Supporting Information (SI)

SI Materials and Methods.

Computational and statistical analyses. Refseq gene annotations were downloaded from the UCSC table browser for both human (hg19) and mouse (mm10). Repeat Masker tracks were also downloaded from the UCSC table browser for both organisms, after which only *Alu* SINEs were retained for humans whereas only B1, B2, B4 and ID SINEs were retained for mouse. The Human:Mouse Orthology table was downloaded from The Jackson Laboratory - Mouse Genome Informatics (http://www.informatics.jax.org) (1).

To identify 3' UTR SINEs, intersections between annotated transcript 3' UTRs and SINEs were calculated, discarding non protein-coding transcripts and transcripts containing intronic SINEs. SINEs in both sense and antisense orientations relative to the transcribed strand were included.

Human:mouse orthologous pairs were divided into four groups: (1) pairs in which both human and mouse orthologs harbor at least one 3' UTR SINE; (2) pairs in which only the human ortholog harbors at least one 3' UTR SINE; (3) pairs in which only the mouse ortholog harbors at least one 3' UTR SINE; and (4) pairs in which neither human or mouse ortholog has a 3' UTR SINE. All computational and statistical analyses of ortho-pair 3' UTRs were performed using the transcript that was annotated to harbor the longest 3' UTR, except when a shorter (rather than the longest) transcript was annotated to harbor a 3' UTR SINE. For measurements of 3' UTR length and nucleotide content, the SINE sequence itself was removed. P-values were calculated using ANOVA followed by the Tukey-Kramer multi-comparison procedure.

Simulations of SINE insertion models were performed using MATLAB. In complementary analyses, RefSeq mRNA databases for human and mouse were acquired from NCBI (www.ncbi.nlm.nih.gov/Ftp). To extract SINE-containing 3' UTRs directly from Refseq, a perl program "extract_3utr.pl" was written to extract sequences downstream of the annotated coding region of each mRNA. RepeatMasker (http://www.repeatmasker.org/cgi-bin/WEBRepeatMasker) and the commandline variants "–human" and "–Alu" were used to uniquely identify *Alu* elements, whereas the commandline variants "–mouse" and "–nolow" were used to identify B1, B2, B4 and ID elements. B1, B2, B4 and ID insertions were uniquely extracted from the mouse output file using the program extract_B-ID.pl.

Human 3' UTR *Alu*s and mouse 3' UTR SINEs were divided into subtypes as defined by Repeat Masker (*Alu*J, *Alu*Y, *Alu*S for human, and B1, B2, B4, ID for mouse). Since a single transcript may contain elements from different families, we selected only 3' UTRs that harbor elements from a single family. The length and AA and TT dinucleotide enrichment analyses were performed as described above for all 3' UTR SINEs.

The cumulative frequency plot (**Fig. S6*A***) and box plot (**Fig. S8*A***) were made in Stata (StataCorp LLC).

Cell culture and transient transfections. Adult human skeletal muscle (hSkMc) myoblasts (Cell Applications) were propagated in SkMc-cell growth medium (PromoCell); mouse C2C12 and rat L6E9 skeletal-muscle myoblasts were propagated in DMEM (GIBCO-BRL) containing 15% fetal bovine serum (FBS); and human HeLa cells

(ATCC), and African Green monkey COS-7 cells were propagated in DMEM (GIBCO-BRL) containing 10% FBS. Where specified, cells were transiently transfected with siRNA using Lipofectamine RNAiMAX (Invitrogen) and/or plasmids using Lipofectamine 2000 (Invitrogen) for C2C12 cells or Lipofectamine LTX (Invitrogen) for HeLa cells according to manufacturer's instructions.

Cell lysis and western blotting. Cells were lysed in RIPA buffer [50mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS 50mM Tris-Cl (pH8.0), and complete protease inhibitors EDTA-free (Roche)], and boiled at 95ºC for 5 min prior to electrophoresis (**Fig. 2*G,F***). Blotting was performed as described (2) using the following antibodies: anti-hSTAU1 (this paper), anti-Calnexin (Enzo Life Sciences), anti-PDK1 (Cell Signalling Technologies), anti-PDHE1a p(Ser^{232}) (Millipore), anti-PDHE1a (Abcam), and anti-UPF1 (2).

RNA purification, and RT coupled to quantitative (q)PCR. RNA was purified from total-cell lysates using TRIzol (Invitrogen) as reported (3, 4, 5). RT-qPCR was performed as described (4) using the designated PCR primer pairs (**Tables S7-9**), the 7500 Fast or StepOne Plus Real-Time PCR System (Applied Biosystems), and the Fast SYBR Green Master Mix (Applied Biosystems). All RT reactions utilized random primers (Invitrogen).

RNA-seq, and mapping and analysis of RNA-seq data. DNA co-purifying with total-cell RNA was digested using Turbo DNase (Life Technologies), and RNA quality was assayed using the 2100 Bioanalyzer (Agilent). Poly$(A)^+$ RNA was selected using

oligo(dT) Dynabeads (Life Technologies). Strand-specific cDNA sequencing libraries were prepared as described in (6) and paired-end sequenced using a HiSeq2000 (Illumina).

All reads were mapped using Tophat2 (7), with Bowtie2 (8) as the underlying alignment tool. The input Illumina fastq files consisting of paired-end reads were trimmed for quality and equivalency to 95x58 base-pairs before mapping. Reads that mapped using Bowtie2 to RepeatMasker elements in mouse or human (as appropriate to the sample) were discarded. The remaining reads were mapped using Tophat2 assisted by gtf files from the UCSC KnownGenes track (9). For multi-mapped fragments, only the highest scoring mapping as determined using Bowtie2 was retained, and only mappings with both read ends aligned were retained. Potential PCR duplicates (mappings of more than one fragment having identical positions for both read ends) were not removed since the duplication rate for all libraries was less than 0.2. The final set of mapped paired-end reads for each sample was converted to position-by-position coverage of the relevant genome assembly using the bedtools 'genomeCoverageBed' function (10). To determine the count of fragments mapping to a gene, the position-by-position coverage was summed over the exonic positions of the gene. The resulting gene total-coverage was divided by a factor of 153, to account for the 98+58 base-pairs of coverage induced by each mapped paired-end fragment, and rounded to an integer. This was calculated for each gene in the UCSC Known Gene set. For input to DESeq (5), all genes with non-zero counts in any sample were considered. Two replicates of each sample were combined per the DESeq methodology. DESeq output and its analysis to identify genes that were

significantly upregulated in both mouse and human myoblasts after STAU1 or UPF1 knockdown are provided (**Dataset S1**).

siRNAs. siRNA sequences are provided (**Table S10**).

Immunoprecipitations. An RNA immunoprecipitation (RIP) (11) protocol was adapted for identification of STAU1—RNA complexes using the α-STAU1 antibody here developed. C2C12 cells ($2x10^7$ cells/150mm-dish) were washed and collected in 1xPBS, and crosslinking was performed by incubating for 10 mins in 1xPBS-0.2% formaldehyde while rocking at room temperature. The reaction was stopped by adding glycine to 0.25M. Cells were pelleted, washed 2x in 1xPBS, and lysed by sonication (Branson Sonifier 250) in Empigen 300 buffer [1% Empigen BB (Sigma), 20mM TrisHCl (pH 7.5), 300mM NaCl, complete mini EDTA-free protease inhibitors (Roche), RNaseOUT (Invitrogen)], and pre-cleared using Protein G-Agarose (Roche). Pre-cleared lysate (1mg of protein) was rotated with 10μg of α-STAU1 antibody (this paper) or mouse IgG (Sigma) for 2 hrs at 4°C and subsequently rotated with 50μL of Protein G-Agarose, that had been pre-blocked with 40μL of 50mg/mL yeast tRNA (Sigma) for 90mins at 4°C in 200μL Empigen 300 buffer, for 90 mins at 4°C. Beads were washed 9x with Empigen 300 buffer, and protein—RNA complexes were eluted by boiling in Laemmli buffer. Crosslinks were reversed by incubating for >1hr at 65°C, and RNA was extracted using 25:24:1 acidic phenol: chloroform: isoamyl alcohol.

mRNA half-life assays. C2C12 or hSkMc cells were treated with 5μg/mL of Actinomycin D (Gibco) 48 hrs after transfection with Control, hSTAU1 or mStau1A siRNA, collected at the indicated time points, and lysed immediately in TRIzol (Invitrogen).

Plasmid constructions. To construct *pCINeo-hPDK1*, a plasmid containing the h*PDK1* open reading frame (ORF) (MGC:24867;ORFeome product ID: ORS09662) was digested with *Afl*II and *Ava*II, and sticky ends were removed using Klenow (NEB). The resulting fragment was blunt-end ligated into Klenow-filled *EcoR*I- and *Xho*I-digested pSport6 to generate *pSport6-hPDK1,* which lacks the h*PDK1* 3' UTR. The missing 3' UTR was generated using cDNA of hMB RNA, primed with random hexamers, in two PCR reactions: one reaction employed the primers 5'-GGTGGATCCTGTCACCAGCCAGAATG-3' and 5'-TCCTCGAGTCTATAAATCTGACTTTAAATAGG-3', where underlined nucleotides specify *BamH*I and *Xho*I sites, and the other reaction employed 5'-AGACTCGAGGAGACAGAAACTGAATCC-3' and 5'-CCCAAGCTTGCACACAACAGAAGATCTTCATTTATGCAAC-3', where underlined nucleotides specify *Xho*I and *Hind*III sites. The resulting PCR products were digested with *BamH*I and *Xho*I or *Xho*I and *Hind*III, respectively, and inserted into *pSport6-hPDK1* to generate *pSport6-hPDK1+3′ UTR*. *pSport6-hPDK1+3′ UTR* was digested with *Kpn*I and *Mlu*I, sticky ends were filled using Klenow, and the resulting blunt-end fragments of the ORF and 3' UTR were together inserted into the *Sma*I site in *pCINeo*.

To construct *pCINeo-hPDK1(ΔAlu),* the *Alu* element was first deleted from *pSport6-hPDK1+3′ UTR* using two sequential steps of PCR. The first step involved two

PCR reactions, both using *pSport6-hPDK1+3′ UTR*. The first reaction employed the primers 5'-GTCAGATTTATAGACTCGAGGAGACAGAAACTGAATC-3' and 5'-CAAGACTTGTGCAATAGACACGTGCAATGATG-3', and the other employed the primers 5'-CATCATTGCACGTGTCTATTGCACAAGTCTTG-3' and 5'-GATTAGCTAAGCAGCTCTGGGCAAAGGATG-3'. In step two, a subsequent round of PCR using the primers 5'-TCAGATTTATAGACTCGAGGAGACAGAAACTGAATC-3' and 5'-GATTAGCTAAGCAGCTCTGGGCAAAGGATG-3' deleted the *Alu*. The resulting PCR product was digested with *Xho*I and *Blp*I and inserted into the corresponding sites in *pSport6-hPDK1+3′ UTR* to generate *pSport6-hPDK1+3′ UTR(ΔAlu)*. *pSport6-hPDK1+3′ UTR(ΔAlu)* was digested with *Kpn*I and *Mlu*I, the sticky ends were made blunt using Klenow, and the ORF and 3' UTR were together inserted into the SmaI site in *pCINeo*.

To construct *pCINeo-mPdk1*, *pSport6-mPdk1* was purchased from Dharmacon (clone ID: 6315111). The ORF and 3' UTR were together inserted into *pCINeo* using *Sal*I and *Not*I restriction sites.

To construct *pCINeo-mPdk1(ΔB1),* the B1 element was first deleted from *pSport6-mPdk1* using two sequential steps of PCR. The first step involved two PCR reactions, both using *pSport6-mPdk1*. The first reaction employed the primers 5'-CAAATATTTCACGTGTGATCTTCACAATTAGTGTTTGC-3' and 5'-CTAACGTTTCTCTTGTAAATCCCTTGCGCTGAC-3', and the other employed the primers 5'-GCGAAGAACAACTACAAGAGAAACGTTAGTC-3' and 5'-CCTTTGGTCACCTGACCTCTCGCTG-3'. In step two, a subsequent round of PCR using the primers 5'-CAAATATTTCACGTGTGATCTTCACAATTAGTGTTTGC-3'

and 5'-CCTTTGGTCACCTGACCTCTCGCTG-3' resulted in deletion of the B1. The resulting PCR product was digested with *Pml*I and *BstE*II and inserted into the corresponding sites in *pSport6-mPdk1* to generate *pSport6-mPdk1(ΔB1)*. *pSport6-mPdk1(ΔB1)* was digested with SalI and NotI, and the ORF and 3' UTR were together inserted into the corresponding sites in *pCINeo*.

Plasmid constructions to generate and screen for a monoclonal anti-STAU1 antibody. Two copies of a DNA fragment, hSTAU1(426-496), which encodes 73 amino acids of the hSTAU1 C-terminus (amino acids 424-496), were tandemly inserted into pGEX6P-1 and pET28a to generate, respectively, pGEX6P-1-hSTAU1(424-496)×2stop, for expressing GST-hSTAU1(424-496)×2, and pET28a-hSTAU1(424-496)×2stop, for expressing 6×His-hSTAU1(424-496)×2.

To construct pGEX6P-1-hSTAU1(424-496)×2stop, hSTAU1(426-496) was PCR-amplified using pRSETB-hStau1[55] (ref (3)) as a DNA template. The PCR product that was generated using the primer-pair hSTAU1(426-496)-sense-*BamH*I (5'-CCGGATCCGAATACAAAGACTTCCCCAAAAAC-3') and hSTAU1(426-496)-antisense- *EcoR*I (5'-CCGAATTCGCACCTCCCACACACAGACATTGG-3') was digested with *BamH*I and *EcoR*I and ligated into the *BamH*I and *EcoR*I sites of pGEX6p-1 at, yielding pGEX6P-1-hSTAU1(424-496). The PCR product that was generated using the primer-pair hSTAU1(426-496)-sense-*Sal*I (5'-CCGTCGACGAGAATACAAAGACTTCCCCAAAAAC-3') and hSTAU1(426-496)-antisense-*Xho*I (5'-CCCTCGAGTCAGCACCTCCCACACACAGACAT-3') was

digested with *Sal*I and *Xho*I and ligated into the *Sal*I and *Xho*I sites of pGEX6P-1-hSTAU1(424-496), yielding pGEX6P-1-hSTAU1(424-496)×2stop.

To construct pET28a-hSTAU1(424-496)×2stop, the hSTAU1(424-496)×2 fragment was excised from pGEX6P-1-hSTAU1(424-496)×2stop using *BamH*I and *Xho*I and ligated into the *BamH*I and *Xho*I sites of pET28a.

Generating monoclonal anti-hSTAU1 antibody. Recombinant proteins were purified from *E. coli* (BL21DE3). Mice were immunized with GST-hSTAU1(424-496)×2, and 6×His-hSTAU1(424-496)×2 was used for screening anti-hSTAU1 monoclonal antibodies.

SI References

1. Mouse Genome Database (MGD), Mouse Genome Informatics, The Jackson Laboratory, Bar Harbor, Maine. World Wide Web (URL: http://www.informatics.jax.org).

2. Gong C, Maquat LE (2011) lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3′ UTRs via Alu elements. *Nature* 470(7333):284–288.

3. Kim YK, Furic L, DesGroseillers L, Maquat LE (2005) Mammalian Staufen1 recruits Upf1 to specific mRNA 3′UTRs so as to elicit mRNA decay. *Cell* 120(2):195–208.

4. Wang J, Gong C, Maquat LE (2013) Control of myogenesis by rodent SINE-containing lncRNAs. *Genes Dev* 27(7):793–804.

5. Anders S, Huber W (2010) Differential expression analysis for sequence count data. *Genome Biol* 11(10):R106.

6. Parkhomchuk D, et al. (2009) Transcriptome analysis by strand-specific sequencing of complementary DNA. *Nucleic Acids Research* 37(18):e123.

7. Kim D, Salzberg SL (2011) TopHat-Fusion: an algorithm for discovery of novel fusion transcripts. *Genome Biol* 12(8):R72.

8. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9(4):357–359.

9. Hsu F, et al. (2006) The UCSC Known Genes. *Bioinformatics* 22(9):1036–1046.

10. Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26(6):841–842.

11. Niranjanakumari S (2002) Reversible cross-linking combined with immunoprecipitation to study RNA–protein interactions in vivo. *Methods* 26(2):182–190.

12. Kusek, G. et al. (2012) Asymmetric segregation of the double-stranded RNA binding protein Staufen2 during mammalian neural stem cell divisions promotes lineage progression. *Cell Stem Cell* 11**,** 505–516.

Supplementary Table S1: Summary of data from two pipelines used to identify 3' UTR SINE-containing human and mouse orthologs

	UCSC RefSeq and MGI orthologs	**RefSeq and ENCODE 1:1 orthologs**
# of human genes with a 3' UTR *Alu* and a mouse ortholog/ total # of identifiable orthologs	2,833/16,064 (17.6%)	2,620/15,736 (16.6%)
# of mouse genes with a 3' UTR B/ID element and a human ortholog/ total # of identifiable orthologs	3,538/16,064 (22.0%)	3,536/15,736 (22.5%)
Observed # of orthologs with a 3' UTR SINE in both human and mouse orthologs/ total # of identifiable orthologs	1,192/16,064 (7.4%)	1,109/15,736 (7.0%)
Expected # of orthologs with a 3' UTR SINE in both human and mouse orthologs/ total # of identifiable orthologs	624/16,064 (3.9%)	589/15,736 (3.7%)

Number and percentage of orthologous protein-coding genes that produce ≥ one 3' UTR *Alu*- or B/ID-containing transcript as determined using two distinct workflows: RefSeq downloaded from UCSC for transcript sequences and Mouse Genome Informatics (MGI) for human:mouse orthologs, or RefSeq for transcript sequences and ENCODE for 1:1 human:mouse orthologs.

Supplementary Table S2: Name and identity of 3' UTR SINEs in 21 orthologous gene pairs that produce putative SMD targets in hMBs and mMBs

Gene Symbol		Encoded protein	3' UTR SINE(s)	
Human	Mouse		Human	Mouse
**C12orf49*	*2410131K14Rik*	UPF0454 protein C12orf49	*Alu*Sx (+); *Alu*Sx (-)	ID
KIAA0408	*9330159F19Rik*	Uncharacterized protein KIAA0408	*AluY* (+)	B1
**AAK1*	*Aak1*	AP2 associated kinase 1	*Alu*Jb (+); *Alu*Sx3 (+); *Alu*Sg (-); *Alu*Sq2 (+); *Alu*Jb (+)	6x B1; B4
ANKRD16	*Ankrd16*	Ankyrin-domain containing protein 16	AluY (+)	ID
CPM	*Cpm*	Carboxypeptidase M	*Alu*Sx1 (+); *Alu*Sz (+)	B2; 4x B1
DCUN1D2	*Dcun1d2*	Defective in Cullin Neddylation 2	*Alu*Sx (+)	B4
GPR155	*Gpr155*	G-protein coupled receptor 155	*Alu*Sx1 (+); *Alu*Sz (+); *Alu*Sz (+); *Alu*Y (+)	B4
GXYLT1	*Gxylt1*	Glucoside xylosyltransferase 1	*Alu*Jo (+)	B2; B1
GK	*Gyk*	Glycerol kinase	*Alu*Sp (+)	B1; B1
HIP1	*Hip1*	Huntington interacting protein 1	*Alu*Sq2 (+); *Alu*Jo (+); *Alu*Sx (+)	3x B1; B2
ING5	*Ing5*	Inhibitor of growth 5	*AluY* (-); *Alu*Sz (-); *Alu*Sx1 (-)	2x B4; 2x B1
LEPREL4	*Leprel4*	Leprecan-like protein 4	*Alu*Sx1 (+); *Alu*Sz (+);	2x B1
NT5DC3	*Nt5dc3*	5' nucleotidase domain containing 3	*Alu*Sp (-); *Alu*Sz6 (-); *Alu*Sq2 (-)	B1; 2x B1
PDK1	*Pdk1*	Pyruvate dehydrogenase kinase 1	*Alu*Jr (-)	B1
PHLPP2	*Phlpp2*	PH domain leucine-rich repeat containing protein phosphatase 2	*Alu*Sx1 (-)	B1
SEC63	*Sec63*	Translocon-associated secretory protein 63	*Alu*Sz6 (-)	B1
SLC7A2	*Slc7a2*	Low affinity cationic amino acid transporter 2	*Alu*Jo (+)	3x B1
SMG1	*Smg1*	SMG1 phosphatidylinositol 3-kinase-related kinase	*AluY* (+)	ID
SNX27	*Snx27*	Sorting nexin 27	*Alu*Jr4 (-); *Alu*Sz (-)	B4
**USP15*	*Usp15*	Ubiquitin carboxyterminal hydrolase	*Alu*Jb (+); *Alu*Jr4 (-)	B1; 3x B4; B2; ID
ZNRF3	*Znrf3*	Zinc and ring finger 3	*Alu*Sp (-)	B1

Human and mouse orthologous gene pairs encoding mRNAs with ≥ one 3' UTR SINE that are putative SMD targets in hMBs and mMBs. SINE subtypes are specified; PB1D refers to a partial B1 element. (+) and (-) indicate, respectively, a sense- or antisense-oriented 3' UTR *Alu*; * indicates a possible inverted-repeat *Alu*. Where more than one of a particular type of SINE is present in the 3' UTR of the mouse ortholog, the number of insertions is indicated (e.g. 2x signifies two copies). Colors correspond to results from **Fig. S2**: yellow indicates that each ortholog encodes an SMD target in hMBs or mMBs; red indicates that the human ortholog encodes an SMD target in hMBs; blue indicates that the mouse ortholog encodes an SMD target in mMBs; gray indicates that neither ortholog encodes an SMD target in hMBs or mMBs; white indicates that orthologs were not interrogated by RT-qPCR.

Supplementary Table S3: Statistical models using 3' UTR length to predict the presence of SINEs in the 3' UTRs of orthologous human and mouse genes, and also the observed numbers of orthologs with 3' UTR SINEs

		Expected # of human orthologs with ≥ one 3' UTR SINE	**Expected # of mouse orthologs with ≥ one 3' UTR SINE**	**Expected # of orthologs with ≥ one 3' UTR SINE in both human and mouse**
Statistical Models	Poisson (fixed probability of insertion per 3' UTR length in nts)	3642±53	5023±72	1606±39
	Poisson + local SINE duplications in a 3' UTR	2812±54	3649±44	959±22
	Fixed number of SINE-containing genes, randomly selected with probability proportional to 3' UTR length	2833	3538	932±16
Observed Results	Observed # of orthologous genes with ≥ one 3' UTR SINE in one or both orthologs	2833	3538	1192

Simulations of three different models for SINE insertion were performed: (1) A Poisson process, where a SINE has a fixed insertion probability per unit 3' UTR length; (2) A combined method whereby a Poisson process, as in (1), is followed by local SINE duplications within the same 3' UTR, with a probability that is independent of 3' UTR length; (3) SINEs are randomly allocated to 3' UTRs by fixing the number of SINE-harboring genes, then selecting them randomly with a probability that is proportional to 3' UTR length.

Supplementary Table S4: Results of a statistical analysis comparing 3' UTR lengths of SINE-containing and SINE-lacking 3' UTRs (data shown in Fig. 3*A*.)

First_Group	Second_Group	p_value_Human	p_value_Mouse
1	2	0.001541719	3.76826E-09
1	3	3.76826E-09	5.52219E-05
1	4	3.76826E-09	3.76826E-09
2	3	4.04392E-09	3.76826E-09
2	4	3.76826E-09	0.130227033
3	4	0.17530978	3.76826E-09

Table showing the calculated P-values from the comparison of 3' UTR lengths shown in Fig. 3*A* using ANOVA followed by the Tukey-Kramer multi-comparison procedure. Group 1 represents those ortho-pairs in which both orthologs are 3' UTR SINE-containing; group 2 represents those ortho-pairs in which only the human ortholog is 3' UTR SINE-containing; group 3 represents those ortho-pairs in which only the mouse ortholog is 3' UTR SINE-containing; group 4 represents those ortho-pairs in which neither ortholog is 3' UTR SINE-containing.

Supplementary Table S5: Results of a statistical analysis comparing 3' UTR lengths of SINE-containing and SINE-lacking 3' UTRs of specific subtypes

	SINE type	Group 1 vs Group 4 (p-value)	Number of Sequences (Group 1)	Number of Sequences (Group 4)
Human SINEs	*Alu*J	0.010922168	183	0
	*Alu*Y	0.109505097	61	0
	*Alu*S	2.32511E-08	517	0
	All	3.76826E-09	1192	10885
Mouse SINEs	B1	3.76844E-09	387	0
	B2	0.040992567	107	0
	B4	0.411981347	96	0
	ID	0.001673608	59	0
	All	3.76826E-09	1192	10885

Table showing the calculated P-values from the comparison of 3' UTR lengths between 3' UTRs with the indicated Alu subtype or SINE type and SINE-lacking 3' UTRs using ANOVA followed by the Tukey-Kramer multi-comparison procedure. Group 1 represents those ortho-pairs in which both orthologs are 3' UTR SINE-containing; group 4 represents those ortho-pairs in which neither ortholog is 3' UTR SINE-containing.

Supplementary Table S6: Results of a statistical analysis comparing 3' UTR AA- or TT-dinucleotide content of SINE-containing and SINE-lacking 3' UTRs of specific subtypes

	SINE type	AA content Group 1 vs Group 4 (p-value)	TT content Group 1 vs Group 4 (p-value)	Number of Sequences (Group 1)	Number of Sequences (Group 4)
	*Alu*J	0.03030948	0.002187358	183	0
	*Alu*Y	0.021807946	0.17233966	61	0
	*Alu*S	8.75095E-07	4.56095E-08	517	0
Human SINEs	All	3.80057E-09	3.76826E-09	1192	10885
	B1	0.468494035	1.79943E-06	387	0
	B2	0.494500956	0.966767349	107	0
	B4	0.927507252	0.837788586	96	0
	ID	0.971399508	0.997289793	59	0
Mouse SINEs	All	0.461629196	2.44191E-07	1192	10885

Table showing the calculated P-values for the comparison of 3' UTR AA or TT dinucleotide content between SINE-containing 3' UTRs, where the *Alu* subtype or SINE type has been indicated, and SINE-lacking 3' UTRs. Group 1 represents those ortho-pairs in which both orthologs are 3' UTR SINE-containing; group 4 represents those ortho-pairs in which neither ortholog is 3' UTR SINE-containing.

Supplementary Table S7: Human quantitative (q)PCR primer-pairs used in Fig. S2

Target	Primer-pairs	
h*GPR155* mRNA	GTGGAGACCAGCAACTGACC	CTTCCAGGCTCTTGGTTGAA
h*GPR155* pre-mRNA	TTGATGGCCATACTGAATCTTTC	TCATTTTCTCTCACCCTCCAAC
h*PDK1* mRNA	CAACTGCACCAAGACCTCGT	ACGTGATATGGGCAATCCAT
h*PDK1* pre-mRNA	ACCATTGCCATCTTAATGCG	ACGCCTAGCATTTTCATAGCC
h*SNX27* mRNA	TGGTGAGAGTGACATCATGC	GTCCATGCCAACCTTTGCTG
h*SNX27* pre-mRNA	TTTCACTCCAGCTTTCTGTGTTT	GCCTGGCACAGCTGATGTAT
h*PHLPP2* mRNA	TGGAACACAAGACACTGGACAT	CATGGCTCCAGAAGGTTGAC
h*PHLPP2* pre-mRNA	GCTGCTCAGTCTTGGTGCAT	GGTTCAGTTCCTCCAATTGCT
h*DCUN1D2* mRNA	CCTTCACCTTCGCTAAGAAC	CTTCTTCATCGTAGTTAGAC
h*DCUN1D2* pre-mRNA	ACAAGGGGAGTCGACTAGGG	TTCAGCTCCTGCTCCAGTCT
h*ING5* mRNA	GCCATGTACTTGGAGCACTA	ATCTCTGCTTTCTTATCTTC
h*ING5* pre-mRNA	CACTGCGCCTTTCTTGTCA	CGACACGAATGAAGGGACA
h*SEC63* mRNA	CTGCAGCTCCCTCATATTGAA	AGTTGCAGTAGAGTGTGACGATCTG
h*SEC63* pre-mRNA	CTTTGTGTTTGCCTTTCCCTT	TTCTTCTTCAGAATCACTGCCCC
h*AAK1* mRNA	TCCTATTATCCACCGGGACC	TTCTGGAATTTGTTGGTGGC
h*AAK1* pre-mRNA	TGGTTGGGGAATGTTTTTCA	TTCTGGAATTTGTTGGTGGC
h*CPM* mRNA	ATCCCAACATGAAGAAAGGAGAC	TGGGCCCAGATGTAGTTGTAA
h*CPM* pre-mRNA	GTCTGATTCATTTGTCCCCG	TGGGCCCAGATGTAGTTGTAA
h*SLC7A2* mRNA	CAATTCCAAAACGAAGACACC	AGGCTGCCACCAGAGAGTAG
h*SLC7A2* pre-mRNA	TGCCTCAAAATGCTATTGCC	ACCCAAGCAGACTCTTTTACTCC
h*GXYLT1* mRNA	TTGATGAACATGACTCGAATGAG	TCTCCCCATTGTAGTCGTACAGTT
h*GXYLT1* pre-mRNA	TGAGCATGATGAAAGTACACATTG	GCTAGATGCATTTTCTCAACAGG
h*ZNRF3* mRNA	TTGACATGGGGATTTTCCTG	AGCCAGCCTGTTCATGGAAT
h*ZNRF3* pre-mRNA	GAATTGTCTGCCTGTCCCAG	AGCCAGCCTGTTCATGGAAT
h*SMG1* mRNA	GCATGGCAGTCGTGCTTTAG	CGACAGTCGAGACTCATCAGAA
h*SMG1* pre-mRNA	CACCAAGATTCTGCTGGCTT	TGTAGTACAGGGGCCAATGC
h*NT5DC3* mRNA	ATTCAAAGCACCTCCACACG	CGAATTGCAAAATTTGGGTC
h*NT5DC3* pre-mRNA	CTGCACTTGGCACATTGTTG	CGAATTGCAAAATTTGGGTC
h*GK* mRNA	GTAGTCTGCCCTTGGGCTTT	GCTCACATCTTGGGAATCCA
h*GK* pre-mRNA	GATTGAATCATCTGCATTGCTTT	GGGAATCCATGAGTTGGTAGG
h*ANKRD16* mRNA	ACCAACCGAGACTACAAGCG	TTCTTCAGGAGTGGATTGGC
h*ANKRD16* pre-mRNA	TCCCCTGGGTTTGAGGTATT	TGACGCCACAGTTGTCTCTG
h*HIP1* mRNA	CCTCTCGGGGAGTGAACC	TCAGCGTCATGCTTGAGAAG
h*HIP1* pre-mRNA	TTTCCACAGAGCATTGGCAT	TCAGCGTCATGCTTGAGAAG
h*PDK1* (Alu-containing 3' UTR isoform only)	TGTGCCACCTGAAAGGTTTT	CGTGTCCACTTCCTGTCCTT

Supplementary Table S8: Mouse quantitative (q)PCR primer-pairs used in Fig. S2

Target	Primer-pairs	
m*Gpr155* mRNA	CAGAGTAATGACTTTGCTTTGGG	GGTGCCACCAAATAAATGTACTG
m*Gpr155* pre-mRNA	TCCAGCCTCTCCACTCAAAG	GGTGACTGCCGTGGTGAT
m*Pdk1* mRNA	AGCCTTCAGGAGTTGCTTGAT	TGGGAATGACATCATTGTGC
m*Pdk1* pre-mRNA	TTTGTGCCAATTTTACAACATGA	TCAAGCAACTCCTGAAGGCT
m*Snx27* mRNA	TGGTGAGAGTGACATCATGC	GTCCATGCCAACCTTTGCTG
m*Snx27* pre-mRNA	GCTAGCATGTAGCAAGGAGAATG	GTTGTTCCATCCGGTAGTGC
m*Phlpp2* mRNA	AAAGTGGCTATGGCAGGAAAC	TCACCGTTTTCAAGTGGTTCA
m*Phlpp2* pre-mRNA	GTTCCTGTCTGTCTTCCCCA	CGCAGGTTGAGGTAGGTGATG
m*Dcun1d2* mRNA	CCTTCACCTTCGCTAAGAAC	CTTCTTCATCGTAGTTAGAC
m*Dcun1d2* pre-mRNA	GCCCTGCCCACTGTGATA	TCCAAGTATCCCTTGGAATTGA
m*Ing5* mRNA	GCCATGTACTTGGAGCACTA	ATCTCTGCTTTCTTATCTTC
m*Ing5* pre-mRNA	TGTACAGTGACATGGCCAGC	CAGCCAGGATGTCGATCTCT
m*Sec63* mRNA	TCATCCAGATAAAGGAGGTGATG	CTTGAGGCCCATCTGGATTT
m*Sec63* pre-mRNA	TTTTGGGGGTATGGCTTTCT	CTTGAGGCCCATCTGGATTT
m*Aak1* mRNA	GCCTGCATCAGTGCAAAACT	TTCTGGAATTTGTTGGTGGC
m*Aak1* pre-mRNA	GGTGCCAACAGTAGCCAGAG	CATTTCACCCCATTGCTTGT
m*Cpm* mRNA	GTCCAGAAGCCCGACTGTTA	TTCTCAAAGGCATCAGGAAAGT
m*Cpm* pre-mRNA	CACCCAAACACCCACACAAT	TTCTCAAAGGCATCAGGAAAGT
m*Slc7a2* mRNA	CAATTCCAAAACGAAGACACC	AGGCTGCCACCAGAGAGTAG
m*Slc7a2* pre-mRNA	TCCCTGAAGTGGCTGACAAT	CCGTAGATGCTTGTTCCGTT
m*Gxylt1* mRNA	GGCGTGAACTCTGGAGTGAT	CCGCGCAGTTGTCATATCAT
m*Gxylt1* pre-mRNA	GGGCTTTGCAGTGGTTCTAA	CCGCGCAGTTGTCATATCAT
m*Znrf3* mRNA	TCTTCGTCGTGGTCTCCTTG	GTTTCCATCTTCTCCAGAGCC
m*Znrf3* pre-mRNA	AACACTGACCATGCTCTGCC	GTTTCCATCTTCTCCAGAGCC
m*Smg1* mRNA	CAGTGTTCAGCATGGCAGTC	CGACAGTCGAGACTCATCAGAG
m*Smg1* pre-mRNA	CCCTGGTTTTGATGCACTTC	CGACAGTCGAGACTCATCAGAG
m*Nt5dc3* mRNA	CTGTGTGAACGAGCACTTCCT	GCTTCGATTGCTCGGTACAT
m*Nt5dc3* pre-mRNA	GGCACTAACTCCAGCTCCCT	GCTTCGATTGCTCGGTACAT
m*Gyk* mRNA	GAACATGGCCTCCTGACAAC	GTTGTCTCTTAGCCAGCGGA
m*Gyk* pre-mRNA	GAAAAGGAAATCCATGGCCTA	CTTCTAGTGCAGCAAAAGCGA
m*Ankrd16* mRNA	ACCAACCGAGACTACAAGCG	TTCTTCAGGAGTGGATTGGC
m*Ankrd16* pre-mRNA	CAATGAGCTGCTTCCTCTGC	CACCTGGACTGCTTCCAAAC
m*Hip1* mRNA	GATCAGTGGACTGACAGGGC	CTGCTCTGCTAGCTCTGCCT
m*Hip1* pre-mRNA	GCCTGGCAGTAGTGAGCTTC	CTGCTCTGCTAGCTCTGCCT

Supplementary Table S9: qPCR primer-pairs used in Fig. 2 and Fig. 4

Target	Primer-pairs	
pCINeo-h*PDK1* mRNA	GGGCTTGCACTAAATTTCACTATC	ACCCTCACTAAAGGGAAGCG
pCINeo-m*Pdk1* mRNA	GCAGGTGTCCACTCCCAG	CCAGCCTCATGGGTCGA
MUP mRNA	CTGATGGGGCTCTATG	TCCTGGTGAGAAGTCTCC
cons*SNX27* mRNA	TCATGTACCTCAACATGCTAAGGAC	AAGTGCGTGATGCTGATGG
cons*PDK1* mRNA	CTCCATGAAGCAGTTCCTGG	TGGCTGGTGACAGGATCC
Mouse and rat 18S rRNA	TCCAGCACATTTTGCGAGTA	CCACATGAGCATATCTCCGC
human and African green monkey 18S rRNA	GGGAAACCAAAGTCTTTGGG	GGAATTAACCAGACAAATCGC

Supplementary Table S10: siRNAs used in Figs 1, 2, 4, S2, S3 and S9

siRNA	Target	siRNA sequence	Reference
Control		Silencer negative control #1	Ambion
STAU1 #1	human STAU1	5'-UUGACUAACUCCUACAGCCdTdT-3'	3
UPF1 #1	human UPF1	5'-GAUGCAGUUCCGCUCCAUUdTdT-3'	3
STAU1 #2 (AKA: STAU1B)	human STAU1	5'-UUGACUAACUCCUACAGCCdTdT-3'	this paper
UPF1 #2 (AKA: UPF1A)	human UPF1	5'-AACGUUUGCCGUGGAUGAGdTdT-3'	3
STAU1 #3 (AKA: mStau1A)	mouse Stau1	5'-CAACUGUACUACCUUUCCAdTdT-3'	3
UPF1 #3 (AKA: mUpf1A)	mouse Upf1	5'-UCAAGGUUCCUGAUAAUUAdTdT-3'	3
STAU1 #4	conserved sequence of STAU1	5'- ACGGUAACUGCCAUGAUAGdTdT-3'	this paper
UPF1 #4	conserved sequence of UPF1	5'- GAAAUACUUCUGGCAGCCAdTdT-3'	this paper
mStau2	mouse Stau2	5'-ACCUCUGGCACAACUCUAAdTdT-3'	12

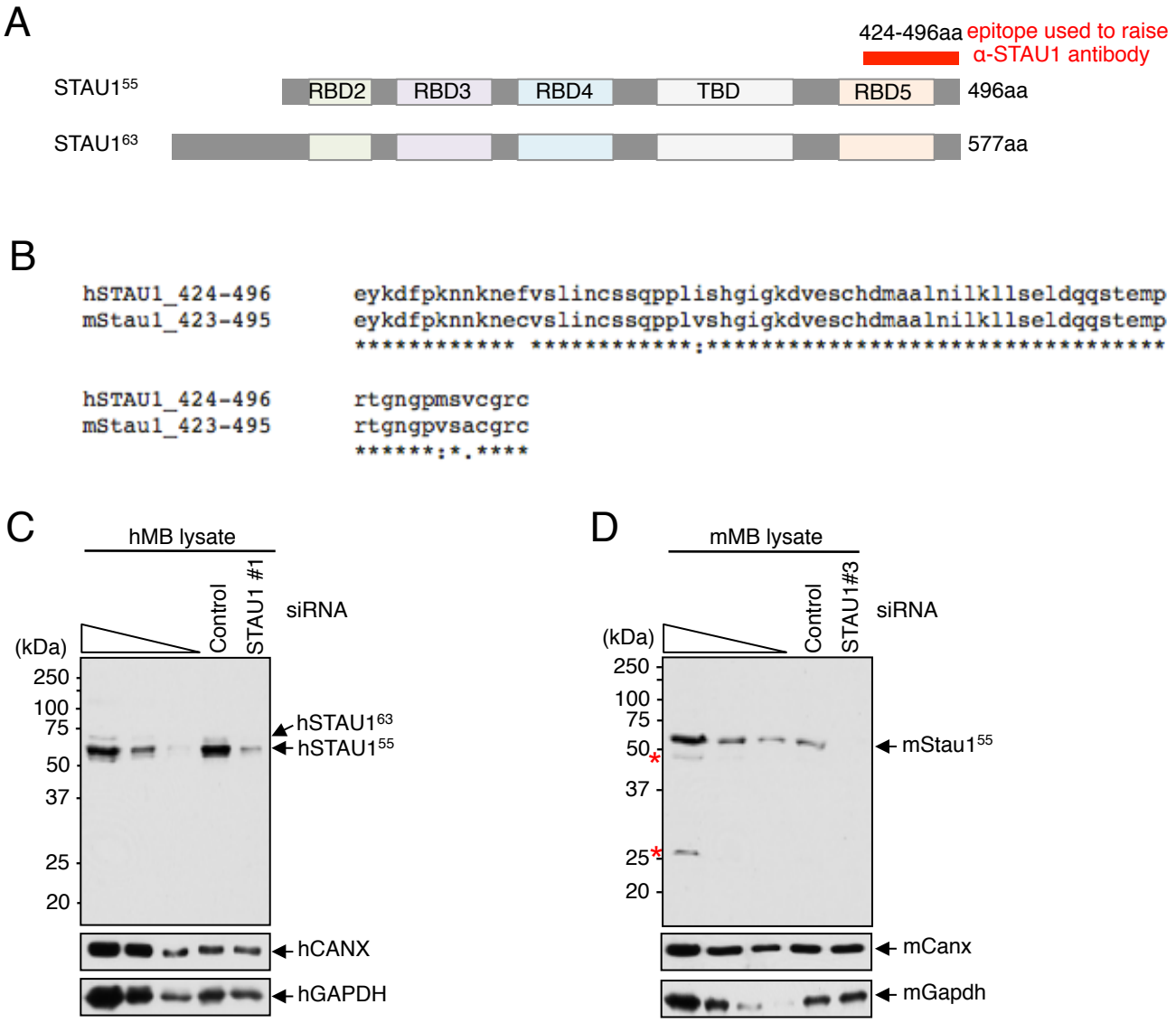


Fig. S1. Generating a monoclonal antibody that reacts specifically with hSTAU1 and mStau1 proteins. (***A***) Diagram showing two isoforms of hSTAU1, where the red horizontal bar indicates the amino acids in the epitope used to raise α-STAU1 antibody in mouse. (***B***) Clustal Omega alignment of the human epitope sequence used to raise the anti-STAU1 antibody aligned to the equivalent sequence in mStau1; asterisks denote exact amino acid matches; colons denote conservation amino acids having strongly similar properties; the period denotes conservation of amino acids having weakly similar properties. (***C***) Full-length western blot of lysates from human myoblasts (hMBs) transfected with Control siRNA or hSTAU1 siRNA showing that the antibody is specific for hSTAU1; hCalnexin (hCANX) and hGAPDH serve as loading controls, and the three leftmost lanes analyze 3-fold dilutions of lysates. Red asterisks indicate possible degradation products. (***D***) As for ***D*,** only using lysates of mMBs treated with control siRNA or mStau1 siRNA. Red asterisks mark mStau1 degradation products.

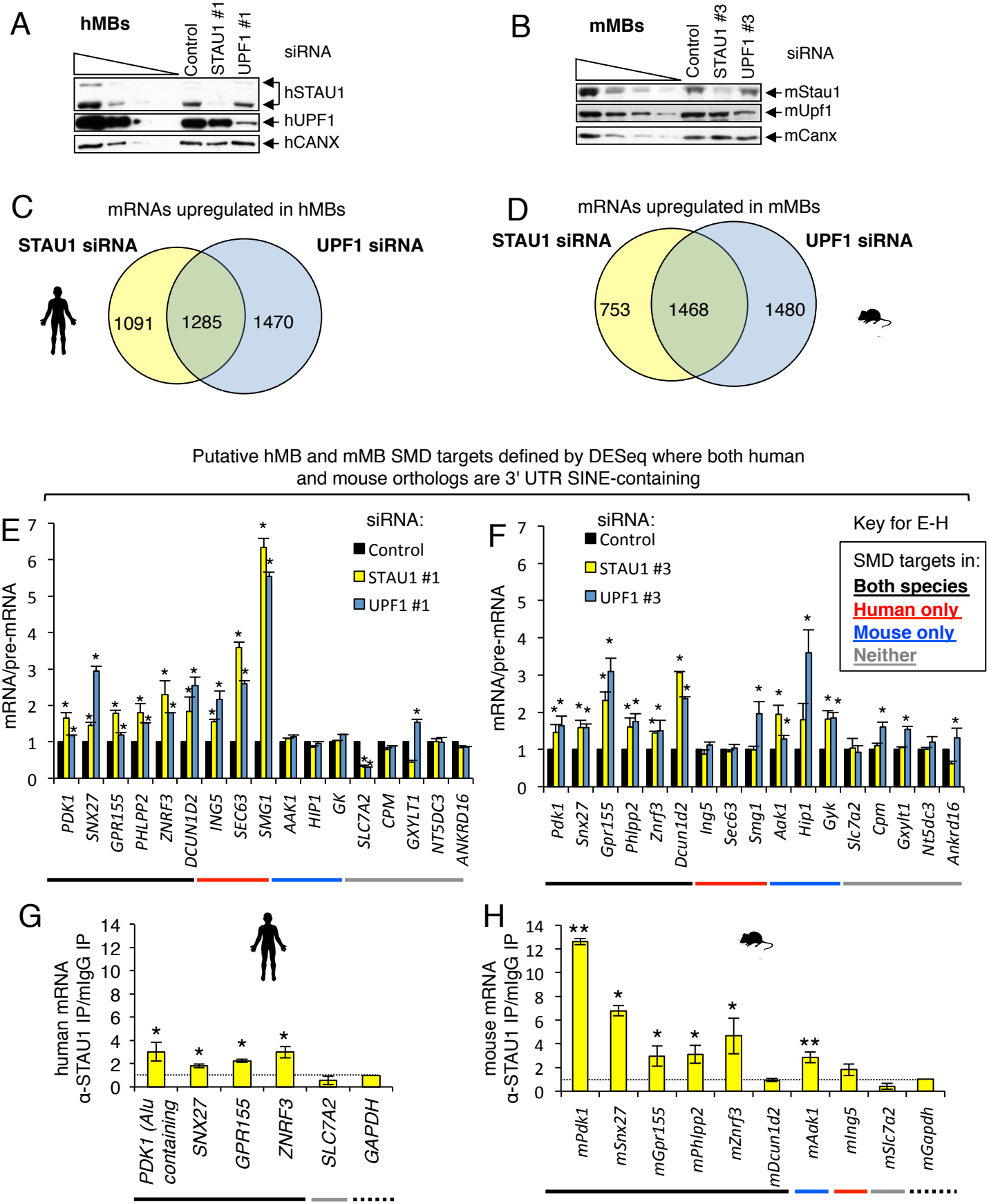

A
hMBs
Control
STAU1 #1
UPF1 #1
siRNA
hSTAU1
hUPF1
hCANX
B
mMBs
Control
STAU1 #3
UPF1 #3
siRNA
mStau1
mUpf1
mCanx
C
mRNAs upregulated in hMBs
STAU1 siRNA
UPF1 siRNA
1091
1285
1470
D
mRNAs upregulated in mMBs
STAU1 siRNA
UPF1 siRNA
753
1468
1480
Putative hMB and mMB SMD targets defined by DESeq where both human and mouse orthologs are 3' UTR SINE-containing
E
mRNA/pre-mRNA
siRNA:
Control
STAU1 #1
UPF1 #1
PDK1
SNX27
GPR155
PHLPP2
ZNRF3
DCUN1D2
ING5
SEC63
SMG1
AAK1
HIP1
GK
SLC7A2
CPM
GXYLT1
NT5DC3
ANKRD16
F
mRNA/pre-mRNA
siRNA:
Control
STAU1 #3
UPF1 #3
Key for E-H
SMD targets in:
Both species
Human only
Mouse only
Neither
Pdk1
Snx27
Gpr155
Phlpp2
Znrf3
Dcun1d2
Ing5
Sec63
Smg1
Aak1
Hip1
Gyk
Slc7a2
Cpm
Gxylt1
Nt5dc3
Ankrd16
G
human mRNA
α-STAU1 IP/mIgG IP
PDK1 (Alu containing
SNX27
GPR155
ZNRF3
SLC7A2
GAPDH
H
mouse mRNA
α-STAU1 IP/mIgG IP
mPdk1
mSnx27
mGpr155
mPhlpp2
mZnrf3
mDcun1d2
mAak1
mIng5
mSlc7a2
mGapdh

Fig. S2. Identification of SMD targets in hMBs and mMBs. (***A*)** Western blot of lysates of hMBs transfected with the specified siRNA, the level of Calnexin controls for loading and recovery (***B***) As in a but using mMB) (***C***) Venn diagram showing all human transcripts significantly upregulated with either STAU1 (yellow) or UPF1 (blue) knockdown in hMBs. Overlap shows those transcripts significantly upregulated in both conditions. (***D***) As for *C* only for mouse transcripts in mMBs. (***E***) Histogram of RT-qPCR results, whereby the expression of 17 hMB mRNAs (selected from the 24 in the overlapping region of Fig. 1e) were analyzed and normalized to their corresponding pre-mRNA. The ratio in the presence of Control siRNA is defined as 1. Color coding as described in Fig 1f. (***F***) As in *E* only analyzing the mouse orthologs using mMBs. (***G***) RT-qPCR quantitation of human mRNAs that co-immunoprecipitate (IP) with STAU1 from human HeLa cells. Results were first normalized to mouse IgG then to *GAPDH* mRNA. **(*H*)** As for g only analyzing the co-IP of mouse mRNAs with STAU1 from mMBs. Results shown relative to mIgG control. Dashed line indicates STAU1 IP negative control mRNAs.
All results derive from ≥ 3 independent experiments. Histograms represent the average ± SD. $*P < 0.05$, $**P < 0.01$.

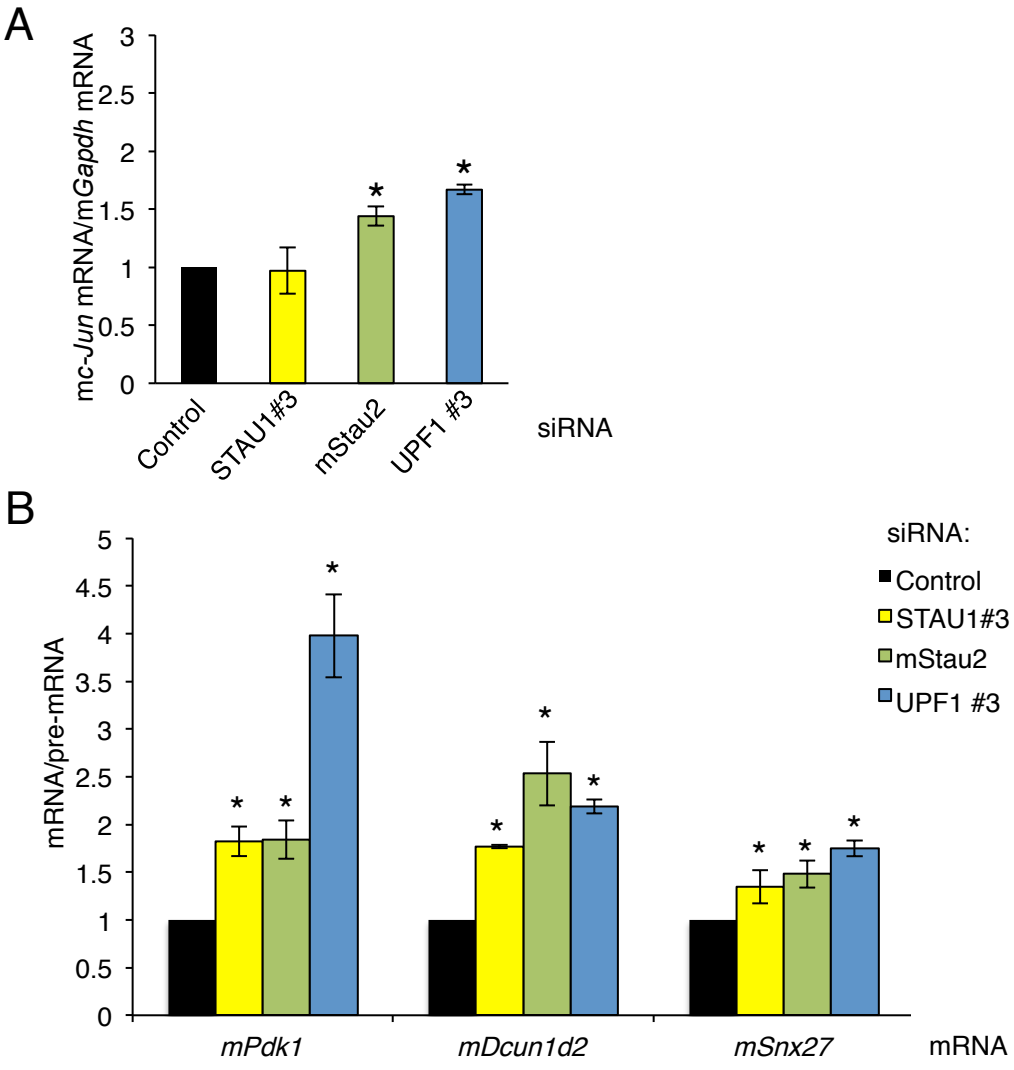


Fig. S3. Evidence that, like hSTAU2, mStau2 can trigger SMD. (***A***) Histogram representation of RT-qPCR results using primers to previously demonstrated SMD target m*c-Jun* mRNA, the level of which was normalized to the level of m*Gapdh* mRNA, show that m*c-Jun* mRNA is upregulated when mStau2 and, independently, mUpf1 are downregulated in mMBs. (***B***) Histogram representation of RT-qPCR results demonstrates that, like Stau1 siRNA and Upf1 siRNA, Stau2 siRNA upregulates the level of *bona fide* SMD targets m*Pdk1* mRNA, m*Dcun1d2* mRNA, m*Snx27* mRNA and m*Gpr155* mRNA. Each mRNA was normalized to the level of its pre-mRNA, and the normalized level in the presence of control siRNA is defined as 1.

All results derive from ≥ 3 independent experiments. Histograms represent the average ± SD. **P* < 0.05.

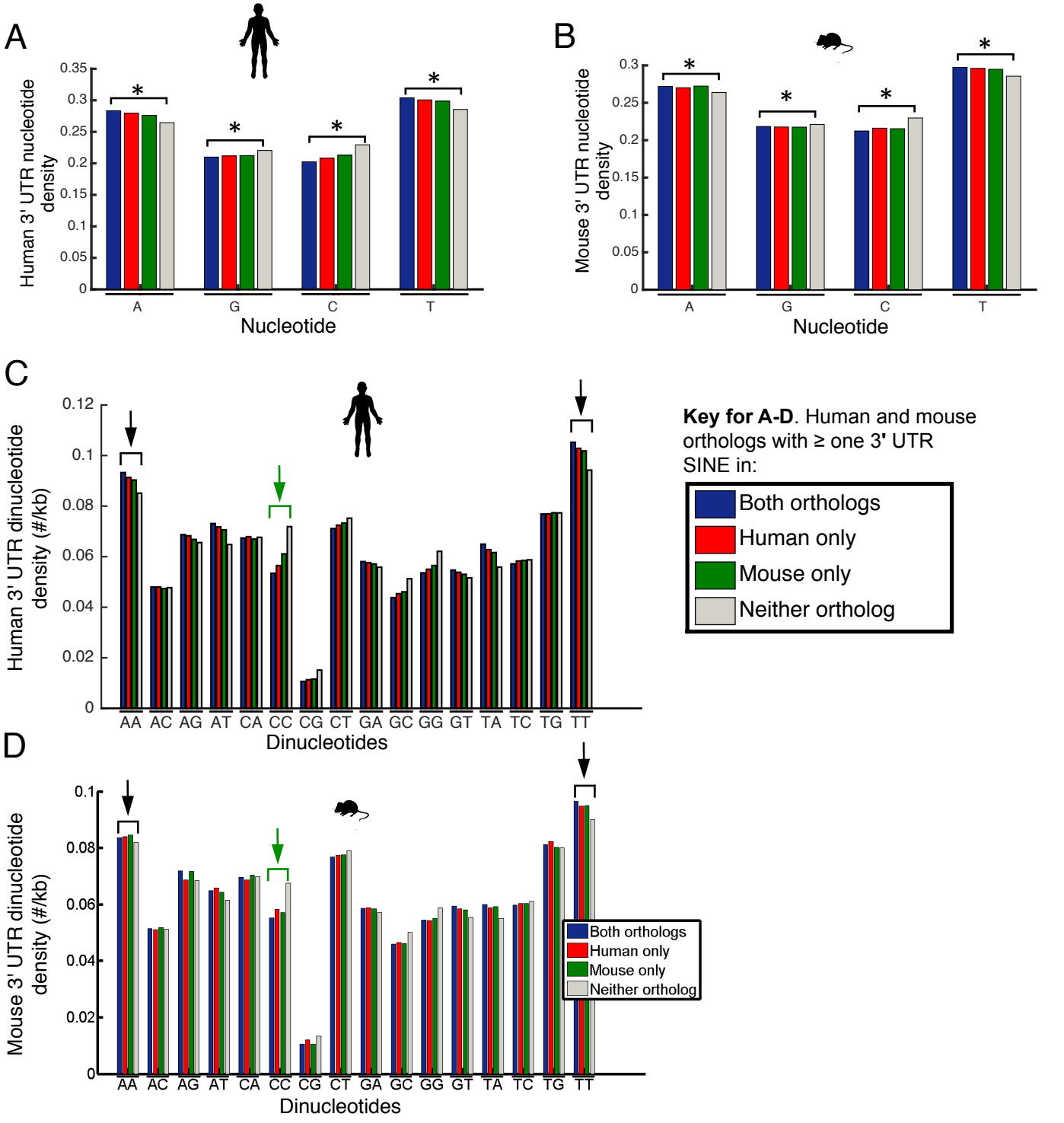


Fig. S4. *Alu* and B/ID elements tend to reside in A+T-rich 3' UTRs. (***A***) Histogram of the density of each nucleotide in the human mRNA isoform with the longest 3' UTR (excluding the *Alu* sequences themselves) for those mRNAs where it and its mouse ortholog, lack a SINE (gray), contain ≥ one 3' UTR SINE in the human ortholog (red), the mouse ortholog (green) or both (blue). (***B***) As for ***A*** but for the mouse isoform with the longest 3' UTR, where B/ID sequences were excluded from analyses. (***C***) Dinucleotide frequency of sequences described in ***A***; black arrows indicate enrichment of AA and TT dinucleotides; green arrow denotes depletion of CC dinucleotides. (***D***) As for ***C***, but using sequences from *B*. * $p<0.05$

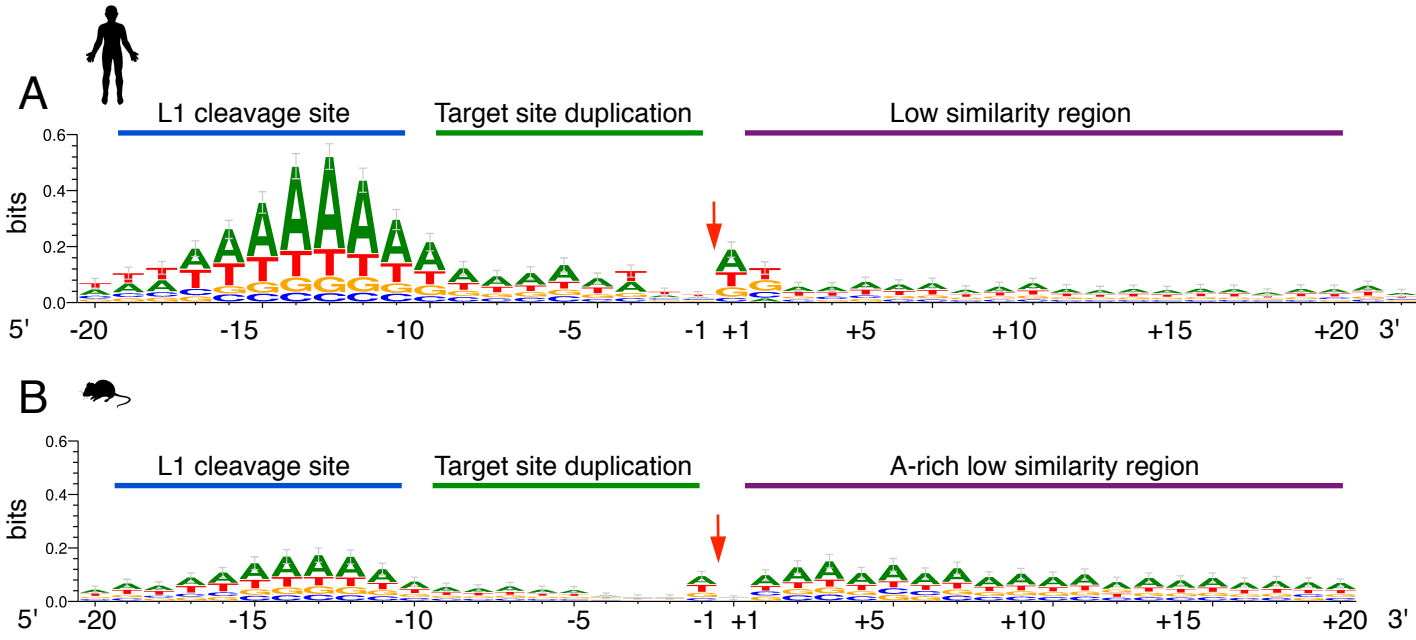
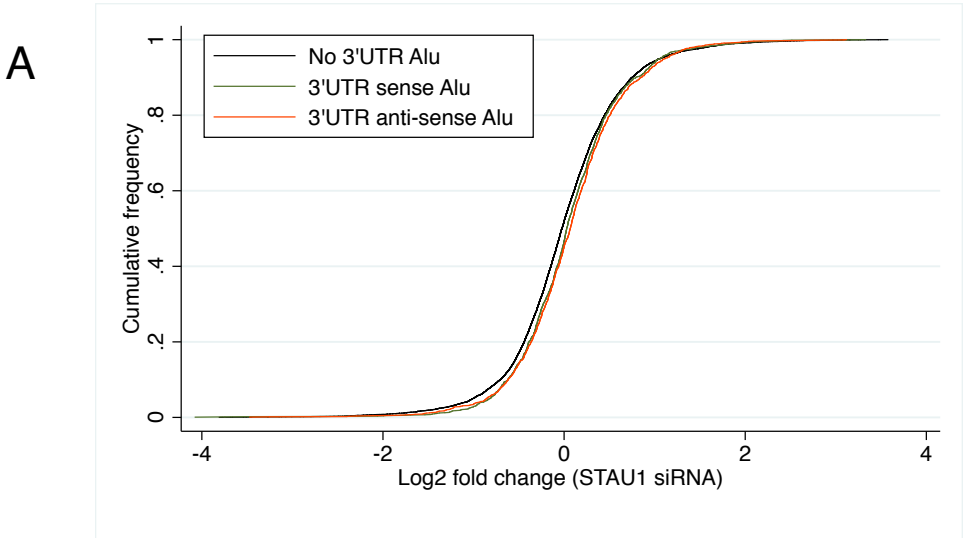
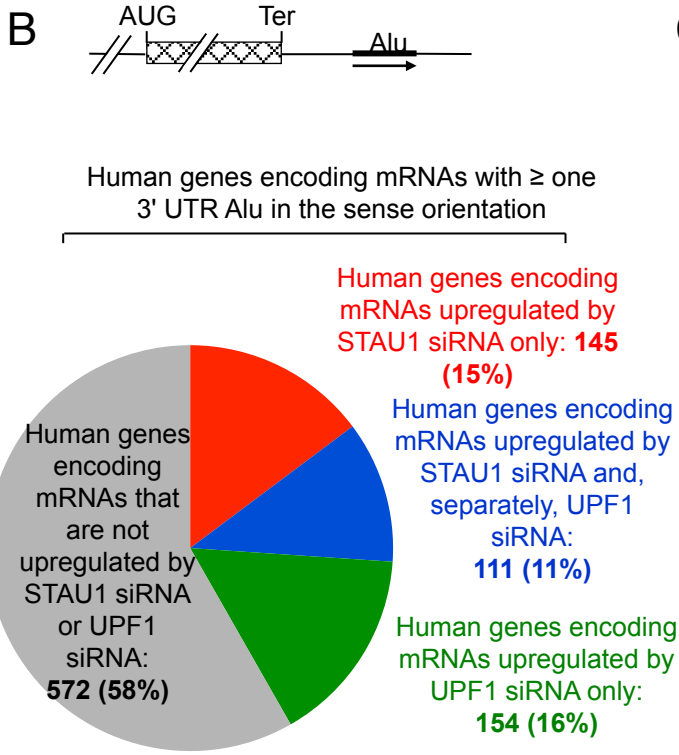


Fig. S5. Nucleotides flanking the SINE in human and mouse 3' UTRs display characteristic features of L1 insertion. (***A***) Alignment of the 20-nucleotides residing 5' (i.e. positions -20 to -1) or 3' (i.e. positions +1 to +20) of 3' UTR *Alu* insertion sites (red arrow) defined by Repeat Masker so as to generate a sequence LOGO using WebLogo to weigh (bits) nucleotide composition at each position. (***B***) As in a but for 3' UTR B/ID SINEs.





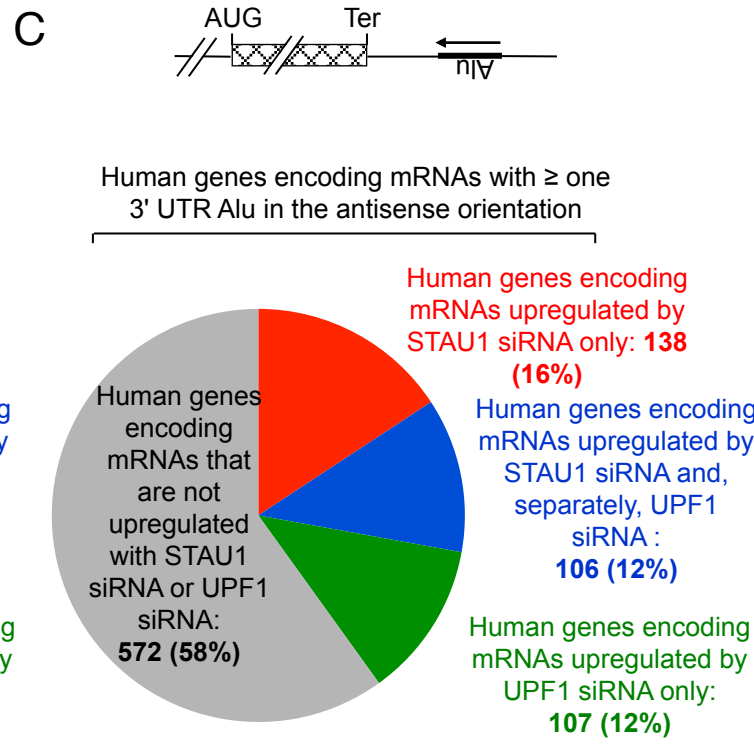


Fig. S6. No preference for *Alu* orientation in the 3' UTRs of putative SMD targets. (***A***) Cumulative frequency plot showing the distribution of log2 fold changes in mRNA abundance in response to STAU1 siRNA for mRNAs that lack a 3' UTR SINE (dark gray), with only sense *Alu*s (green) or only anti-sense *Alu*s (orange). (***B***) Diagram showing the orientation of a sense *Alu* relative to the gene (upper). Pie chart showing the proportion of sense *Alu*s in the 3' UTRs of mRNAs that are upregulated in response to STAU1 siRNA only (red), UPF1 siRNA only (green), both siRNAs (blue) or neither siRNA (gray). (***C***) As for *B*, only showing anti-sense *Alu*s.

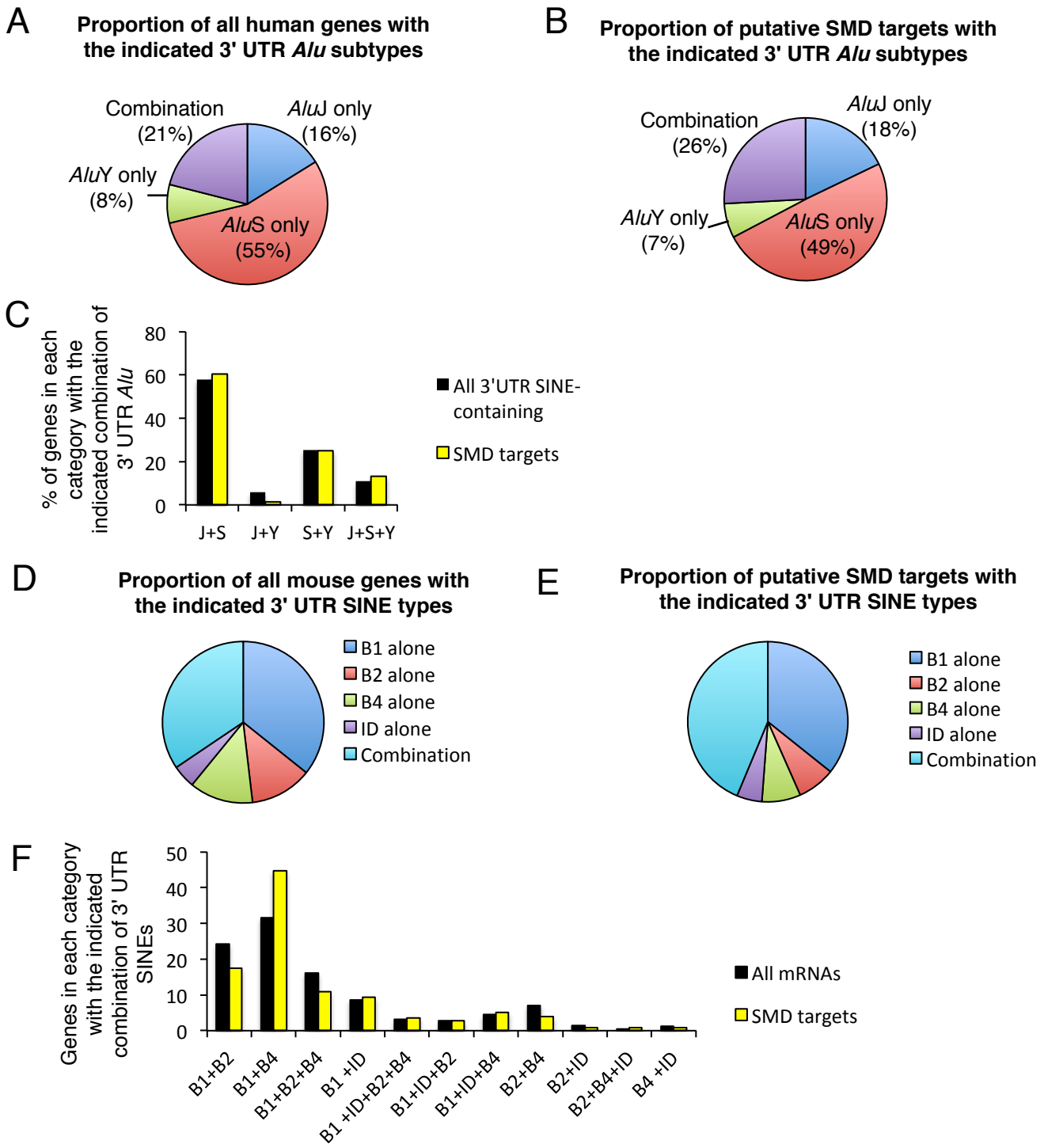


Fig. S7. No particular *Alu* or SINE subtype is overrepresented in the 3' UTR of putative SMD targets. Combinations of SINE types that have the potential to form a duplex are overrepresented in putative SMD targets. (***A***) Pie chart showing the proportion of the indicated *Alu* subtypes or a combination in all human mRNA 3' UTRs. (***B***) As for *A*, except only showing those *Alu*s in the 3' UTRs of mRNAs that are upregulated when STAU1, and, independently, UPF1 is downregulated (i.e. putative SMD targets). (***C***) Histogram showing the difference in the percentage of each indicated combination of *Alu*s in all 3' UTRs (black) or the 3' UTRs of putative SMD targets (yellow). (***D***) Pie chart showing the proportion of the indicated SINE types or a combination in all mouse mRNA 3' UTRs. (***E***) As for *D*, except only showing those SINEs in the 3' UTRs of mRNAs that are putative SMD targets. (***F***) As for *C*, only showing combinations of mouse SINEs in all mouse mRNA 3' UTRs (black) or putative SMD targets (yellow).

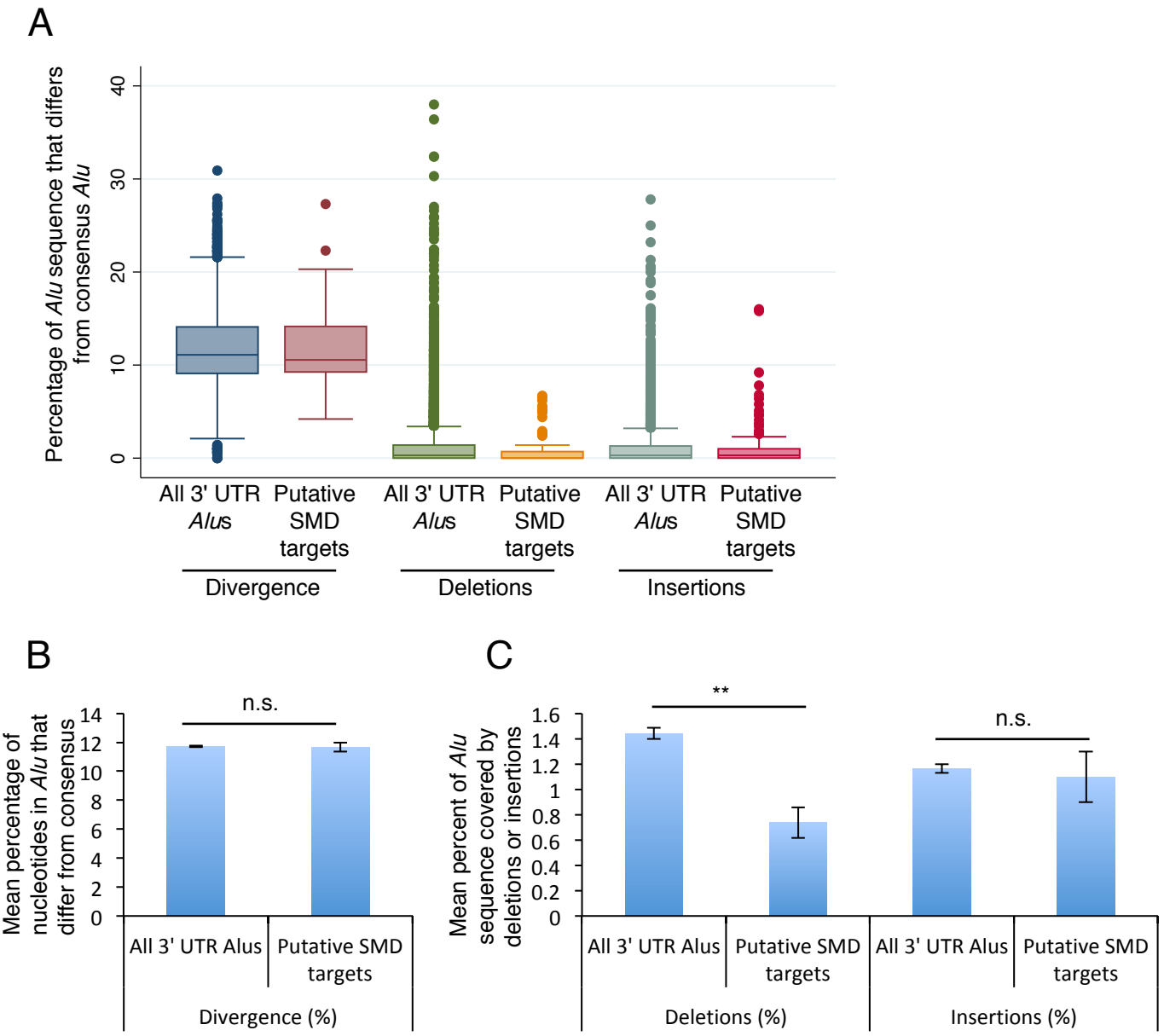


Fig. S8. The percent of the *Alu* sequence deleted relative to the consensus *Alu* sequence is less in 3' UTR *Alu*s residing in putative SMD targets relative to all 3' UTR *Alu*s. (***A***) Boxplots show the median and interquartile range of the percentage divergence (left), deletions (middle) and insertions (right) relative to the consensus *Alu* sequence as determined by Repeat Masker. (***B***) Histogram showing the mean ± std error of the percentage divergence relative to the consensus *Alu*. (***C***) As for *B* but showing the percentage of the *Alu* sequence with deletions (left) or insertions (right) relative to the consensus *Alu*.

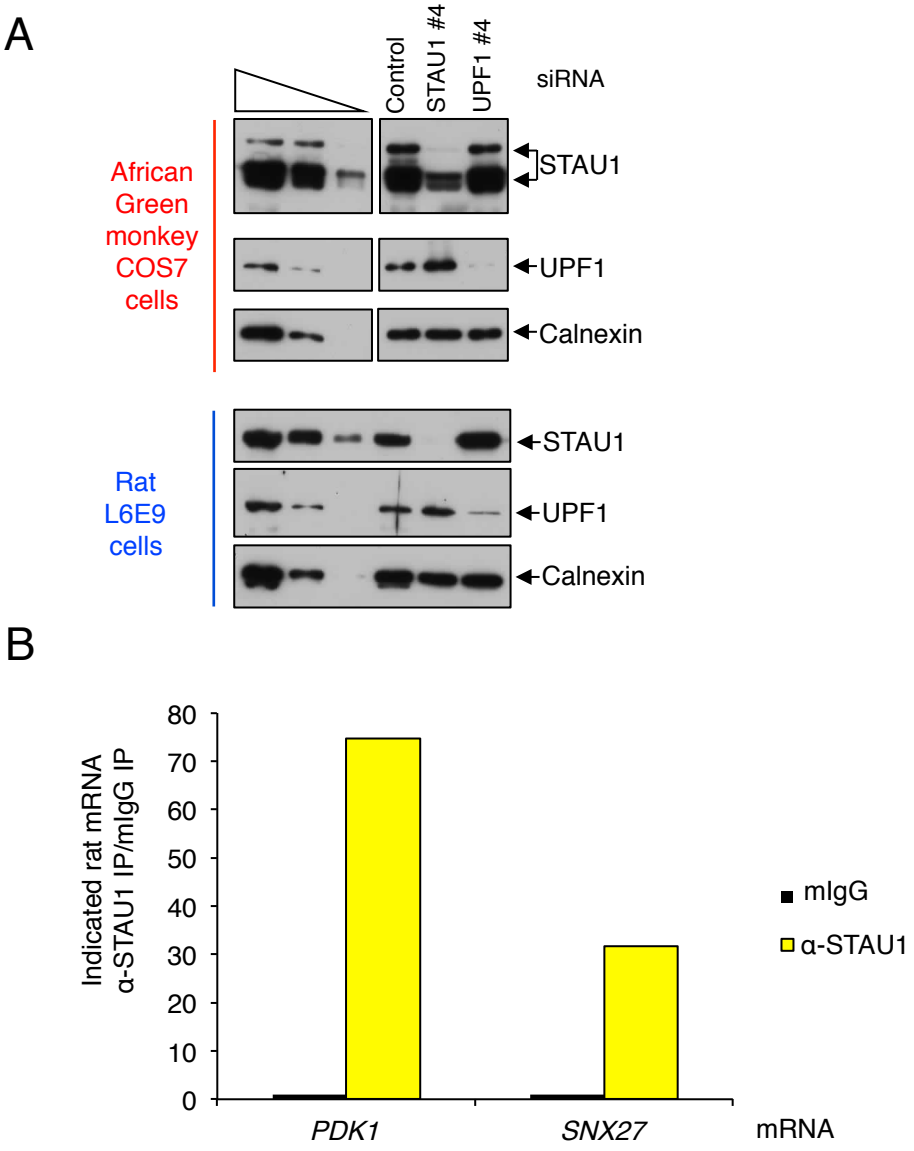


Fig. S9. (***A***) Western blots showing knockdown of STAU1 or UPF1 in cells from African Green monkey and rat using the same samples analyzed in Fig. 4. (***B***) IP showing the co-immunoprecipitaiton of rat *PDK1* and rat *SNX27* mRNAs with STAU1 from L6E9 cells.

Fig. S10. Clustal Omega alignments showing the divergence in the 3' UTR SINE sequence in (***A***) human and African Green monkey *PDK1* genes; (***B***) mouse and rat *PDK1* genes; (***C***) human and African Green Monkey *SNX27* genes or (***D***) mouse and rat *SNX27* genes. Asterisks indicate conserved nucleotides, dashes denote nucleotide gaps.