SUPPLEMENTARY MATERIALS

Supplemental Methods.

Supplemental Fig S1. Compare read distribution of RNA fragments from different doses of RNase treatments.

Supplemental Fig S2. Compare RNase footprinting and ribosome profiling read distribution across mRNAs.

Supplemental Fig S3. Compare the read distribution of RNase footprinting (high dose) and ribosome profiling in noncoding regions.

Supplemental Fig S4. Compare experimental procedures of our RNase footprinting vs. conventional ribosome profiling.

Supplemental Fig S5. Low-input RNase footprinting for 50,000 and 1,000 K562 cells.

Supplemental Fig S6. The RUST codon metafootprint profiles showing adapter ligation differences across HEK293T ribosome footprinting datasets.

Supplemental Fig S7. Translational regulation after TUFM knockout in HeLa cells.

Supplemental Fig S8. RNase footprinting quantifies RNA translation in primary blood cells.

Supplemental Fig S9. Differential RNA expression and translational regulation across immune cells.

Supplemental Table S1. The RUST ratios (log2(observed/expected)) showing the codon occurrence at the 5'-end and 3'-end of 28-nt ribosome footprints.

Supplemental Table S2. Cytosolic genes showing dynamic regulation of translation efficiency after TUFM knockout. For the genes with increased translation efficiency, we annotated those with uORFs and >10 supporting reads.
Supplemental Table S3. Gene Ontology analyses results for genes showing up-regulation or down-regulation of ribosome occupancy (RFP), RNA expression, and translation efficiency (TE) after TUFM knockout. The -log₁₀(P-values) were shown.

Supplemental Table S4. Genes showing dynamic regulation of ribosome occupancy (RFP), RNA expression, and translation efficiency (TE) among human immune cell types.

Supplemental Table S5. Gene Ontology analyses results for genes showing regulation of ribosome occupancy (RFP), RNA expression, and translation efficiency (TE) comparing immune cell types. The -log₁₀(P-values) were shown.

Supplemental Table S6. The RibORF-identified translated ORFs in immune cells. ORFs were grouped based on their genomic locations. uORF: ORF in 5’UTR; dORF: ORF in 3’UTR; internal: ORF within a canonical ORF; truncation: ORF with the same stop codon of a canonical ORF and with a downstream start codon; extension: ORF with the same stop codon of a canonical ORF and with an upstream start codon.

Supplemental Table S7. Non-canonical ORFs showing dynamic regulation of ribosome occupancy among human immune cell types.

Supplemental Table S8. The primers used for low-input RNase footprinting library preparation.

Supplemental Table S9. The antibodies used for the western blots to examine protein expression.

Supplemental Table S10. Genesets used to examine the regulation of indicated pathways among immune cell types.

Supplemental Code.
**Supplemental Methods**

**Cell culture.** HEK293T and HeLa cells (ATCC) were cultured in DMEM medium (Gibco, cat. no. 11965092) supplemented with 10% FBS (Gibco, cat. no. 26140079). K562 cells (ATCC) were cultured using IMDM medium (Gibco, cat. no. 12440061) supplemented with 10% FBS.

**Mouse 4T1 breast tumor model.** All mouse experiments were in compliance with relevant ethical regulations and were approved by the Institutional Animal Care and Use Committee (IACUC) (protocol IS00010046). The 4T1 mouse breast cancer cells were purchased from the Karmanos Cancer Institute. Briefly, 4-6-week-old, female Balb/c mice were used for 4T1 xenograft experiments. 100,000 cells were resuspended in 50 μl DMEM mixed with 50 μl Matrigel (Corning 354234), and implanted into the fourth mammary gland subcutaneously. After 3 weeks, mice were CO₂ euthanized, and tumor and spleen tissues were collected and were snap-frozen in liquid nitrogen. Snap-frozen mouse tissues were powdered using a pre-cooled mortar and pestle under the continuous addition of liquid nitrogen. For each tissue type, we collected powdered tissue for RNase footprinting and RNA-seq.

**Human immune cell isolation.** Fresh human whole peripheral blood was purchased from Stemcell Technologies, and was a standard blood draw from one healthy donor (61-year-old female). PBMCs were isolated using Ficoll-Paque PREMIUM (GE Healthcare) following the manufacturer’s protocol from the 10 ml blood sample. The cells were split into four groups: one was used as the mixed PBMC control, and the other three groups were used to isolate immune cell types. CD3⁺ T cells, CD3⁻CD19⁺ B cells, and CD14⁺CD16⁻ monocytes were isolated from PBMCs using immunomagnetic isolation kits (Stemcell Technologies, cat. no. 17951, 17954, and 19359,
respectively) following the manufacturer’s instructions. We then generated RNA sequencing and replicates of RNase footprinting using the isolated cells. All sequencing libraries were generated using cells from the same one-time blood draw.

**Gene knockout.** To knock out TUFM in HeLa cells, two TUFM-targeting single guide RNAs (sgRNAs) (5’-CTAGGAGCTTGCCATGCCCG-3’, and 5’-CTAGGAGCTTGCCATGCCCG-3’) and two non-targeting control sgRNAs (5’-CGTTGGTACGTCCTCCCCG-3’, and 5’-TAAGGGGCTAACTTGGTCCC-3’) were cloned into the pLentiCRISPR V2 plasmid (Addgene, cat. no. 52961). Lentivirus was produced by transfecting HEK293T cells with pMD2.G, pPAX2, and pLentiCRISPR V2 target plasmid. After 24 h of lentiviral transduction, HeLa cells stably expressing sgRNAs were selected by treating cells with 0.5 µg/ml of puromycin for 7 days. We refreshed the medium with puromycin every two days and passaged cells when necessary. TUFM knockout efficiency was confirmed by western blotting.

**Ribosome profiling using sucrose cushion.** Ribosome profiling libraries for 1.5 million HEK293 and K562 cells were prepared based on a recently updated protocol (McGlincy and Ingolia 2017), including the experimental procedures of ribosome purification and sequencing adapter ligation. Because the original RNase recommended by the protocol is out of the market, we adjusted the RNase digestion condition according to the RNase 1 unit definition by Lucigen. Lysates containing 10 ug of total RNA were digested with 150 units of RNase 1 (Lucigen) at 25 °C for 1 h.

**Mitochondrial and cytosolic ribosome profiling using sucrose gradient.** We used a previously published method (Rooijers et al. 2013) to examine mitochondrial and cytosolic ribosome
protected footprints. Approximately 20 million HEK293T cells were treated with 100 µg/ml cycloheximide and were then lysed. Lysates were centrifuged at 1,300g and the supernatant was treated with 5 U/µl of RNase 1 (Ambion) for 1 h at room temperature (25 °C). Lysates were fractionated on a linear sucrose gradient (5-45%) using the SW-41Ti rotor at 210,000g for 2.5 h. Live absorption at 254 nm was used to track the fractions. Fractions enriched in 55S mitoribosomes and 80S cytosolic ribosomes were identified by western blotting to examine the expression of ribosome proteins, including RPS6, RPL36a, and MRPL11. These fractions were then pooled to perform footprint extraction. Ribosome footprints between 15-40 nucleotides were gel-purified and a sequencing library was prepared using the A-tailing and SMARTer oligo-based template switching method described above for RNase footprinting.

**Sequencing libraries to compare RNA fragments after digestion with different doses of RNase.** We used 0.5 million HEK293T cells for the experiments. We treated the cell lysate containing 10 µg total RNA at room temperature (25 °C) for 1 h with different RNase doses (cat. no. N6901K, Lucigen): low (0.05 U/µg), medium (0.5 U/µg), and high (50 U/µg), respectively. For the medium and high dose conditions, we selected RNA fragments with 10-40 nt to generate the sequencing library, using the same adapter ligation procedures described for ribosome profiling (McGlincy and Ingolia 2017). For the low dose RNase condition, we extracted all RNA fragments and used A-tailing and the SMARTer oligo-based template switching described above to generate the sequencing library.

**RNA sequencing library preparation.** Total RNA was isolated using the Direct-zol RNA kit (Zymo Research), and mRNAs were then purified using oligo(dT)25 magnetic beads (New England
Biolabs) according to the manufacturer’s instructions. Purified mRNAs were fragmented with NEBNext Mg2+ RNA fragmentation module (New England Biolabs) at 94 °C for 4 min. Fragmented RNAs were then precipitated overnight at -20 °C by adding 0.1 volume of 3 M sodium acetate, 10 mg of glycolblue, and 1.2 volumes of isopropanol. Purified RNA fragments were used for the library construction using the A-tailing and SMARTer oligo-based template switching method described above for RNase footprinting.

**Western blotting.** *TUFM* knockout and control cells were briefly rinsed twice with ice-cold PBS and lysed with RIPA buffer (Thermo) supplemented with protease inhibitors cocktail (Roche). Lysates were collected and the protein concentration was determined by the BCA Protein Assay kit (Pierce). Calibrated samples were diluted with 4 × LDS sample buffer (Thermo), and equal amounts of total protein were separated by SDS-PAGE using NuPAGE 12% Bis-Tris precast gels (Invitrogen). Proteins were transferred to nitrocellulose membranes using a Trans-blot Turbo Transfer System (Bio-Rad). The membranes were blocked with 5% nonfat milk in TBST, incubated with primary antibodies overnight at 4 °C, washed three times with TBST at room temperature, incubated with HRP-secondary antibodies, and imaged using the Bio-Rad Chemidoc imaging system. Protein blot intensity in Figure 4E was quantified by Image J. The antibodies used for the western blots are shown in Supplemental Table S9.

**Measurement of nascent protein synthesis.** We used AHA (L-azidohomoalanine) (Invitrogen, cat. no. C10102) for the metabolic labeling of nascent proteins and detected them using alkyne-modified fluorophores through the “click” reaction.
To detect global nascent protein synthesis rate (Fig. 4I) (Dieterich et al. 2007), cells with 70-80% confluency in a 6-well plate were rinsed with warm PBS, cultured in the methionine-free medium for 1 h, and then labeled with AHA in culture medium for 3 h. AHA labeled total proteins were clicked with the Alexa Fluor 488 alkyne (Invitrogen, cat. no. C10267), and analyzed by flow cytometry (BD LSRFortessa) and microscopy imaging (Nikon A1R confocal microscope).

To detect cytosolic or mitochondrial nascent protein synthesis rate independently (Supplemental Fig. S7GH) (Xiao et al. 2015; Shih and Hsueh 2016), mitochondrial translation inhibitor (100 µg/ml chloramphenicol, Sigma-Aldrich, cat. no. C1919) or cytosolic translation inhibitor (100 µg/ml emetine, Sigma-Aldrich, cat. no. 324693) was supplemented to the culture medium during AHA labeling. AHA-labeled cytosolic or mitochondrial translation products were clicked with TAMRA (Invitrogen, cdoseat. no. C33370), and analyzed by gel electrophoresis. In brief, protein lysates (200 µg) were incubated with Click-iT TAMRA for 2 h at room temperature and then precipitated by methanol/chloroform. Precipitated proteins were resolubilized in sample loading buffer and equal amounts of protein were subjected to NuPAGE Novex 4-12% Bis-Tris protein gel separation. TAMRA labeled gels were imaged by Bio-Rad Chemidoc using 532 nm excitation. The gel was stained with Coomassie to confirm equal sample loading.

Ribosome profiling and RNA-seq analyses. To make the analysis results comparable, we used the same RNase footprinting analysis procedures described above for the ribosome profiling and RNA-seq read mapping and gene expression calculation. The one exception was that we used all uniquely mappable reads with different fragment lengths (not just those from 18-35 nt) from RNA-seq to calculate RNA expression levels.
Comparing RNase footprinting with ribosome profiling datasets. To evaluate the performance of our RNase footprinting, we compared the obtained results with the ribosome profiling datasets generated by the conventional methods. Besides the data generated in this study which used the sucrose cushion to isolate ribosomes, we downloaded published HEK293T and K562 ribosome profiling datasets and selected two high-quality datasets for each cell type for the comparative analyses. ~29 nt footprints of these datasets show clear 3-nt periodicity across canonical ORFs. The accession numbers for the HEK293T datasets are GSE125218 (SRX5256546; using size-exclusion chromatography) and GSE70211 (SRX1070870; using sucrose cushion), and the K562 datasets are GSE125218 (SRX5256556; using size-exclusion chromatography) and GSE129061 (SRX5604287; using size-exclusion chromatography) (Iwasaki et al. 2016; Calviello et al. 2020; Martinez et al. 2020). Notably, these published studies only surveyed the subpopulation of ribosome footprints around 29 nt and did not include the 21 nt population in the sequencing library construction.

Analyses of codon usage levels, uniformness of read distribution, and PCR duplication rates. We used the published software CONCUR (Frye and Bornelov 2020) to select the read fragments showing strong 3-nt periodicity across canonical ORFs, correct read locations to ribosomal P-sites, and count the number of reads occupying each codon. Then we calculated the relative codon usage level for each codon by normalizing the occupying read count to the mean count number across codons. We used the combination of read sequence and UMI to define whether a sequencing read is from PCR duplication. The PCR duplication rate is defined as the ratio of read count after removing duplication reads vs. total.
We used our previously published method based on the percentage of maximum entropy (PME) values (Ji et al. 2015) to model the uniformness of the read distribution across an ORF region. A high PME value (close to 1) indicates uniform read coverage across codons, while a low value (close to 0) represents a highly localized distribution. We observed that a few genes show consistently low PME values across samples. We manually examined the reads for these genes and found that they are due to sequencing errors and read misalignment. We excluded these 45 genes with PME values < 0.3 from further expression analyses. Additionally, we used RUST software (O'Connor et al. 2016) to compare linker ligation differences among ribosome footprinting datasets. Only 28 nt footprints were included in the analyses with the offset correction parameter 15.

**Analyses of translation efficiency.** The translation efficiency of a gene is calculated as the ratio of RNase footprinting read density vs. RNA-seq read density in the coding region. Suppose for gene \(i\), the TPM value of ribosome footprints (RFP) is \(RFP_i\) and the TPM value of RNA expression is \(RNA_i\). The translation efficiency (TE) of gene \(i\) is calculated as follows: 
\[
TE_i = \frac{RFP_i + 1}{RNA_i + 1}
\]
The pseudocount 1 was added to mitigate the technical variance of lowly expressed genes. Genes included in the differential translation efficiency analyses should show high RNA expression levels with TPM values >3 across all compared samples, and RNase footprinting with TPM > 3 in at least one experimental condition.

**Translational regulation and RNA expression after TUFM knockout.** The regulation of ribosome occupancy is the combined effect of RNA expression and translation efficiency. Genes showing dynamic regulation of ribosome occupancy were defined based on the following criteria:
1) TPM >3 in at least one condition; 2) >1.2-fold expression change in both replicates with an average expression change >1.5-fold. The same cutoffs were used to select genes showing RNA expression regulation. Genes showing regulation of translation efficiency were selected using the following cutoff: >1.2-fold change in both replicates and an average expression change >1.5-fold. To further support the regulation of ribosomal proteins and histone genes, we examined the relative expression of all genes in the pathway comparing TUFM knockout vs. control, and used the Wilcoxon rank sum test to get the P-values. The curated gene lists are shown in Supplemental Table S10.

To examine the regulation of ribosome elongation after TUFM knockout, we used different cutoffs to split the coding regions of the mitochondrial gene into two segments: 10/90, 20/80, 30/70, 40/60, 50/50, 60/40, 70/30, 80/20, and 90/10. Then we used the Fisher’s exact test to examine whether the read ratio in two segments showed significant changes in TUFM knockout vs. control samples. Finally, we picked 20/80 as the cutoff, because it showed the most robust results across genes to capture ribosome pausing at the 5’-end of transcripts. We excluded the gene MT-ATP8 from these analyses because it did not have enough reads in unique gene regions for the calculation.

**Translational regulation and RNA expression across immune cell types.** To examine the regulation of translation efficiency, we first identified genes showing dynamic regulation from the pairwise comparisons of T cells, B cells, and monocytes, using a cutoff of >1.5-fold change in both replicates (Supplemental Fig. S9A-C). Then for the genes showing dynamic regulation in any of the pairwise comparisons, we further compared their relative expression across the immune cell types by normalizing their translation efficiency (log2 values) to the mean of T cells, B cells, and monocytes. We separated the genes into 4 groups: “T and B cell high” (the normalized
value > log₂(1.1) in both T cells and B cells, and <0 in monocytes); “T cell high” (the normalized value > log₂(1.1) in T cells, and <0 in both B cells and monocytes); “B cell high” (the normalized value > log₂(1.1) in B cells, and <0 in both T cells and monocytes); and “monocyte high” (the normalized value > log₂(1.1) in monocytes, and <0 in both T cells and B cells). The genes showing higher translation efficiency in the “T and B cell high” group capture lymphocyte-specific regulation, while the genes in other groups identify cell type-specific regulation. We found that this grouping method effectively captured cell type-specific gene regulation, and included 92% of all regulated genes from the pairwise comparison.

Using the similar method, we examined the regulation of ribosome occupancy and RNA expression across the immune cell types, except that only genes showing TPM > 3 in any one cell type were included in the analyses, and a 2-fold change was used as the cutoff to select genes showing RNA expression regulation from the pairwise comparison. To further examine the regulation of a pathway, we compared the relative expression of all genes in the pathway as one geneset across immune cells using the Wilcoxon rank sum test. We manually curated the gene lists for “ribosomal proteins” and “translation initiation factors”. The following Gene Ontology-defined genesets were also included in the analyses: “GO:0007049 cell cycle”, “GO:0006396 RNA processing”, “GO:0051276 chromosome organization”, “GO:0045449 regulation of transcription”. The gene lists are shown in Supplemental Table S10.

The analyses of non-canonical ORFs. We used the RibORF software we developed (Ji et al. 2015; Ji 2018) for the analyses. We grouped RNase footprinting or ribosome profiling reads based on fragment lengths, and selected enriched reads showing clear 3-nt periodicity across canonical ORFs for further analysis to identify translated non-canonical ORFs. These reads should show >50%
assigned to the 1st nucleotides of codons. We adjusted their 5’ end locations to the ribosomal A-sites based on defined offset distances. For the analyses of HEK293T cells in Fig. 1, we used the following read fragment lengths and corresponding offset distances: 20 nt (+15), 21 nt (+15), 22 nt (+16), 28 nt (+15), 30 nt (+16), 31 nt (+16). These parameters are the same for RNase footprinting and ribosome profiling. To identify genome-wide translated ORFs in blood cells, we merged reads (28-31 nt) in different blood cell types for the analyses with the following offset correction parameters: 28 nt (+15), 29 nt (+15), 30 nt (+16), 31 nt (+16).

The algorithm uses the following read distribution features to distinguish the in-frame actively translated ORFs vs. off-frame ORFs: (1) 3-nt periodicity across ORFs measured by the fraction of reads in 1st nucleotides of codons; (2) the uniformness of read distribution across ORFs measured by the PME value. It uses in-frame translated canonical ORFs as positive examples and internal off-frame candidate ORFs as negative examples to train the logistic regression model to identify genome-wide translated ORFs. It randomly picked 1,000 positive examples and 2,000 negative examples for training and used another 1,000 positive examples and 2,000 negative examples for testing. We obtained the candidate ORFs (any possible ORF with a start codon NUG/ACG and a stop codon) using transcripts defined by GENCODE (v28) (Frankish et al. 2019). We used the ROC curve to measure the algorithm performance for classifying positive and negative examples in the testing set. The translated ORFs were selected using the following cutoffs: translation probability >0.7 and >10 supporting reads.

To examine the differential ribosome occupancy in non-canonical ORFs, we used footprinting reads with 18-35 nt for the analysis and required that the ORFs should contain >10 reads in at least one immune cell type (i.e. T cells, B cells, and monocytes). For overlapping uORFs, we excluded the regions overlapping with canonical ORFs from the analyses. We used the similar
analyses steps presented above for canonical coding regions to identify non-canonical ORFs showing >1.5-fold differential ribosome occupancy among the immune cells.

**Gene Ontology analyses.** Gene Ontology analyses were conducted using the DAVID database (Huang da et al. 2009).

**References**

Calviello L, Hirsekorn A, Ohler U. 2020. Quantification of translation uncovers the functions of the alternative transcriptome. *Nat Struct Mol Biol* **27**: 717-725.

Dieterich DC, Lee JJ, Link AJ, Graumann J, Tirrell DA, Schuman EM. 2007. Labeling, detection and identification of newly synthesized proteomes with bioorthogonal non-canonical amino-acid tagging. *Nat Protoc* **2**: 532-540.

Frankish A, Diekhans M, Ferreira AM, Johnson R, Jungreis I, Loveland J, Mudge JM, Sisu C, Wright J, Armstrong J et al. 2019. GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res* **47**: D766-D773.

Frye M, Bornelov S. 2020. CONCUR: quick and robust calculation of codon usage from ribosome profiling data. *Bioinformatics* doi:10.1093/bioinformatics/btaa733.

Huang da W, Sherman BT, Zheng X, Yang J, Imamichi T, Stephens R, Lempicki RA. 2009. Extracting biological meaning from large gene lists with DAVID. *Curr Protoc Bioinformatics* **Chapter 13**: Unit 13 11.

Iwasaki S, Floor SN, Ingolia NT. 2016. Rocaglates convert DEAD-box protein eIF4A into a sequence-selective translational repressor. *Nature* **534**: 558-561.

Ji Z. 2018. RibORF: Identifying Genome-Wide Translated Open Reading Frames Using Ribosome Profiling. *Curr Protoc Mol Biol* **124**: e67.
Ji Z, Song R, Regev A, Struhl K. 2015. Many lncRNAs, 5'UTRs, and pseudogenes are translated and some are likely to express functional proteins. *Elife* 4: e08890.

Martinez TF, Chu Q, Donaldson C, Tan D, Shokhirev MN, Saghatelian A. 2020. Accurate annotation of human protein-coding small open reading frames. *Nat Chem Biol* 16: 458-468.

McGlincy NJ, Ingolia NT. 2017. Transcriptome-wide measurement of translation by ribosome profiling. *Methods* 126: 112-129.

O'Connor PB, Andreev DE, Baranov PV. 2016. Comparative survey of the relative impact of mRNA features on local ribosome profiling read density. *Nat Commun* 7: 12915.

Rooijers K, Loayza-Puch F, Nijtmans LG, Agami R. 2013. Ribosome profiling reveals features of normal and disease-associated mitochondrial translation. *Nat Commun* 4: 2886.

Shih YT, Hsueh YP. 2016. VCP and ATL1 regulate endoplasmic reticulum and protein synthesis for dendritic spine formation. *Nat Commun* 7: 11020.

Xiao L, Xian H, Lee KY, Xiao B, Wang H, Yu F, Shen HM, Liou YC. 2015. Death-associated Protein 3 Regulates Mitochondrial-encoded Protein Synthesis and Mitochondrial Dynamics. *J Biol Chem* 290: 24961-24974.
RNase concentrations

A

Marker  RNA ladder 1  RNA ladder 2
Low  Medium 1  Medium 2  High

1000 500 150

B

Relative read enrichment

CDS 5'UTR 3'UTR Intron

Low  Medium  High

Ribosome profiling
RNA-seq

polyA site of AKT1S1

C

D

Relative enrichment to coding regions

Flanking polyA sites (nt)

Low  Medium  High

RNA-Seq

E

F

G

18-24 nt

25-35 nt

18-35 nt

R=0.882

R=0.882

R=0.986

R=0.994

R=0.945

R=0.988

R=0.994

R=0.994

H

RNase high-dose

RNase high-dose

RNase high-dose

 RNase high-dose

- replicate 1

- replicate 2
Supplemental Fig S1. Compare read distribution of RNA fragments from different doses of RNase treatments.

(A) The gel showing the size distribution of RNA fragments after different doses of RNase treatment.

(B) Relative enrichment of CLIP-seq reads in different regions of mRNAs. We analyzed K562 CLIP-seq data for translation factors generated by ENCODE, including EIF3G (GSM5369830), EIF4G2 (GSM2423297), RPS11 (GSM2423243), and RPS3 (GSM5369826). The calculation method is same as in Figure 1D.

(C) The example gene AKT1S1. The enlarged region shows the read distribution around the polyadenylation (polyA) site with a low dose RNase treatment. RNA-seq and ribosome profiling data are shown for comparison.

(D) The read distribution surrounding the genome-wide polyA sites.

(E-G) Correlation between RNase treatments and ribosome profiling in HEK293T cells: (E) Low-dose RNase treatment vs. ribosome profiling; (F) Medium-dose RNase treatment vs. ribosome profiling; (G) High-dose RNase footprinting vs. ribosome profiling. The reads were grouped based on their indicated length ranges: 18-24 nt and 25-35 nt. The X-axis and Y-axis represent log2(read count+1) in coding regions.

(H) Correlation between two biological replicates of high-dose RNase footprinting.
Supplemental Fig S2. Compare RNase footprinting and ribosome profiling read distribution across mRNAs.

(A-C) Read Distribution around start codons and stop codons of mRNAs: (A) RNA-seq data; (B) Low-dose RNase; (C) Medium-dose RNase (32-nt fragments).

(D) Distribution of high-dose RNase footprinting reads grouped based on fragment sizes (18-35 nt). Y-axis represents the Read per Million (RPM) value. The read location was represented by 5’-end of the alignment position.

(E) Similar to (D), the read distribution of conventional ribosome profiling data were shown.

(F) The log$_2$ ratio of read density in coding regions vs. 3'UTRs for each fragment size between 18-35 nt for high-dose RNase footprinting data.

(G) The ROC curve measuring the performance using HEK293T RNase footprinting (high-dose) data to classify in-frame translated canonical ORFs vs. internal off-frame ORFs. The area under the ROC curve (AUC) value is shown.
Supplemental Fig S3. Compare the read distribution of RNase footprinting (high dose) and ribosome profiling in noncoding regions.

(A) The fractions of reads aligned to different types of RNAs.

(B) The ratios of RNase footprinting and ribosome profiling reads (measured by RPM values) for different types of small noncoding RNAs.

(C) The fractions of reads mapped to different small noncoding RNAs using the RNase footprinting data.

(D-E) Example genes showing the read distribution in small noncoding RNAs: MIR1248 (D) and SNORA3A (E). The functional domains of the RNAs are highlighted in yellow. The Y-axis represents read per million (RPM) values.

(F) Read distribution patterns representing translated ORFs and non-ribosomal RNA-protein complexes.

(G) Distribution of PME values measuring the uniformness of read distribution across candidate ORFs in the indicated types of transcripts. We included ORFs with >10 reads in the analyses. A PME value closer to 1 indicates a higher uniformness of read distribution. Small noncoding RNAs show low PME values indicating localized read distribution, representing the binding of non-ribosomal RNA-protein complexes.

(H) The ROC curve showing the accuracy using the PME values to classify candidate ORFs from ribosomal footprints in mRNAs vs. non-ribosomal footprints in small noncoding RNAs. The area under the ROC curve is shown.

(I) The length distribution of footprints across different RNA regions. We also plotted the fragment size distribution for rRNAs, as 78% of total sequencing reads were aligned to rRNAs.
Supplemental Fig S4. Compare experimental procedures of our RNase footprinting vs. conventional ribosome profiling.
A. Frequency distributions of read per million across different coding regions.

B. Ribosome occupancy (log2).

C. Codon usage levels (log2).

D. PME measuring uniformness of read distribution in canonical ORFs.

E. Genes detected vs. sampling reads.

F. log2(number of unique footprints).

G. HEK293T log2(TPM value) frequency distribution.

H. K562 log2(TPM value) frequency distribution.

I. Cutoff HEK293T K562
   TPM > 0 12,532 11,926
   TPM > 3 9,646 8,994
Supplemental Fig S5. Low-input RNase footprinting for 50,000 and 1,000 K562 cells.

(A) RNase footprinting reads (using 1,000 K562 cells) grouped by fragment sizes. The read location was represented by 5’ end of the alignment position.

(B) Correlation of ribosome occupancies inferred by our RNase footprinting and conventional ribosome profiling datasets. The X-axis and Y-axis represent log2(RPM+1) in coding regions. The Pearson correlation coefficient values are indicated in the plots.

(C) Correlation of codon usage levels (log2) at the ribosomal P-sites inferred by our RNase footprinting and conventional ribosome profiling datasets. The Pearson correlation coefficient values are indicated in the plots.

(D) The distribution of the percentage of maximum entropy (PME) values measuring the uniformness of read coverage across codons of canonical ORFs.

(E-F) We downsampled of the sequencing reads and then calculated the number of genes (E) and unique footprints (F) detected by different datasets with a fixed number of reads.

(G-H) The distribution of log2(TPM) values for coding genes measured by our RNase footprinting data in HEK293T (G) and K562 (H) cells.

(I) The number of genes detected using two different cutoffs: TPM > 0 and TPM > 3.
Supplemental Fig S6. The RUST codon metafootprint profiles showing adapter ligation differences across HEK293T ribosome footprinting datasets (A-D). The 28-nt footprints were used in the analyses with the offset correction distance 15 nt. The grey lines are 61 curves showing the RUST ratio values of 61 sense codons across mRNAs. The greater RUST ratio indicates higher observed occurrence than expected. The corresponding Kullback–Leibler divergence (K–L) is shown in blue. The higher K-L divergence value in -5 or +3 codon positions indicates stronger linker ligation bias in 5'-end or 3'-end of the footprints. We showed the detailed RUST ratio values in these positions in Supplemental Table S1.
Supplemental Fig S7. Translational regulation after *TUFM* knockout in HeLa cells.

(A) The distribution of footprint lengths in protein-coding regions.

(B) The log$_2$ ratio of read density in coding regions vs. 3’UTRs.

(C) Read distribution around the start and stop codons of mRNAs (29-nt footprints). We adjusted the 5’-end genomic locations of the reads to the ribosomal A-sites.

(D) The correlation between the differential regulation of translation efficiency (TE) vs. ribosome pausing. The Pearson correlation coefficient and the linear regression $P$-value are shown.

(E) The regulation of ribosome elongation of cytosolic-translated genes after *TUFM* knockout. The ribosome pausing level was calculated as the ratio of RNase footprinting read density in the first 20% of a transcript vs. the remaining.

(F) The regulation of ribosome footprint (RFP) levels, RNA expression, and translation efficiency (TE) of mitochondrial respiratory chain proteins encoded by the nuclear genome.

(G-H) The pulse labeling AHA assay to show the downregulation of mitochondrial (G) and cytosolic (H) nascent protein synthesis levels, respectively, comparing *TUFM* knockout cells vs. control. The Coomassie blue stainings for total proteins were shown as loading controls.

(I) RNA expression regulation of cytosolic genes. Red: up-regulated genes; blue: down-regulated genes. The total number of differentially regulated genes is indicated.

(J) The regulation of ribosome occupancy (RFP) of cytosolic genes.
Supplemental Fig S8. RNase footprinting quantifies RNA translation in primary blood cells.

(A) Protein synthesis rates in primary PBMCs vs. K562 cancer cells are compared in density plots based on AHA fluorescence levels.

(B) The distribution of footprint lengths in coding regions of mRNAs using footprinting data from different blood cells.

(C) The log₂ ratio of read density in coding regions vs. 3'UTRs for each fragment size (18-35 nt) for the footprinting data.

(D) Adjusted read distribution (29-nt fragments) around the start and stop codons of mRNAs using the PBMC footprinting data.

(E) Correlation of ribosome occupancy levels between two PBMC replicates. The X-axis and Y-axis represent log₂(read count+1) in coding regions.

(F-H) The marker gene expression confirms the purification of T cells, B cells, and monocytes. The read distribution of RNase footprinting and RNA-seq across the marker genes: CD3 (CD3E and CD3D) (F), CD19 (G), and CD14 (H).
Supplemental Fig S9. Differential RNA expression and translational regulation across immune cells.

(A) Change in mRNA and ribosome footprint (RFP) levels in T cells vs. monocytes. The genes showing significant differential translation efficiency are highlighted. Yellow: down-regulated in T cells; purple: up-regulated in T cells.

(B) As in (A), comparing B cells vs. monocytes.

(C) As in (A), comparing T cells vs. B cells.

(D-F) Venn diagrams showing the overlap of genes showing significant regulation of RNA expression and translation efficiency (TE) comparing T cells vs. monocytes (D), B cells vs. monocytes (E), and T cells vs. B cells (F).

(G) Heatmap showing genes with significant regulation of ribosome occupancy (ribosome footprints, RFP) across different types of immune cells.

(H-J) The relative expression levels of ribosome occupancy (RFP), RNA expression, and translation efficiency (TE) of indicated genesets comparing T cells, B cells, monocytes, and PBMCs: “RNA processing” (H), “chromatin organization” (I), and “regulation of gene transcription” (J). The Wilcoxon rank sum test was used to examine whether the geneset has significantly lower expression in monocytes, compared to both T cells and B cells.

(K) Relative RNA expression levels of ribosomal proteins comparing classical monocytes vs. naive CD8+ T cells, CD4+ T cells, B cells, and PBMCs from four individuals. The TPM expression values were obtained from the NCBI GEO database (accession number: GSE107011). The Wilcoxon rank sum test was used to show monocytes have lower RNA expression of ribosomal proteins than other cell types.