this population of 2279 that increased or decreased in abundance at least 1.4-fold (FDR < 0.05) with UPF1$_{LL}$-specific knockdown following treatment with 25, 50, or 100 μg/mL puromycin. At least 3-fold more genes were selectively up-regulated with siUPF1$_{LL}$ compared to those that were down-regulated with UPF1$_{LL}$-specific depletion in puromycin.

B. Scatterplot of changes in relative abundance of the 2279 mRNAs decreased in abundance upon treatment with 50 μg/mL puromycin, as determined by RNA-seq following knockdown of UPF1$_{LL}$ in HEK-293 cells and treatment with puromycin, comparing 50 μg/mL vs 25 μg/mL of drug treatment (left) or 50 μg/mL vs 100 μg/mL of drug treatment (right). Genes are categorized as being up or down-regulated 1.4-fold (FDR < 0.05) in response to UPF1$_{LL}$ knockdown in puromycin treatment. Dashed lines indicate the center-point of zero on both axes.

Appendix Table S1. NMD-inducing features of genes common and unique to $UPF1_{total}$ or $UPF1_{LL}$-specific depletion under normal cellular conditions.

		Response to $UPF1_{total}$ or $UPF1_{LL}$-specific knockdown in HEK-293 cells under normal cellular conditions			
NMD-inducing feature		Up >1.4x in $siUPF1_{LL}$ only	Up >1.4x in $siUPF1_{total}$ only	Up >1.4x in $siUPF1_{total}$ & $siUPF1_{LL}$	Other
Gene does not have PTC	Count	892	1000	379	8747
	Column %	86.85491724	74.57121551	74.0234375	84.4957496
	Chi-Square statistic	1.58699552	12.18014599	5.254724602	1.87448916
	Chi-Square probability	$P > 0.10$	$P > 0.005$	$P > 0.10$	$P > 0.10$
Gene has PTC	Count	135	341	133	1605
	Column %	13.14508277	25.42878449	25.9765625	15.5042504
	Chi-Square statistic	7.897703992	60.61465607	26.15020581	9.32841988
	Chi-Square probability	$P > 0.10$	$P < 0.005$	$P < 0.005$	$P > 0.025$
Gene does not have uORF	Count	392	408	157	4099
	Column %	38.16942551	30.42505593	30.6640625	39.5962133
	Chi-Square statistic	0.000451227	21.27172543	7.630635798	5.20298701
	Chi-Square probability	$P > 0.995$	$P < 0.005$	$P > 0.05$	$P > 0.10$
Gene has uORF	Count	635	933	355	6253
	Column %	61.83057449	69.57494407	69.3359375	60.4037867
	Chi-Square statistic	0.000279037	13.1543351	4.718749339	3.21750273
	Chi-Square probability	$P > 0.995$	$P < 0.005$	$P > 0.10$	$P > 0.10$
Gene does not have long 3'UTR	Count	719	982	371	7274
	Column %	70.0097371	73.22893363	72.4609375	70.2666151
	Chi-Square statistic	0.056265363	1.280608911	0.242527829	0.19545803
	Chi-Square probability	$P > 0.99$	$P > 0.10$	$P > 0.95$	$P > 0.975$
Gene has long 3'UTR (> 1686 nt)	Count	308	359	141	3078
	Column %	29.9902629	26.77106637	27.5390625	29.7333849
	Chi-Square statistic	0.13532066	3.079920454	0.583290039	0.47008511
	Chi-Square probability	$P > 0.975$	$P > 0.10$	$P > 0.90$	$P > 0.90$
Gene does not have NMD-inducing feature	Count	266	256	93	2748
	Column %	25.9006816	19.09023117	18.1640625	26.5455951
	Chi-Square statistic	0.095058947	21.11092262	10.59341163	5.20032654
	Chi-Square probability	$P > 0.995$	$P < 0.025$	$P > 0.10$	$P > 0.10$
Gene has one NMD-inducing feature	Count	471	609	233	4624
	Column %	45.8617332	45.41387025	45.5078125	44.6676971
	Chi-Square statistic	0.225803577	0.088893323	0.04663951	0.09303052
	Chi-Square probability	$P > 0.995$	$P > 0.995$	$P > 0.995$	$P > 0.995$
Gene has two NMD-inducing features	Count	263	404	162	2628
	Column %	25.60856865	30.12677107	31.640625	25.3863988
	Chi-Square statistic	0.10526875	8.215418485	5.959624576	2.16777296
	Chi-Square probability	$P > 0.995$	$P > 0.10$	$P > 0.10$	$P > 0.975$
Gene has three NMD-inducing features	Count	27	72	24	352
	Column %	2.629016553	5.369127517	4.6875	3.40030912
	Chi-Square statistic	2.640811425	11.82717624	1.718632979	1.03526576
	Chi-Square probability	$P > 0.975$	$P > 0.10$	$P > 0.995$	$P > 0.995$

Appendix Table S2. qPCR primers

Gene	Forward Primer (5' to 3')	Reverse Primer (5' to 3')
AIFM2	CCTTGAGGGACTTTGTGGTT	CATAGTCGGTGGCAGACATTAG
ARID3A	ACAGAAAGAGGCAGACGTTTC	ACACAGTGCAGGGATGTATG
ATF4	TCAAACCTCATGGGTTCTCC	GTGTCATCCAACGTGGTCAG
CDK16	CATGTTCACCTGCCCACTT	GACCAAGTGAAGGAGTGATGAG
CDK6	CAGTTCTTCCATCGGTGTAGTT	GCTTGTGTTGGATCTGGATTTC
COL4A1	CCTGCTTCATTGACCTCTACTT	TGGAGTTCTCACTTCACACATC
CSPG4	AACCAGGGTAACCTCCTACA	TCCTTCTCCTTGCCCTCTTA
CSRP1	ACTTGAACTTGGGCATCTGG	ACCTAGACCCAAAACACTGAGG
CUL2	AGACTAGTACCCTAGCTCTGTG	CTGTTGGAGGAGGGTTGTAAG
DCP2	CTCCTCAAGCCTTACCTTTCTC	GTGCAAACTCGTGTTTCTTTCT
DICER1	CAGTCTTGCACTCCCATCAA	CACCTCTGCTCAGCTTTCTT
EFNB3	ATGACAGCTACCATGAGAAGAAG	GCATCAGACCAAGTCCATACA
eIF5A2	GTGAGGTCAAGTGCCTGAAA	AGGTGAAAGCCTGGTATGC
ERBB2	GGAGTCTTTGTGGATTCTGAGG	CTTCCCTAAGGCTTTCAGTACC
FAM171B	CTGCTGTCTCGTGCTGTTTA	GAGACATGTCCAAGAACCATGA
FGF9	CACGCGGTGGGTTCTTATT	CCATCCAAGCCTCCATCATAC
FMR1	CTTAGCACTTCAGGGCAGATT	TGATCACCCAATGCAGGAAA
GAPDH	AACATCATCCCTGCCTCTACTGG	GTTTTTCTAGACGGCAGGTCAGG
GCNT1	TCGAGATGATGAGGGTGGTA	GCTAAGAGAAGCATGGGAGAA
GIGYF2	GCACTGGCTGGTTCTGTATT	CAGGAGCTGATCCACATGTTAG
GJB2	CCTGTTTCAGAGGCTCAGATT	CAGCTGTGCTGCAAAGTATTC
GPR161	CTCTGGATGGACCACAAGATAC	CTCACTGCCTCCTCTTTATAGC
HDAC1	GGTCTTCAAGGATCTCCTGTTT	CAGTGTTTCTTGGCTCACTTG
HERC3	GCTGAGTCTGGGTCCATTATT	GCCATGGGAAAGTTGCTATTC
LDLR	GCCTCTGAAATGCCTCTTCT	CCCAGAAGCCACTCATACATAC
MAP1A	GCTGTGGCTTGGTGTTTAATG	CAGGATGGCATAGTGGAATGT
MED17	AAGGACAGCTGAAGGCAATAG	GTCTCAACTGGTGAGGTCAATC
METTL7A	CCGAGATCATGCCATTGTACT	CTGTCTCCCTGTCTCTACTTCT
NFKB1	TCCCTCTGCTACGTTCCTATT	ATGGCACATCAAGTGACTCTC
NXT2	CACTGTGAACCCAGCCTATT	CCACTTCTTCCCAGAGTGTTAT
OSBPL8	TGCCAGGACATTTCTAGCTTT	GACTACTGTCAGGTAGGAGTACA
PTEN	TGTAATCAAGGCCAGTGCTAAA	AGCATCCACAGCAGGTATTATG
RICTOR	TATGCAGAAGCACCCTTCAC	CAGCCTTAGAGACACTGATTCC
SMG1	AAATGACCCTGCAGCGATAC	GCTTCCTGTTTCAAGCGTTC
SMG5	GCCTGGATTTGCTGAAGAAG	TCAAAGCTCTTTCCCACCTC
SMG6	AGGCTTCCTTACCAAGAATGAG	CCCACTGGTCCCATTAAGATT
SMG7	CTGGGAATGAAGGCTCCATAAA	CCTCTTGACAGCGAGAAACA
SRPR	CATCTGCAAGGAAGGCCTAATC	TGAGTCTGTAGGCAGAGTGAAG
SRSF2 (PTC+)	GGCGTGTATTGGAGCAGATGTA	CTGCTACACAACTGCGCCTTTT
SRSF3 (PTC+)	TGGAACTGTCGAATGGTGAAA	GAGACATGATGGTGACTCTGC
SRSF6 (PTC+)	GGAAGCCGCATGACCAA	GGCAAGGGTCACACAATGTA
STAT2	CTGGACAATCTCACACCTTAGC	TGCAGGGCCTTTCCTTTATATC
TBL1XR1	CATCAAGAGAGGCAGTCATTCA	TCCTACGTAAGTCACAGGGTAA
TGOLN2	ATACCGCACATCAGAGGAAAG	CTCATCCACACAATCGCTAGAA
TIMP2	CTGCATCGTGGAAGCATTTG	AAGACGGGAGACGAATGAAAG
TMEM165	CCTGAGCCAGTAAACAGTAGTT	CAGTTCTTGTTGTGCCATGTTAT
TMEM19	CTCAGCCTCCCAAAGATCTATTAC	AGTGTGAAACCCAGACATCAG
$UPF1_{total}$	GGACCTGGGAGATTTGAGAAG	AAGAAGCCACTGGAAGCTAAA
$UPF1_{LL}$	GACTCTGGTAATGAGGATTTAGTCATA	CGTGGCCGATCCCTTTC
UPF2	CAAAAGCCAGCTGAGGAAAG	GACAAACGGCATTGTGTCTG
TNFRSF10D	CATGGCATCAAGAGGGAAGA	CAGTTCACAGACCACCAGAA
VPS37D	CATGTCCTGAGCCTACCTTTC	GCATGTGTGGACACGTAGTAG