Appendix

Accession	Tissue Type
PRJNA544411	**human induced pluripotent stem cells (iPSCs), iPSC-derived cardiomyocytes, human foreskin fibroblasts (HFFs), and HFFs infected with cytomegalovirus**
PRJNA515538	**293T, HeLa-S3, K562**
PRJNA604580	**Late pancreatic progenitor (PP2) cells obtained from six independent differentiations of H1 human embryonic stem cell (hESC)**
PRJNA756018	**Brain, Embryonic stem cells, Fat, Fibroblast, Human aorta endothelial cells (HA_EC), Human coronary artery endothelial cells (HCAEC), Hepatocytes, Human vascular endothelial cells (HVEC), Vascular smooth muscle cells (VSMC)**
PRJNA552552	**ESCs**
PRJNA812862	**Primary subcutaneous white, beige (from subcutaneous WAT), and brown adipocytes were generated from freshly isolated subcutaneous white adipose tissue (WAT) and brown adipose tissue (BAT).**
PRJNA375080	**embryonic (E15.5) and adult (P42) Brain and Liver Tissue**

Accession	Experimental Approach Employed
PRJNA544411	Briefly, cells were rapidly harvested and lysed. Clarified cell lysates were treated with RNase I (Invitrogen) to digest RNA not protected by ribosomes. 80S ribosomes were isolated by centrifuging lysates through a 34% sucrose cushion at 100,000×g for 1 hour at 4°C. RNA was then purified from the ribosome pellet using the Direct-zol RNA kit (Zymo Research). The RNA was then resolved by electrophoresis through a denaturing gel, and the fragments corresponding to 28 to 34 bp were extracted from the gel. The 3' ends of the ribosome footprint RNA fragments were then treated with T4 polynucleotide kinase (NEB) to allow ligation of a pre-adenylated DNA linker with T4 Rnl2(tr) K227Q (NEB). The DNA linker incorporates sample barcodes to enable library multiplexing, as well as unique molecular identifiers (UMIs) to enable removal of duplicated sequences.

	To separate ligated RNA fragments from unligated DNA linkers, 5'-deadenylase (Epicentre) was used to deadenylate the pre-adenylated linkers, which were then degraded by the 5'-3' ssDNA exonuclease RecJ (NEB). After rRNA reduction using the Ribo-Zero Gold rRNA removal kit (Illumina), The RNA-DNA hybrid was used as a template for reverse transcription, followed by circularization with CircLigase (Epicentre). Finally, PCR of the cDNA circles attached suitable adapters and indices for Illumina Sequencing. The library was sequenced on an Illumina HiSeq 4000 sequencer with a single-end 50 base pair run.
PRJNA515538	For all ribosome footprinting experiments, adherent cells were grown to about 80% confluency in 10-cm or 15-cm diameter tissue culture dishes, and suspension cells were grown to a density of approximately 500,000 cells per ml. Cells were washed with 5 ml ice-cold PBS with 100 μg ml−1 cycloheximide (CHX) added. Immediately after removing PBS, 400 μl of ice-cold lysis buffer (20 mM Tris–HCl, pH 7.4, 150 mM NaCl, 5 mM MgCl2, 1% Triton X-100, with 1 mM DTT, 25 U ml−1 Turbo DNase (Thermo Fisher, catalog no. AM2238) and 100 μg ml−1 CHX added fresh) was dripped onto the plate or added to the cell pellet. Cells were incubated on ice in lysis buffer for 10 min with periodic vortexing and pipetting to disperse the cells. The lysate was then clarified by centrifugation at 15,000g for 10 min. Cell lysates were flash-frozen in liquid nitrogen and stored at −80 °C for up to 5 d before ribosome footprinting. For experiments profiling translation initiation, the same procedure was followed except for the addition of either 2 μg ml−1 harringtonine (Abcam, catalog no. ab141941) for 2 min or 20 μg ml−1 lactimidomycin (MilliporeSigma, catalog no. 506291) for 30 min to media, before PBS wash and lysis. A variety of digestion conditions were tested in this study and are summarized in Supplementary Data 1. Briefly, RNA digestions using 250 U RNase I (Thermo Fisher, catalog no. AM2294) per 100 μl lysate were used in the low-resolution 293T and HeLa-S3 experiments. For high-resolution experiments, 15–30 U TruSeq Nuclease (Illumina) was used to digest 30–60 μg RNA in up to 300 μl lysate. Digestion reactions were run for 45–60 min at room temperature and

	quenched with 100–200 U Superase-In RNase I inhibitor (Thermo Fisher, catalog no. AM2694) on ice. Following digestion, RPFs were purified from small RNA fragments using MicroSpin S-400 HR columns (GE Life Sciences) according to the TruSeq Ribo Profile kit (Illumina). Low-resolution experiments were cleaned up with Zymo RNA Clean & Concentrator-25 kit, while high resolution experiments were purified by acid phenol:chloroform extraction followed by isopropanol precipitation. Ribosomal RNAs were depleted from RPF fragments by Ribo-Zero Mammalian kit (Illumina) following the manufacturer's protocol. cDNA sequencing libraries were then prepared using the TruSeq Ribo Profile kit (Illumina) following the manufacturer's protocol. Single-end 50-base reads were collected for each library on an Illumina HiSeq 2500 with no more than four samples sequenced on a single lane. Each Ribo-seq experiment was prepared from a different biological replicate, except for K562 HiRes1 and 2, which were prepared from the same lysate using different digestion conditions. For K562 HiRes3, CHX was added to the media before pelleting cells and washing with PBS.
PRJNA604580	Ribosome profiling was performed on PP2 cells obtained from six independent differentiations of H1 hESCs with the Rezania et al., 2014 protocol, yielding an average of 89% PDX1-positive cells. Ribosome footprinting and sequencing library preparation was done with the TruSeq Ribo Profile (Mammalian) Library Prep Kit (Illumina, Cat# RPYSC12116, currently out of production) according to the TruSeq Ribo Profile (Mammalian) Reference Guide (version August 2016). In short, 50 mg of PP2 aggregates were washed twice with cold PBS and lysed for 10 min on ice in 1 mL lysis buffer (1 × TruSeq Ribo Profile mammalian polysome buffer, 1% Triton X-100, 0.1% NP-40, 1 mM dithiothreitol, 10 U ml−1 DNase I, cycloheximide (0.1 mg/ml) and nuclease-free H2O). Per sample, 400 µL of lysate was further processed according to manufacturer's instructions. Final library size distributions were checked on the Bioanalyzer 2100 using a High Sensitivity DNA assay (Agilent Technologies), multiplexed and sequenced on an Illumina HiSeq 4000

	producing single end 1 × 51 nt reads. Ribo-seq libraries were sequenced to an average depth of 85M reads.
PRJNA756018	For brain and adipose tissue, 80-100mg of tissue were placed in chilled tubes containing zirconia beads (11079110zx, Biospec) and 1ml cold lysis buffer supplemented with 0.1 mg/mL cycloheximide (formulation as in TruSeq Ribo Profile Mammalian Kit, RPHMR12126, Illumina) and lysed using the Magnalyser machine (Roche) in pulses of 20s at 6000g so that the sample would remain cold. Samples were then centrifuged at 20,000g for 10min at 4C to pellet debris. Fibroblasts, hepatocytes, hESCs and endothelial cells were grown to 90% confluence in a 10cm culture dish, while VSMCs were pooled from 5 wells of a 6-well culture plate at baseline conditions in basal M231 medium (M-231-500) for 24h, and pelleted before snap-freezing in liquid nitrogen and stored at -80C prior to lysis for Ribo-seq. It was ensured that primary cells at low passage (% passage 4) were used for these experiments. Cell lysis occurred in the presence of 0.1 mg/mL cycloheximide in 1ml cold lysis buffer (formulation as in TruSeq Ribo Profile Mammalian Kit, RPHMR12126, Illumina). After immediate repeated pipetting and multiple passes through a syringe with a 21G needle, sample lysates were cleared as described above. 400-800ul of supernatant recovered from homogenized and cleared lysates were then footprinted with Truseq Nuclease (Illumina). Ribosomes were purified using Illustra Sephacryl S400 columns (GE Healthcare), and the protected RNA fragments were extracted with a standard phenol:-chloroform:isoamylalcohol technique. Following ribosomal RNA removal (Mammalian RiboZero Magnetic Gold, Illumina), sequencing libraries were prepared out of the footprinted RNA according to the TruSeq Ribo Profile (Mammalian) Reference Guide, with the additional modification of 8% PAGE purification following the PCR amplification of the final library. The final RNA-seq and ribosome profiling libraries were quantified using KAPA library quantification kits (Roche); the quality and average fragment size of the final libraries were determined using a LabChip GX HT DNA HiSens Reagent Kit (Perkin Elmer). Libraries with unique indexes were

	pooled and sequenced on a Hiseq / NextSeq 500 Illumina sequencer using 75-bp paired-end [RNA-seq: NextSeq 500 High Output kit v2 (150 cycles)] or 50-bp single-end [Ribo-seq: NextSeq 500 High Output kit v2 (75 cycles)] sequencing chemistry.
PRJNA552552	Ribosome profiling was performed as previously described65, with minor modifications. Briefly, 500 µg of the ribonucleoproteins (two biological replicates per sample) were treated with 1000 U RNase I (Ambion Cat# AM2295) at 4 °C for 50 min with gentle end-over-end rotation followed by incubation with SuperaseIn (Ambion, Cat# AM2696). Monosomes were pelleted by ultracentrifugation in a 34% sucrose cushion at 70,000 r.p.m. in a TLA-120.2 rotor (Beckman Coulter) at 4 °C for 3 h. RNA fragments were extracted with acid phenol (2×), once with chloroform, and precipitated with isopropanol at −20 °C in the presence of NaOAc and GlycoBlue (Invitrogen). Purified RNA samples were resolved on a denaturing 15% polyacrylamide-urea gel and the sections corresponding to 28–32 nucleotides containing the RFPs were excised, eluted, and precipitated by isopropanol. Purified RFPs were dephosphorylated using T4 polynucleotide kinase (New England Biolabs) for 1 h at 37 °C. Denatured fragments were re-suspended in 10 mM Tris (pH 7) and quantified using the Bio-Analyzer Small RNA assay (Agilent). A sample of 10 pmol of RNA was ligated to the 3 ´ -adaptor with T4 RNA ligase 1 (New England Biolabs) for 2 h at 37 °C. Reverse transcription was carried out using oNTI223 adapter (Illumina) and SuperScript III reverse transcriptase (Invitrogen) according to the manufacturer's instructions. Products were separated from the empty adaptor on a 10% polyacrylamide Tris/Borate/EDTA-urea (TBE-urea) gel and circularized by CircLigase (Epicentre). Ribosomal RNA amounts were reduced by subtractive hybridization using biotinylated rDNA complementary oligos65. The RFP libraries were amplified by PCR (11 cycles) using indexed primers and quantified using the Agilent BioAnalyzer High-Sensitivity assay. DNA was then sequenced on the HiSeq-2000 platform with read length of 50 nucleotides (SR50) according

	to the manufacturer's instructions, with sequencing primer oNTI202 (5CGACAGGTTCAGAGTTCTACAGTCCGACGATC).
PRJNA812862	For sample collection of BAT/WAT/Beige cultures for Ribo-Seq and total mRNA-Seq used for smORF discovery, each culture was washed directly out of the incubators 2x with ice-cold PBS supplemented with 100 mg/mL cycloheximide (CHX; Fisher Scientific, AAJ66004X). After the last wash, liquid nitrogen was gently ladled onto the surfaces of the cells and plates were stored at 80C prior to ribosome footprinting and preparation of sequencing libraries. Each biological replicate for Ribo-Seq in our experiments represents 11 wells of a 6-well dish leaving one well to be lysed with Trizol reagent (Thermo #15596026) for the bulk mRNA-Seq used in the de novo transcriptome assembly of the cells under analysis. For the Ribo-Seq analysis, two separate primary isolations along with their preps for both RNA-Seq and Ribo-Seq were performed for the discovery work. Cells were lysed with 400 mL of ice-cold lysis buffer (20 mM Tris–HCl, pH 7.4, 150 mM NaCl, 5 mM MgCl2, 1% Triton X-100, with 1 mM DTT, 25 U ml 1 Turbo DNase (ThermoFisher, catalog no. AM2238) and 100 mg mL 1 CHX added fresh) was dripped onto the plate. Cells were incubated on ice in lysis buffer for 10 min with periodic vortexing and pipetting to disperse the cells. The lysate was then clarified by centrifugation at 15,000 g for 10 min. Cell lysates were flash-frozen in liquid nitrogen and stored at -80C for up to 7 days before ribosome footprinting. For each cell type, ribosome footprinting was carried out by digesting 40–60 mg of RNA in 200–300 mL lysate with 0.375 U/mg RNase I (Lucigen, N6901K) for 50 min at room temperature. Digestion reactions were quenched with 200 U Superase-In RNase I inhibitor (Thermo Fisher, AM2694) on ice. Following digestion, monosomes were purified from small RNA fragments using Micro-Spin S-400 HR columns (GE Life Sciences), and ribosome protected RNA fragments (RPFs) were extracted by acid phenol chloroform and isopropanol precipitation. Sequencing libraries were prepared as in McGlincy and Ingolia with some modifications. First, the Ribo-Zero Mammalian Kit (Illumina) was used to

	deplete rRNA after RPF extraction and just prior to RPF size selection by gel extraction. Second, the Zymo clean & concentrator step after adaptor ligation is omitted and the reaction was carried over straight into reverse transcription. For the reverse transcription step to form cDNA, Episcript RT (Lucigen, ERT12910K) was used. Following reverse transcription, excess primer was degraded using Exonuclease I (Lucigen, X40520K) and the RNA templates were degraded using Hybridase (Lucigen, H39500). For the cDNA circularization step, CircLigase I (Lucigen, CL4111K) was used. PCR amplification was then carried out using Phusion Hot Start II High-Fidelity Master Mix (Thermo Fisher, F565L) for 9–12 cycles.
PRJNA375080	Frozen tissue samples were lysed using 1 ml of mammalian lysis buffer (200 µl of 5× Mammalian Polysome Buffer, 100 µl of 10% Triton X-100, 10 µl of DTT (100 mM), 10 µl of DNase I (1 U/µl), 2 µl of cycloheximide (50 mg/ml), 10 µl of 10% NP-40 and 668 µl of nuclease-free water). After incubation for 20 min on ice, the lysates were cleared by centrifugation at 10 000 × g and 4°C for 3 min. For each tissue and replicate sample, the lysate was divided into 300- and 100-µl aliquots. For the 300-µl aliquots of clarified lysates, 5 units of ARTseq Nuclease were added to each A260 lysate, and the mixtures were incubated for 45 min at room temperature. Nuclease digestion was stopped by the addition of 15 µl of SUPERase·In RNase Inhibitor (Ambion). Subsequently, the lysates were applied to Sephacryl S-400 HR spin columns (GE Healthcare Life Sciences), and ribosome-protected fragments were purified using the Zymo RNA Clean & Concentrator-25 kit (Zymo Research). Ribosomal RNA was depleted using the Ribo-Zero magnetic kit (Epicentre). Sequencing libraries of ribosome-protected fragments (RPFs) were generated using the ARTseq™ Ribosome Profiling Kit (Epicentre, RPHMR12126), according to the manufacturer's instructions. From the 100-µl aliquots of clarified lysates, poly(A)+ RNAs were extracted and purified, and sequencing libraries of poly(A)+ RNAs were then generated using the VAHTSTM mRNA-seq v2 Library Prep Kit from Illumina (Vazyme Biotech, NR601-01) according to the manufacturer's instructions. The resulting 48 barcoded libraries were

	pooled and sequenced using an Illumina HiSeq 2500 instrument in single-end mode.

Table 1. List of Experimental Employed For Ribosome Profiling

Accession	Bioinformatics Methods Employed
PRJNA544411	For processing of ribosome profiling data, linker sequences were removed from sequencing reads and samples were de-multiplexed using FASTX-clipper and FASTX-barcode splitter (FASTX-Toolkit). Unique molecular identifiers and sample barcodes were then removed from reads using a custom Python script. Bowtie v1.1.2 was used to filter out reads aligning to rRNAs and contaminants, and all surviving reads were aligned to the custom transcriptome described above with Tophat v2.1.1 using the --b2-very-sensitive --transcriptome-only --nonovel-juncs --max-multihits=64 flags. These alignments were assigned a specific P-site nucleotide using a 12-nt offset from the 3' end of reads.
PRJNA515538	Ribo-seq and accompanying short fragment total RNA-seq reads were first trimmed of excess 3'adapter sequences using the FASTX toolkit. Trimmed Ribo-seq reads aligning to transfer RNA and ribosomal RNA sequences were then removed using STAR v.2.5.2b. Next, the remaining Ribo-seq reads were aligned to the UCSC hg19 human genome assembly containing chromosomes 1–22, X and Y, with the hg19 refGene transcript annotation using STAR. Up to two mismatches were allowed during alignment, keeping only uniquely mapped reads. Ribo-seq and RNA-seq alignments were checked for overall quality using the CollectRnaSeqMetrics script from the Picard Tools software suite.
PRJNA604580	Prior to mapping, ribosome-profiling reads were clipped for residual adapter sequences and filtered for mitochondrial, ribosomal RNA and tRNA sequences. Next, all mRNA and ribosome profiling data were mapped to the Ensembl 87 transcriptome annotation of the human genome hg38 assembly using STAR 2.5.2b (Dobin et al., 2013) in 2-pass mapping

	mode. To avoid mRNA-seq mapping biases due to read length, the 2 × 101 nt mRNA-seq reads were next trimmed to 29-mers, and those mRNA reads were processed and mapped with the exact same settings as the ribosome profiling data. For the mapping of 2 × 101 nt RNA-seq reads six mismatches per read were allowed (default is 10), whereas two mismatches were permitted for the Ribo-seq and trimmed mRNA-seq reads. To account for variable ribosome footprint lengths, the search start point of the read was defined using the option --seedSearchStartLmaxOverLread, which was set to 0.5 (half the read, independent of ribosome footprint length). Furthermore, --outFilterMultimapNmax was set to 20 and --outSAMmultNmax to 1, which prevents the reporting of multimapping reads.
PRJNA756018	Ribo-seq and RNA-seq data were pre-processed as described previously Raw sequencing data were demultiplexed with bcl2fastq V2.19.0.316 to obtain fastq format files. The fastq file was processed to remove adaptors and low-quality bases using Trimmomatic V0.36. Demultiplexing and trimming of adaptors were carried out for both Ribo-seq and RNA-seq reads. Reads that were shorter than 20 nucleotides for Ribo-seq and 35 nucleotides for RNA-seq were discarded. RPFs represent the actively translated mRNA and the ribosomal RNA (rRNA), mitochondrial RNA (mtRNA) and transfer RNA (tRNA) sequences are considered contaminant sequences. Trimmed Ribo-seq reads were aligned using Bowtie2 to sequences present on RNACentral (release 5.0) database (The RNA central Consortium et al., 2017), a database of known rRNA, mtRNA and tRNA sequences. The reads aligned to these contaminant sequences were discarded and the remaining unaligned reads were retained for further processing. After these pre-processing steps, both RNA-seq and Ribo-seq reads were aligned using STAR (Dobin et al., 2013) to the human genome (hg38) using the combined transcript models from Ensembl. Ribo-seq datasets were further screened for their 3nt-periodicity signal across known ORFs to select high-quality data for smORF detection. RiboTISH (Zhang et al., 2017a) was used to quantify the transcriptome-

	wide periodicity near the start-codon and stopcodon in the Ribo-seq dataset.
PRJNA552552	**Analysis of ribosome profiling and RNA-seq data.** To analyze the ribosome profiling data, FASTQ reads were processed as previously described65. Briefly, adaptor sequences were removed using fastx_clipper (fastx_toolkit-0.0.14) and by employing the following parameters clipper -Q33 -a CTGTAGGCACCATCAAT -l 25 -c -n -v. Reads were then trimmed using fastx_trimmer (fastx_toolkit-0.0.14) and -Q33 -f 2 parameters. Both RNA-seq reads and the trimmed RFP reads were then aligned against mouse rRNA, tRNA, snRNA, snoRNA, mtRNA sequences using Bowtie and -seedlen=23 to deplete these small RNA contamination. Unmapped reads (cleaned) were then mapped against mouse genome (mm9) using GSNAP and by employing the following parameters: -B 5 -t 15 -N 1 -E 100 -w 100000 -n 10 -s mm9refGene_splice. Mapped reads were counted in the coding sequence using HTseq-count68 with the settings: -m union -s reverse -t CDS. RFP and RNA-Seq read counts were normalized for quantification of the TE quantification and differentially translation analysis using Xtail algorithm v.1.1.5 with minMeanCount=50. Genes were considered differentially translated if they had at least 50 normalized reads in RNA, 25 normalized reads in RFP, displayed absolute fold change >1.5 and FDR ≤ 0.1. The DESeq package70 was also used to normalize the RFP and RNA-Seq read counts (with similar results to Xtail normalization) and the results were employed to correlate the variation in RNA, RFP, and protein levels. Graphs were generated using data.table and ggplot2 R packages in R version 3.5.1 on Ubuntu 16.04.5 LTS. **Quality control of Ribo-Seq data.** To analyze the position of initiating ribosome in relation to the annotated translation initiation sites (TISs), we employed a custom pipeline as reported previously52. Briefly, BAM files were converted to strand specific 5´ end wiggle files using a custom Python script. Wiggle files were then converted to a format suitable for the Batch PositionConverter Interface in Mutalyzer 2.0.beta-32. These converted files were then manually loaded into Mutalyzer to retrieve

	positions relative to the annotated TIS. We analyzed the first position of the aligned reads to transcript coordinates and relating those coordinates to annotated TISs positions located up to −15 nt surrounding the TIS were counted as positions in coding regions For all samples, a major peak was observed at −12 nt from the annotated TIS. To calculate the read distribution and coverage at 5 ´ UTR, CDS, and 3 ´ UTR regions, we first downloaded all the 5 ´ UTR, CDS, and 3 ´ UTR sequences from Ensemble BIOMART mm9 database. CDS sequences were extended by 50 nt upstream for 5 ´ UTR analysis and by 200 nt downstream for 3 ´ UTR analysis. We used CDS sequences with more than 1000 nt for this analysis. The generated BED file for 5 ´ UTR-CDS or 3 ´ UTR-CDS were binned into 10 nt bins and the number of reads from RNA-seq and RFP-seq libraries were counted using BEDtools, Samtools and in house script (peakstats.py).
PRJNA812862	After trimming adapters, removal of mm10 rRNA and tRNA sequences, and alignment to the mm10 genome with STAR v2.53b,88 instead of Cufflinks,89 Stringtie v2.1.489 and MAPS v1.090 were run using default parameters on combined alignments from total and mRNA RNA-Seq libraries from BAT, beige, and WAT tissues, followed by 3-frame translated using a custom script GTFtoFasta (see Martinez et al.4 for details). The resulting 3-frame translated ORF databases were then scored for translation using RibORF28 with the pipeline described by Martinez et al.4 which briefly included shifting reads of each length to obtain base-pair resolution, and then scoring each candidate smORF using RibORF, keeping only highly scoring (scoreR0.7), short (%150 aa), and novel (not overlapped with known coding regions in RefSeq, and a maximum blastp alignment evalue to SwissProt. The resulting microproteins were then collated into a non-redundant table and were further annotated using HOMER91 (genomic location of their ORFs), and using PhyloCSF92; conservation from a multi-way alignment of mammals) as described previously.

	Candidate microproteins were then split into categories based on the genomic location of their ORF with respect to genomic features: uORFs exist upstream of known genes, non-uORFs are all other ORFs, iORFs specifically exist in intergenic regions, while dORFs exist downstream of known genes. These categories of ORFs were then further tested for expression in RNA-Seq datasets and visualized using read pileup tracks.
PRJNA375080	**Sequencing data preprocessing.** The raw sequence reads were demultiplexed using CASAVA (v1.8.2), and the 3´-end adapter was clipped using Cutadapt (v1.8.1) (with the parameters '-a AGATCGGAAGAGCACACGTCTGAACTCCAGTCA -match-read-wildcards -m 6´). Low-quality sequences were trimmed using Sickle (v1.33) (with the parameters '-q 20´). The trimmed reads were filtered by length based on the ranges [25, 34] for ribosome-associated footprints and [20, 50] for mRNA. The retained reads that mapped to reference mouse rRNAs or tRNAs were then removed, and the remaining reads were aligned to the mouse reference genome (downloaded from GENCODE, Release M18: GRCm38.p6) using Tophat2 (v2.0.14) (13) with the following command: 'tophat2 -g 20 -N 2 –transcriptome-index [index_file] -G [gtf_file] [fastq_file] -o [output_directory]'. Only those uniquely mapped reads were extracted for gene expression determination. The number of reads per gene was counted using the Subread R package-featureCounts (v1.6.2). **Triplet periodicity analysis.** Briefly, footprint profiles within coding sequences (CDSs) of canonical protein-coding genes were produced by assigning ribosomal P-sites to each nucleotide position per codon, that is, reading frames 1, 2 and 3. The average footprint density of metagene profiles along the CDS was calculated by dividing the number of P-sites in each of the three reading frames by the total number of P-sites within the CDS. In contrast to the RNA-seq reads that mapped evenly to the three sub-codon positions, the ribosome-associated footprints mapped primarily to the first nucleotide of the codon, that is, reading frame 1.

PRJNA507253	Raw sequencing data were demultiplexed with bcl2fastq V2.19.0.316, and the adaptors were trimmed using Trimmomatic V0.36 retaining reads longer than 20 nucleotides postclipping. RNA sequencing reads were further clipped with FASTX Toolkit V0.0.14 to 29 nucleotides, to allow comparison directly with ribosome profiling reads. Reads were aligned using bowtie2 to known mitochondrial RNA, ribosomal RNA, and transfer RNA sequences (RNACentral,22 release 5.0); aligned reads were filtered out to retain only ribosome protected fragments (RPFs). Alignment to the human genome (hg38) was carried out using STAR. Gene expression was quantified on the coding sequence region using uniquely mapped reads (Ensembl database release GRCh38 v86 combined with additional transcripts from RefSeq GRCh38, downloaded January 2018) with feature counts.24 Genes with mean transcripts per million mapped reads (TPM) <1 in either RNA sequencing or ribosome profiling across all conditions were removed prior to downstream analysis. Ribotaper was used to obtain the in-frame reads around the start and stop codon.25 These peptidyl-sites (P-sites) were then visualized across samples and genes. Heatmap for the ribosome drop-off was generated using the pheatmap, 1.0.8, R package.

Table 3. List of Bioinformatics Methods Employed For Ribosome Profiling Preprocessing