PNAS
www.pnas.org

Supplementary Information for

Records of RNA locations in living yeast revealed through covalent marks

Hugo C. Medina-Muñoz[a], Christopher P. Lapointe[a,b] , Douglas F. Porter[a,c], and Marvin Wickens[a,d]

Hugo Cesar Medina-Munoz
hugo.c.medina@gmail.com

Marvin Wickens
Email: wickens@biochem.wisc.edu

This PDF file includes:

www.pnas.org/cgi/doi/10.1073/pnas.1921408117

SUPPORTING INFORMATION APPENDIX

SUPPLEMENTARY DATASETS

Dataset S1. [In excel file.] RNA species reproducibly recorded by PUP alone (+PAB) across three biological replicates.

Dataset S2. [In excel file.] RNA species reproducibly recorded by PUP alone (-PAB) across three biological replicates.

Dataset S3. [In excel file.] Yeast (BY4742) transcriptome abundance measurements.

Dataset S4. [In excel file.] RNAs selectively enriched at the ER by the ER-anchored PUP (ER-PUP) relative to PUP alone (unanchored PUP).

Dataset S5. [In excel file.] RNAs selectively enriched by PUP alone (unanchored PUP) relative to the ER-anchored PUP (ER-PUP).

Dataset S6. [In excel file.] ER-PUP- or PUP alone-enriched RNAs: GO Terms.

Dataset S7. [In excel file.] mRNAs enriched at the ER by ER proximity-specific ribosome profiling (ER profiling).

Dataset S8. [In excel file.] mRNAs enriched at the outer mitochondrial membrane by mitochondria proximity-specific ribosome profiling (mitochondrial profiling).

Dataset S9. [In excel file.] RNAs selectively enriched at the mitochondria by the mitochondria-anchored PUP (Mito-PUP) relative to PUP alone (unanchored PUP).

Dataset S10. [In excel file.] RNAs selectively enriched by PUP alone (unanchored PUP) relative to the mitochondria-anchored PUP (Mito-PUP).

Dataset S11. [In excel file.] Mitochondria-PUP- or PUP alone-enriched RNAs: GO Terms.

Dataset S12. [In excel file.] 133 mRNAs common to Mitochondria recording and Mitochondrial Profiling: GO Terms.

Dataset S13. [In excel file.] Mitochondria recording Unique vs Mitochondrial Profiling: GO Terms.

Dataset S14. [In excel file.] Mitochondrial Profiling Unique vs Mitochondria recording: GO Terms.

Dataset S15. [In excel file.] Mitochondria-detected RNAs determined to localize to the mitochondria in a Puf3-dependent (Class I) or independent (Class II) manner, and those that are thought to not localize to the mitochondria at all (Class III).

Dataset S16. [In excel file.] ER recording Unique vs Mitochondria recording: GO Terms.

Dataset S17. [In excel file.] Mitochondria recording Unique vs ER recording: GO Terms.

Dataset S18. [In excel file.] ER- and mitochondria-detected (dual-recorded) RNAs by Localized RNA recording and proximity-specific ribosome profiling in yeast, and by APE-Seq in HEK293T cells.

Dataset S19. [In excel file.] Yeast Dual-recorded RNAs: GO Terms.

Dataset S20. [In excel file.] 45 RNAs recorded about equally by both ER- and Mito-PUP.

Dataset S21. [In excel file.] RNA species labeled in proximity of the ERM by APEX-Seq in HEK293T cells.

Dataset S22. [In excel file.] RNA species labeled in proximity of the OMM by APEX-Seq in HEK293T cells.

Dataset S23. [In excel file.] Dual-detection conservation in yeast and HEK293T cells.

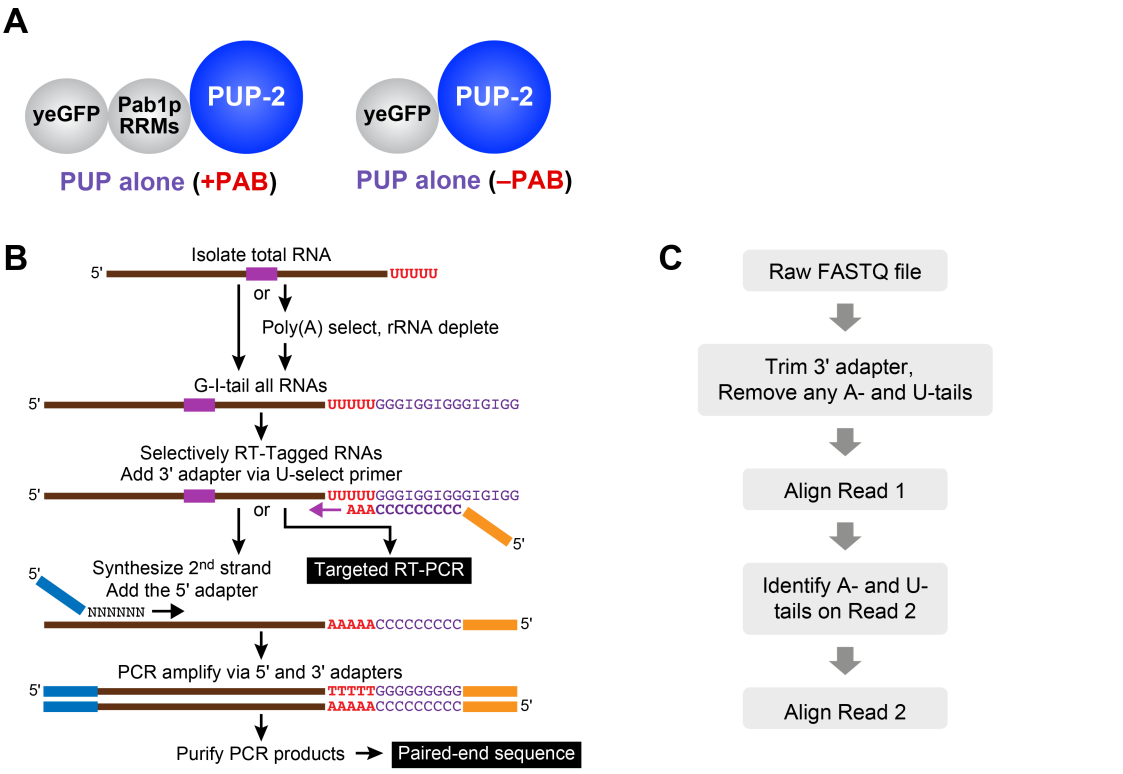
Dataset S24. [In excel file.] Dual-, and ER- or Mito-PUP-recorded complex components.

Dataset S25. [In excel file.] RNAs selectively enriched by Bfr1p-PUP (Bfr1p) relative to PUP alone (unanchored PUP).

Dataset S26. [In excel file.] RNAs selectively enriched by Puf3p-PUP (Puf3p) relative to PUP alone (unanchored PUP).

Dataset S27. [In excel file.] Recorded non-mRNAs.

SUPPLEMENTARY FIGURES



D

PUP alone (+PAB)

Replicate	Total Reads	Uniquely Mapped Reads	Number of Genes Identified
1	26,116,295	4,941,967	5,999
2	25,021,422	3,780,003	5,982
3	21,865,004	4,163,971	5,991

PUP alone (–PAB)

Replicate	Total Reads	Uniquely Mapped Reads	Number of Genes Identified
1	16,143,092	8,368,921	2,090
2	20,430,380	10,462,818	2,208
3	16,713,968	8,274,280	2,080

Fig. S1. Detecting U-recorded RNAs.
A) PUP-2 ("PUP alone") chimeras with or without Pab1p RRMs. **B)** Experimental protocol. From a preparation of total yeast RNA, RNAs with 3' terminal uridines were identified using sequential depletion of rRNAs and oligo(dT) selection, G/I-tailing, reverse transcription with a U-selective primer, and amplification (1). Sequences of the 3' ends and tails were identified by paired-end sequencing (Illumina HiSeq2500). **C)** Computational analysis. Flowchart to identify RNAs recorded *in vivo*. **D)** Representative data. Sequencing statistics with PUP (+PAB) and PUP (-PAB). Experiments were performed with a minimum of three biological replicates and a minimum of two technical replicates for each one.

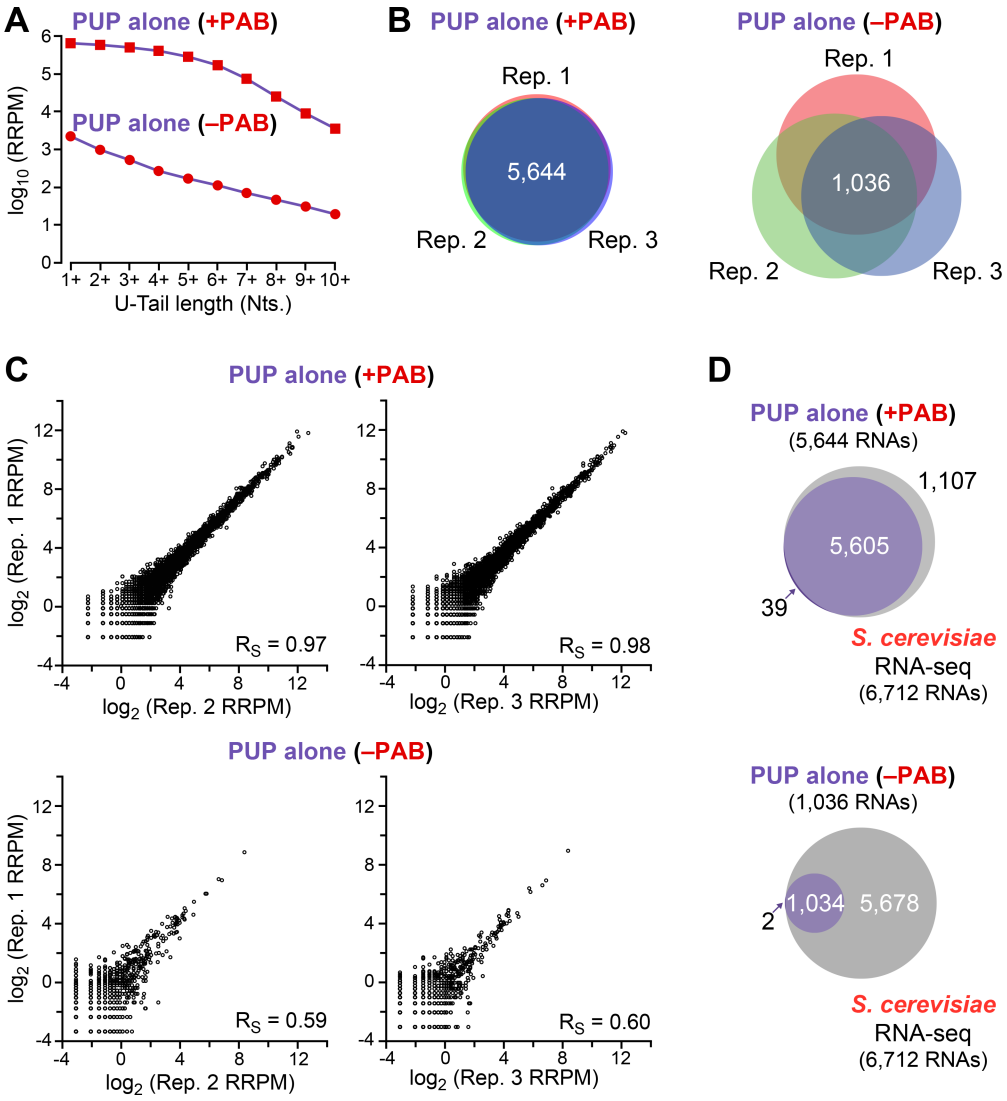


Fig. S2. PUP recording efficiency.
A) U-tail length intervals (1U-10U) and analogous PUP variant U-tail reads (RRPM). **B)** Reproducibility across three PUP alone biological replicates for each variant. **C)** recorded reads per million (RRPM) for both chimeras. Each dotrepresents a single RNA species. **D)** Comparison of the number of RNAs identified with PUP alone, with (top) or without (bottom) PAB vs all yeast RNAs (1).

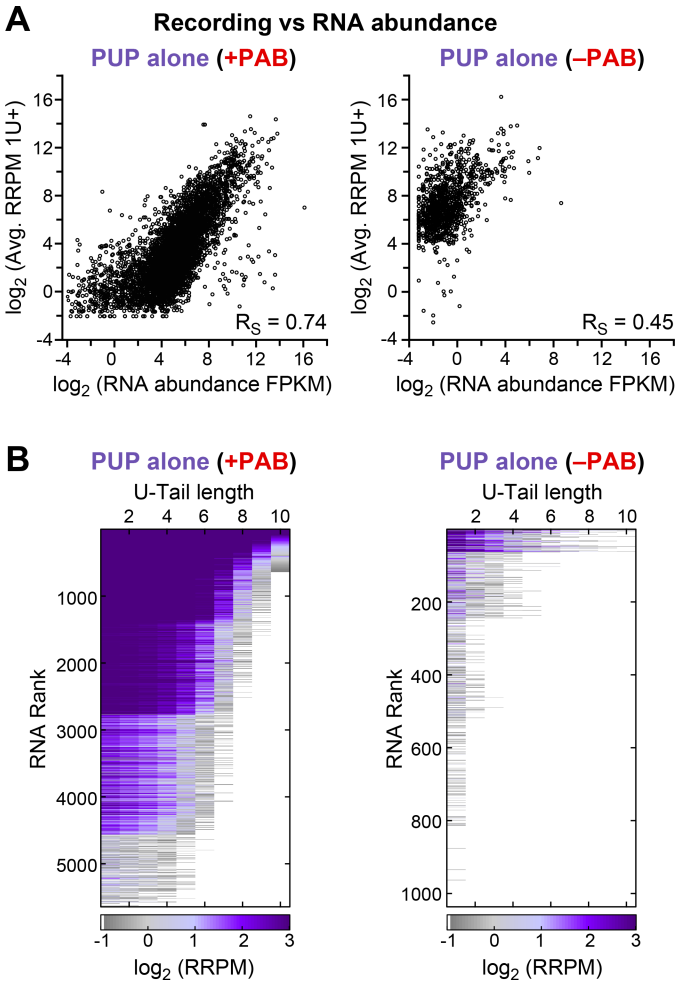
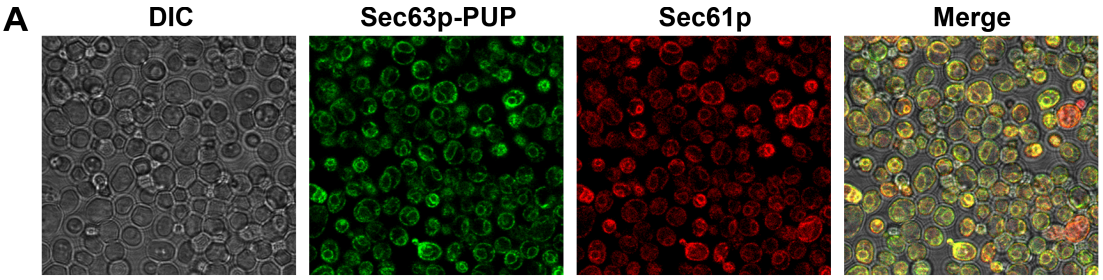


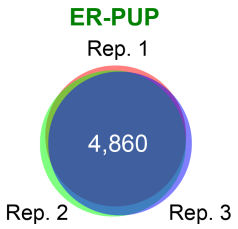
Fig. S3. Broad specificity cytoplasmic RNA recording: a baseline.
A) Recorded reads vs RNA abundance (1) for PUP alone with (left) or without (right) PAB. **B)** Ranked U-recorded RNAs (y-axis) with and without RRMs. Each row indicates an individual RNA species; columns represent U-tail length intervals, from 1 to 10 U's added. Color relates to abundance of reads, as indicated in the key (bottom): purple indicates frequent reads; grey, low reads.

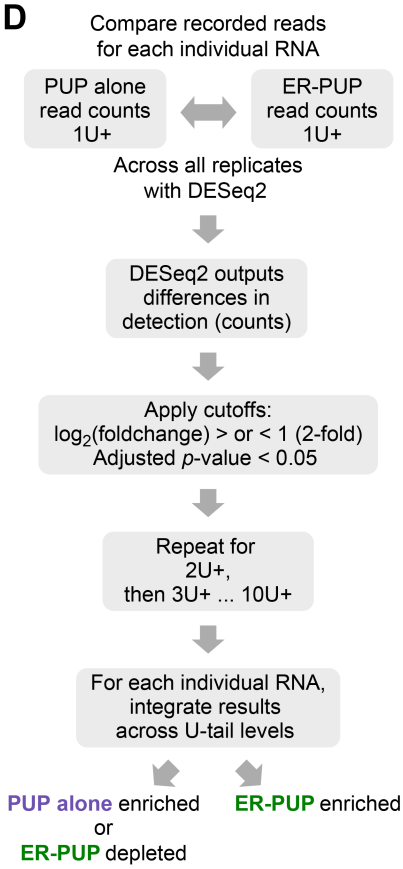


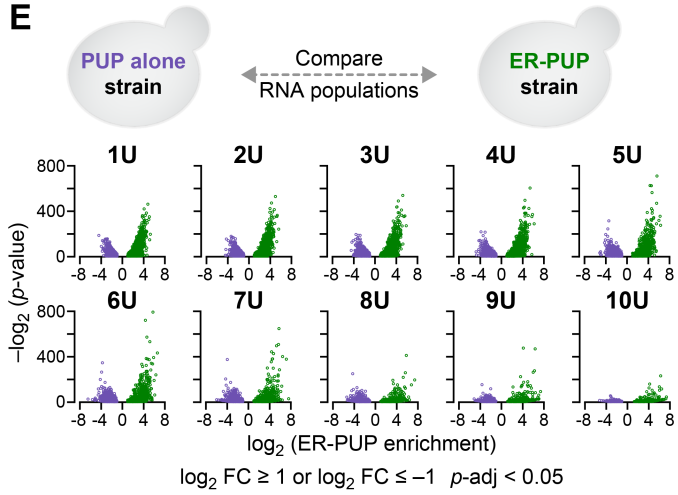
B

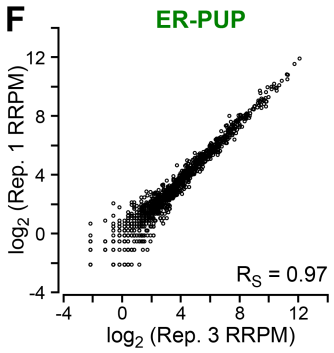
ER-PUP

Replicate	Total Reads	Uniquely Mapped Reads	Number of Genes Identified
1	15,779,019	4,455,578	5,285
2	20,712,699	4,716,985	5,684
3	15,357,796	4,345,423	5,506









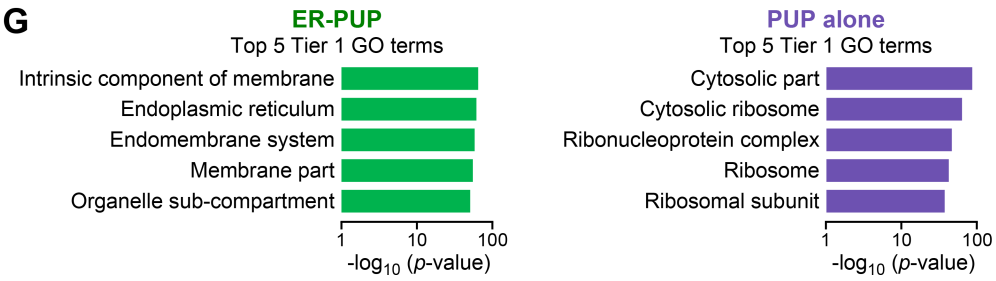


Fig. S4. ER recording and enriched events.
A) Subcellular distribution: ER-PUP GFP fluorescence (green) vs an ER marker (2, 3) in a field of cells. **B)** *Statistics.* ER-PUP sequencing statistics across three biological replicates. **C)** *Reproducibility.* Data as in (B). **D)** *Flowchart of computational analysis.* The steps used to identify RNAs whose recording was enriched with ER-PUP or PUP alone are depicted. DESeq2 (4) was used to identify statistically significant differences (adjusted *p*-value < 0.05), ($\log_2(\Delta$ recorded reads$) \geq 1$)). **E)** Enrichment of individual RNAs vs number of U's added. Differences (x-axis) and significance (y-axis) values distinguish individual RNA species (dots). Each U-tail length (1U-10U) was analyzed separately. RNAs to the right of zero on the x-axis are enriched by ER-PUP (green dots), while the ones on the left are depleted from ER-PUP (aka enriched by PUP alone, purple dots). **F)** *Reproducibility.* Comparison of data from two biological ER-PUP replicates. **G)** *Functional enrichments.* Top five ER-PUP (left) and PUP alone (right) gene ontology (GO) terms (5, 6) are depicted.

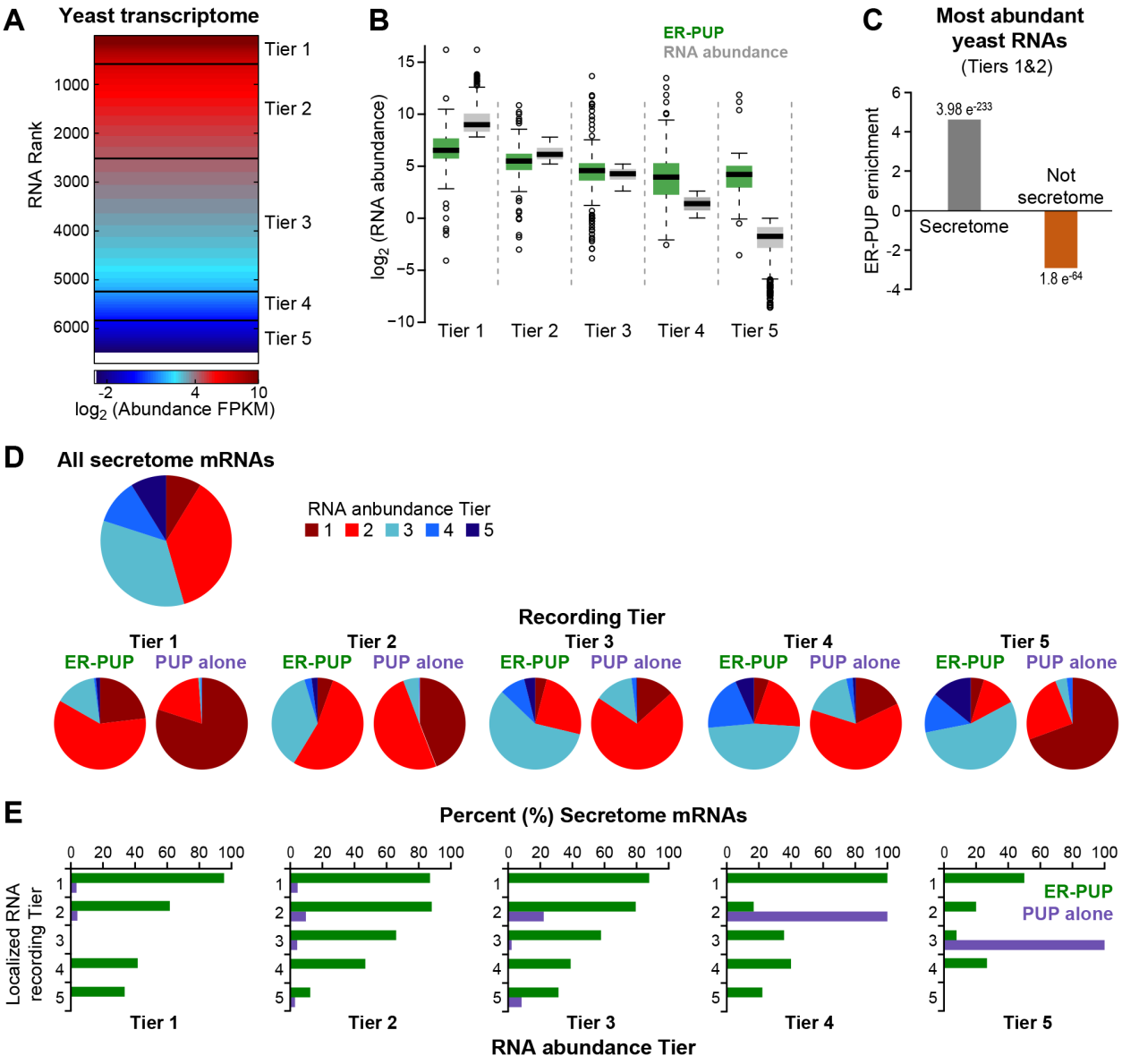


Fig. S5. Analysis of recording and abundance.
A) The yeast transcriptome comprises five RNA abundance (1) (FPKM) tiers (clusters, K-means), ranked from most (Tier 1) to least abundant (Tier 5). **B)** Per tier RNA abundance of ER-enriched RNAs (green) and of all yeast transcripts (grey). **C)** Relationship of ER-PUP enrichment and secretome mRNAs in abundance Tiers 1 & 2. Hypergeometric distribution significance (*p*-values) are reported. **D)** Per Tier ER-PUP and PUP alone abundance composition. **E)** Recorded RNAs populate five distinct RNA abundance bins (bar graphs). For each, the ER-PUP (green) and PUP alone Tiers (purple) (y-axes) project the proportion (%, x-axes) of secretome mRNAs.

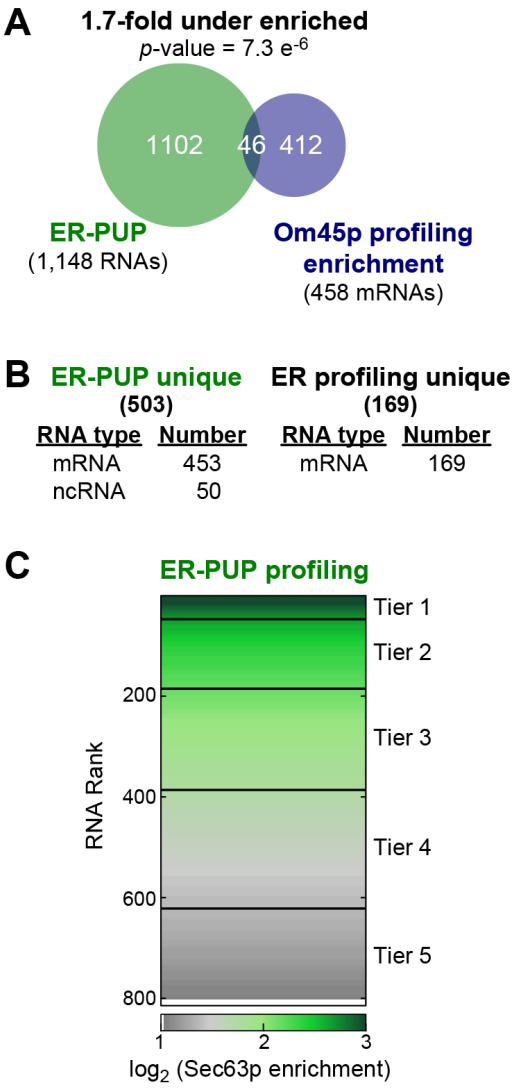


Fig. S6. ER recording and ribosome association.
A) Comparison of ER-recorded RNAs with mRNAs identified by mitochondrial (Om45p) profiling (7). **B)** RNAs identified uniquely by either recording or profiling**. C)** ER (Sec63p) profiling data (8) comprises five enrichment clusters (tiers from K-means), from most (Tier 1) to least ribosome association (Tier 5).

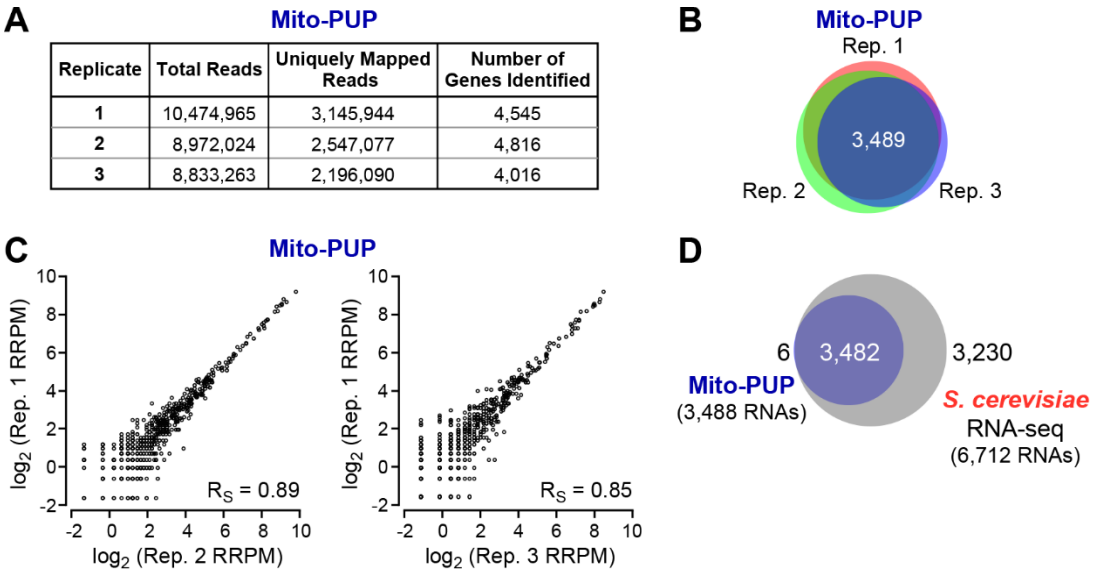


Fig. S7. Mito-PUP recording statistics and reproducibility.
A) *Statistics*. Data from three biological replicates. **B)** RNA species detected reproducibly across three replicates. **C)** Reproducibility of Mito-PUP recording data. Relative abundance of Mito-PUP recorded RNAs (expressed as RRPM) (black dots) in pairs of biological replicates. **D)** Mito-PUP recorded RNAs vs the yeast transcriptome (1).

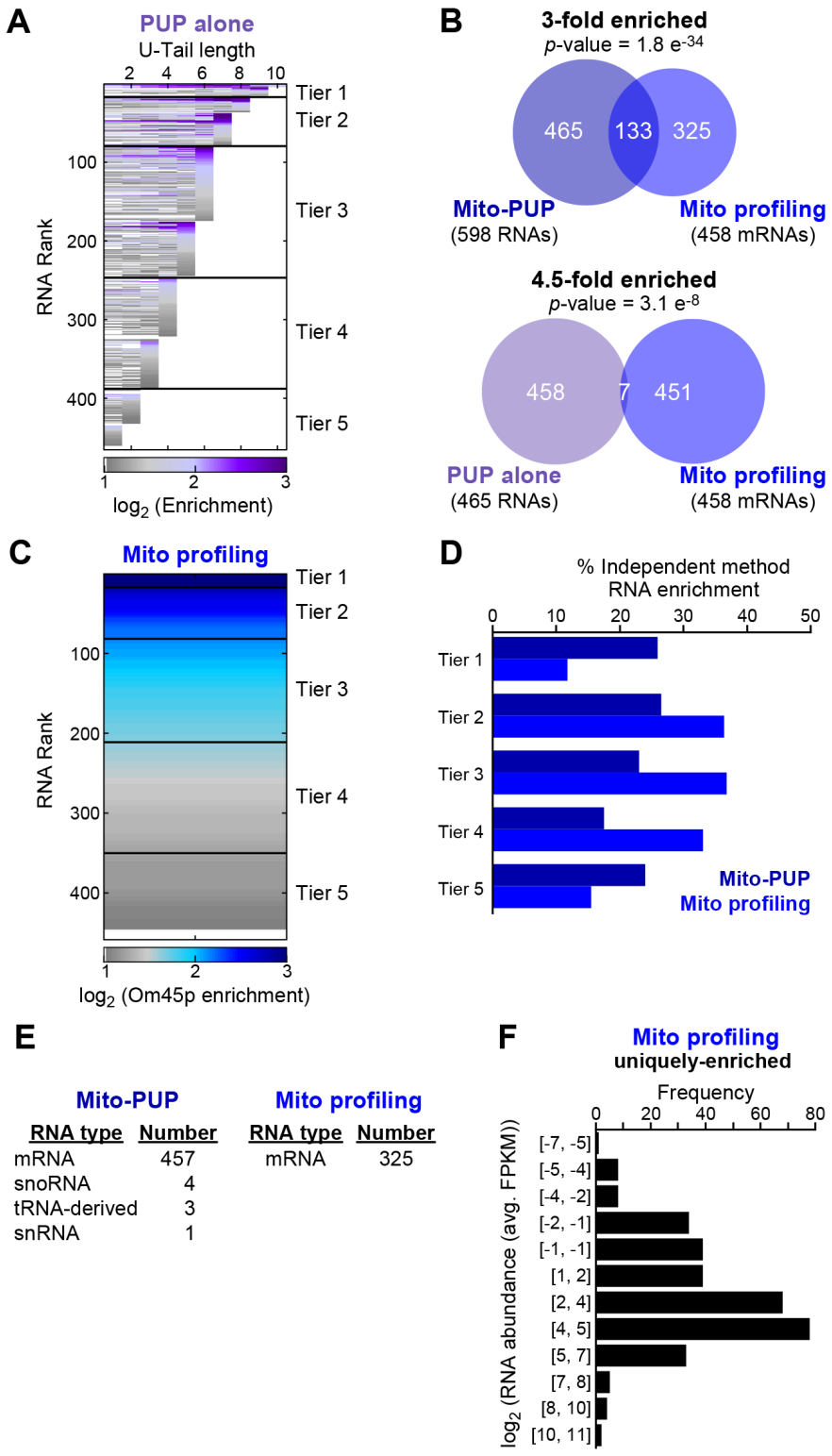


Fig. S8. Mito-PUP recording and ribosome association.
A) PUP alone-enriched RNAs (relative to Mito-PUP) clustered and ranked. Longer U-tails yield a higher rank, from Tier 1 (longest) to Tier 5 (shortest). **B)** RNAs recorded by Mito-PUP (top) or

PUP alone (bottom) vs mitochondria-proximal ribosome associated mRNAs (obtained using Om45p as an anchor (7)). **C)** Mitochondrial ribosome profiling data (7) used to define five (K-means) RNA clusters (Tiers) from the highest (Tier 1) to lowest (Tier 5) association with ribosomes. **D)** RNAs detected by both mitochondrial recording and profiling. For RNAs detected by both recording and profiling, the rank in each tier of the two individual methods is depicted. **E)** RNAs unique to recording and profiling. **F)** Distribution of mRNAs unique to mitochondrial profiling across a series of RNA abundance (1) bins. Each bin (y-axis) represents a range of RNA abundance (log_2(FPKM)) limits. The number of RNAs that fall within those limits is projected by a black bar along the x-axis.

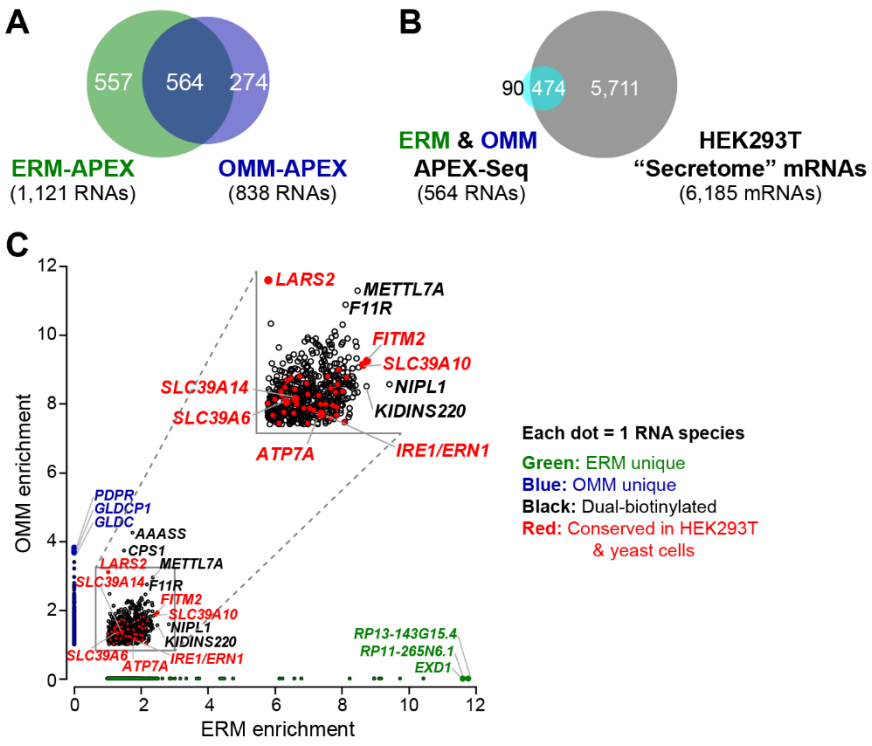


Fig. S9. Conservation of dual-recorded RNA species in yeast and human cultured cells. **A)** RNAs detected by APEX-seq in human (HEK293T) cells at ER membrane (ERM) vs outer mitochondrial membrane (OMM) (data reprocessed from *ref (9)*). **B)** Dual-labeled secretome mRNAs from APEX-Seq. **C)** RNAs identified in recording vs APEX-seq: conservation of RNAs at both ER and mitochondria. Each mRNA is represented by a dot, and plotted vs enrichment in ERM (x-axis) and OMM (y-axis). Organelle-specific mRNAs are green (ERM) or blue (OMM), and lie on the x- or y-axes, respectively. Dual localized mRNAs in HEK293T cells are black. Red indicates RNAs that are identified at both the ER and mitochondria in both yeast and HEK293T cells. Blow-up insert highlights dual-biotinylated RNAs.

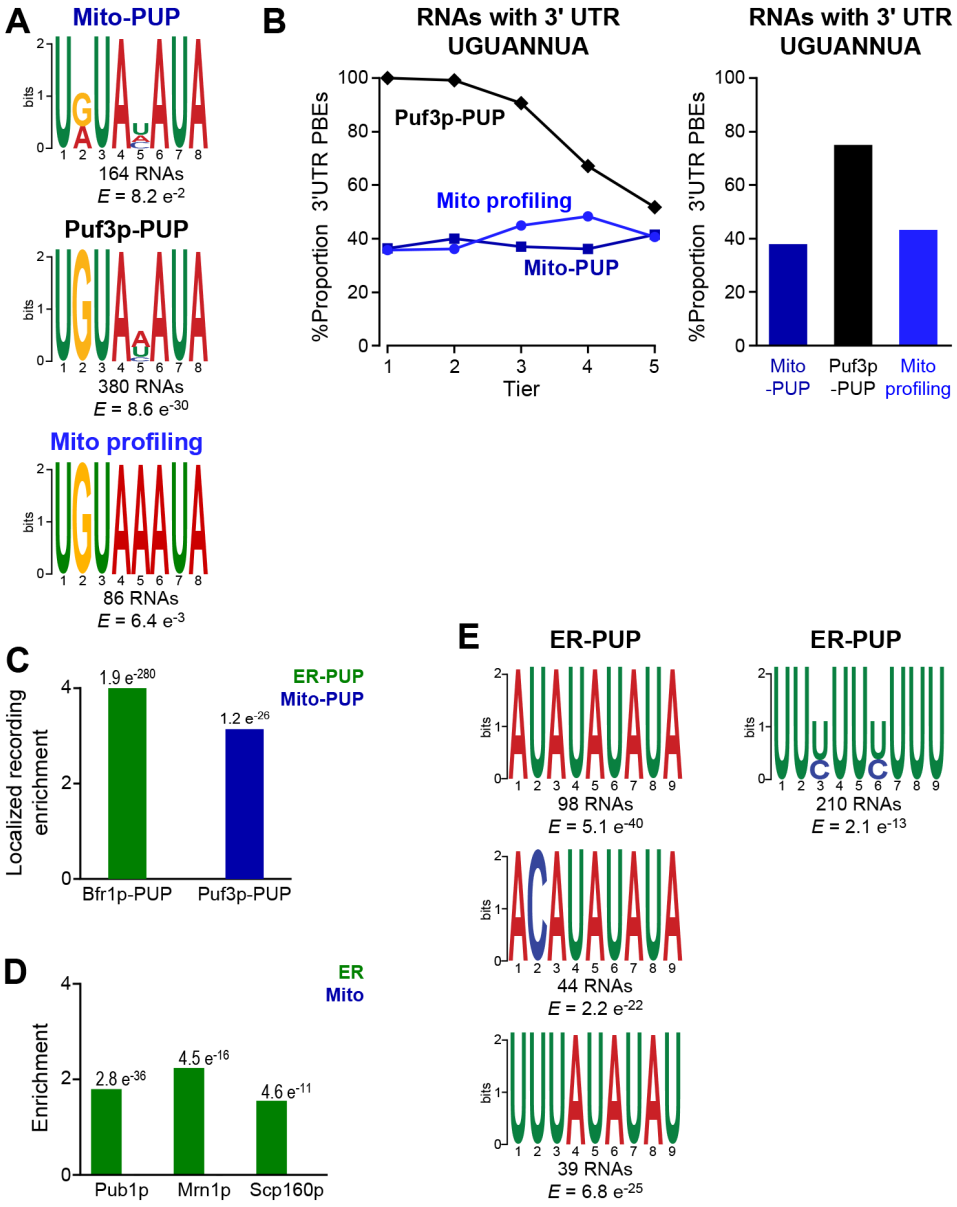


Fig. S10. Sequence elements correlated with ER or mitochondrial localization.
A) Enrichment of motifs associated with mRNAs recorded by Mito-PUP (top), physically associated with Puf3p-PUP (1) (center), or bound by mitochondria-proximal ribosomes inferred from profiling (7) (bottom). **B)** fraction of mRNAs that with one or more Puf3p-binding element (PBE) in their 3' UTR, as a function of their tier in Mito-PUP recording (dark blue), Puf3p-tagging (1) (black) or mitochondrial profiling (7) (light blue) tier Bar graph shows percent (%) of all RNAs detected that contain at least one PBE in the 3' UTR. **C)** RNAs that physically interact with Bfr1p and Puf3p (targets of each RBP) (1) from among ER-PUP- (green) and Mito-PUP-recorded (blue) RNAs. Hypergeometric distribution significance (*p*-values) for enrichment are shown. **D)** ER-PUP enriches additional RNA-binding protein targets (identified in RIP-chip (10)). **E)** Enrichment of motifs associated with mRNAs recorded by ER-PUP.

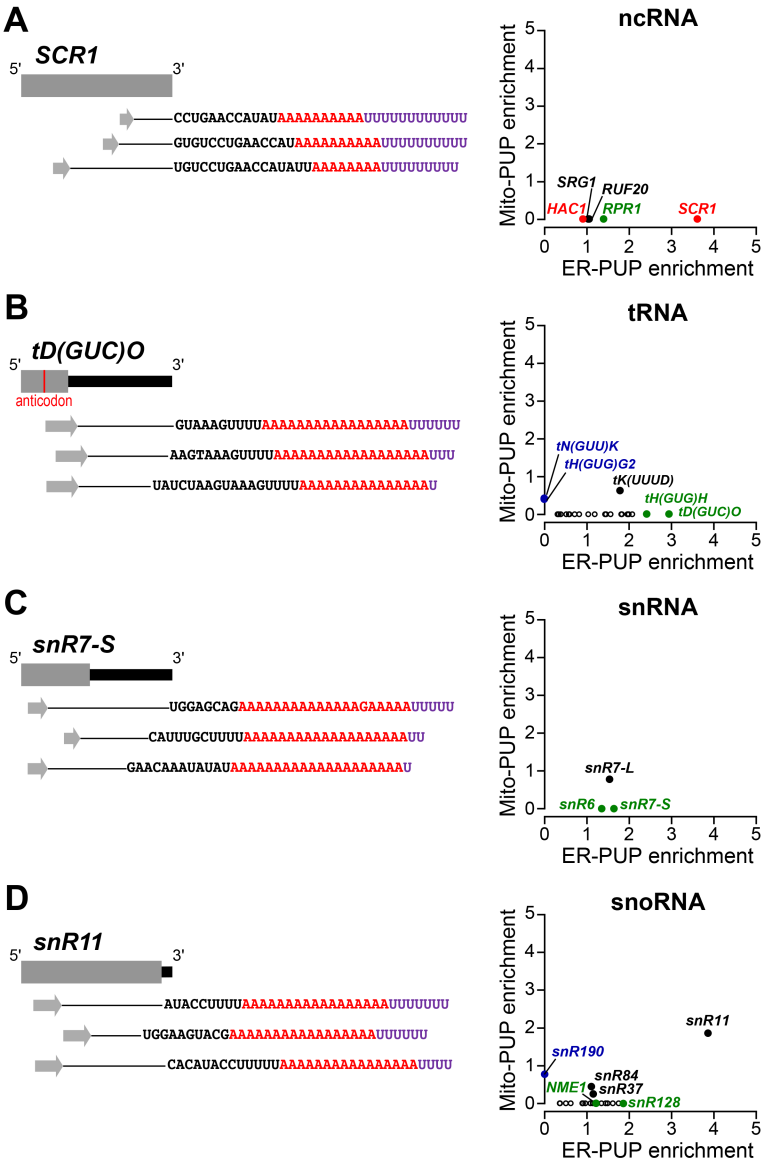


Fig. S11. ncRNAs enriched at the ER and mitochondria.
ER- and Mito-PUP scatter analyses on right demonstrate the enrichment distributions for each of the four classes of non-protein coding RNA genes, which includes RNAs derived from **A)** ncRNAs, **B)** tRNAs **C)** snRNAs, and **D)** snoRNAs. Each RNA is represented by a single dot, and the dots are colored to reflect unique enrichment by ER-PUP (green) or Mito-PUP (blue), or both (black). Larger circles highlight notable RNAs such as the top RNAs for each site or, for the ncRNAs, RNAs characteristic of the ER (red dots). Three RNAs from each class are diagrammed to the left. Forward reads are represented by gray arrows, while black sequences represent the DNA-encoded bit, presumably the 3' end of the RNA, detected in the reverse read. This is followed by the sequence of non-templated adenosines (red A's) and uridines (purple U's) that followed the 3' sequence. The line between these two features represents the section of the RNA inferred from the mapped paired-end reads.

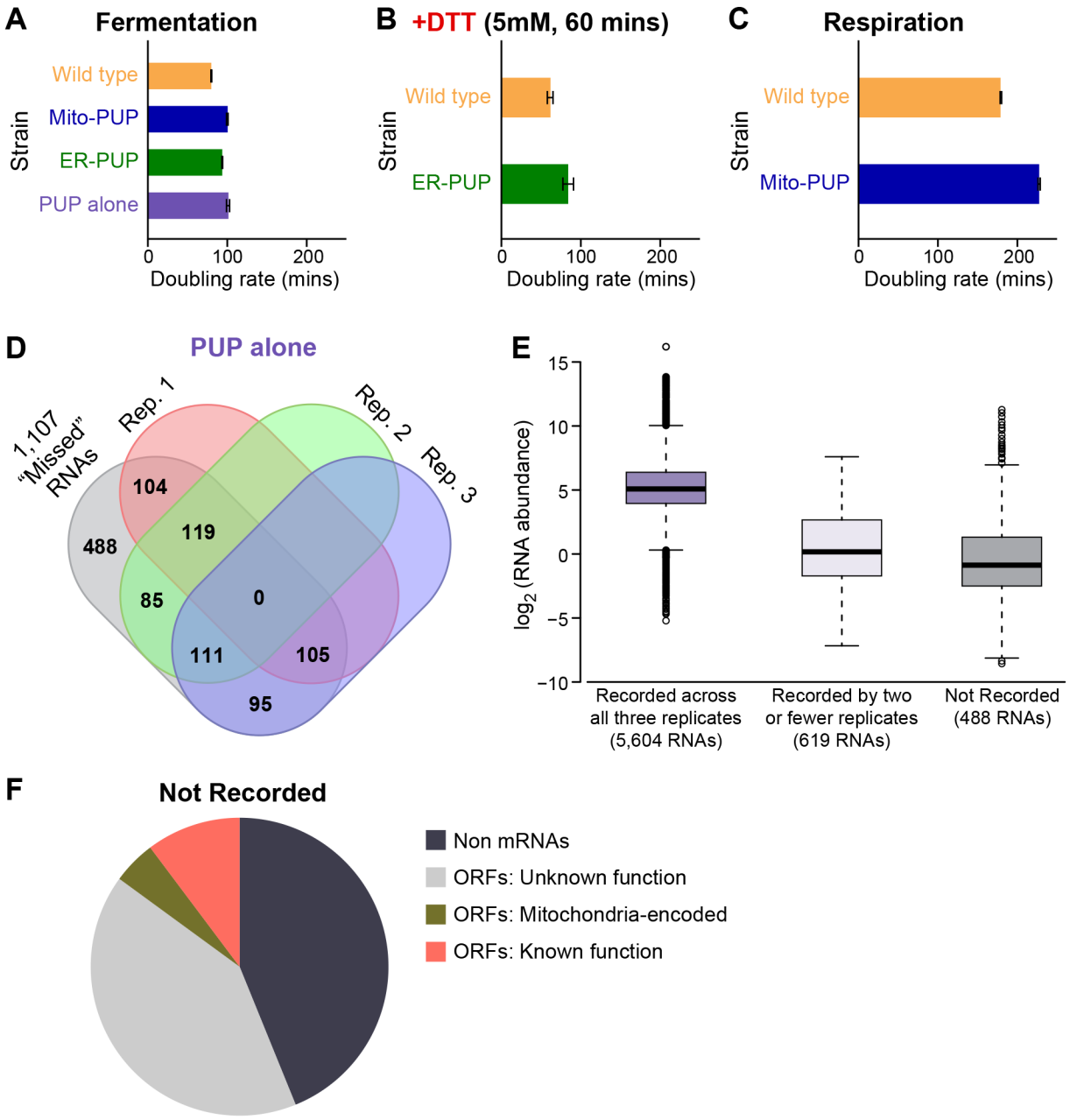


Fig. S12. Recording effects on cellular growth, and detection.
Approximate doubling rates (in minutes) for each strain grown in **A)** Fermentable media, **B)** Fermentable media with 5mM DTT for 60 mins, and **C)** non-fermentable media. All cultures were grown to approximately similar densities (see Methods). **D)** Detection for 1,107 ”Missed” RNA in each PUP alone replicate. **E)** RNA abundance (log_2(Average FPKM)) for RNAs recorded by PUP alone across three replicates (purple), two or fewer replicates (light purple), or not recorded at all (grey). **F)** Types of RNAs not recorded by PUP alone.

SUPPLEMENTARY METHODS

RNA extraction and library preparation

Total RNAs and libraries were prepared in accordance with previous reports (1).

Growing cells for RNA extraction

All yeast strains were grown to mid-log phase (OD_{660} ~ 0.5-1) in 50 mL of YPAD at 30°C and with orbital shaking (180 RPM). The cultures were then transferred to 50 mL conical tubes and placed on ice for 5 minutes. After, the cells were pelleted with centrifugation at 3,200 RPM and 4°C for 5 minutes. The media was decanted, and the pellet was resuspended in 50 mL of ice-cold and sterile deionized water, followed by centrifugation at 4°C for 5 minutes at 3,200 RPM. The supernatant was then decanted, and RNA extracted from the pellets.

Total RNA extraction

The washed Yeast pellets were resuspended in 500 µL of RNA ISO buffer (0.2 M Tris-HCl pH 7.5 0.5 M NaCl, 0.01 M EDTA, 1% (v/v) SDS). Acid-washed beads (425-600 µM; ~200 µL) were added to each tube, followed by the addition of 500 µL of Phenol:Chloroform: Isoamyl Alcohol ("PCA" at 25:24:1 and pH 6.7, ThermoFisher Scientific BP1752). The tubes were then loaded onto a tube holder and vortexed for 20 s at room temperature followed by 20 s on ice. Both steps repeated for a total of ten cycles. After, the lysates were each divided into two new 1.5 mL microfuge tubes, and 375 µL PCA and 375 µL of ISO buffer were added to each. The samples were then mixed by gentle inversion and spun at 4°C at ≥ 13,200 RPM for 10 minutes. The aqueous phase was transferred to new tubes, an equal volume of PCA added and mixed by gentle inversion, and spun at 4°C at ≥ 13,200 RPM for 10 minutes. Again, the aqueous phase was transferred to new tubes, but this time an equal volume of chloroform was added. The samples were mixed by gentle inversion and spun at 4°C at ≥ 13,200 RPM for 10 minutes. The aqueous phase was again transferred to new tubes with 1 mL of ice-cold 100% EtOH, mixed gently by inversion, and incubated overnight at -80°C. The next morning, the samples were spun at 4°C at ≥ 13,200 RPM for 20 minutes to pellet the nucleic acids. The pellets were washed twice with ice-cold 80% ethanol and spun at 4°C at ≥ 13,200 RPM for 20 minutes in between each wash. The pellets were then resuspended in 43 µL of nuclease-free water, and the split samples recombined into a single 1.5 mL tube with 10 µL of 10X TURBO DNase Buffer and 4 µL (8 U) of TURBO DNase (ThermoFisher Scientific, AM1907M). The samples were then incubated at 37°C for 60

minutes, and the RNA was isolated using the GeneJET RNA Purification kit with the "RNA cleanup protocol" (ThermoFisher Scientific, K0731). The RNA was eluted from the columns with 30 µL of nuclease-free water. 1:10 dilutions of each sample were prepared to determine RNA concentration by spectrophotometer, and the RNA was stored at –80°C.

Library preparation

Poly(A)-enrichment

Starting with 50 µg of total RNA, the samples were subjected to poly(A) selection with the Dynabeads mRNA purification kit (ThermoFisher Scientific, cat. # 61006). The poly(A)-enriched RNAs were then subjected to rRNA depletion using the Ribo–Zero Magnetic Gold kit for yeast (Epicentre/Illumina, cat. # MRZY1324). The samples were then cleaned using Agencourt RNAclean XP beads (Beckman Coulter, A63987). The Agencourt beads were allowed to come to room temperature and were mixed by vortexing before use. 160 µL of beads were added to 85-90 µL of RNA from the previous steps with immediate pipette mixing (10 times) and gentle vortexing. The samples were then allowed to sit at RT for 15 minutes and were then placed on a magnetic stand for 5 minutes. The supernatant was removed, and 400 µL of freshly-prepared 80% EtOH was added to the samples while on the magnetic stand and allowed to sit for 1 minute. This step was repeated for a total of two times, and EtOH was removed after the last step, and the beads allowed to air dry for 5 minutes while still on the magnetic stand. The samples were then removed from the stand and RNA eluted from the beads with the addition of 12 µL of nuclease-free water and thorough pipette mixing. The samples were then placed on the magnetic stand for 5 minutes, and 10 µL of elution removed to a new tube, making sure not to carry over any beads.

GI-Tailing

The poly(A)-enriched RNA was first subjected to an enzymatic reaction to add guanosine (G) and inosine (I) to the 3' ends (GI-tailing). These GI tails protect the 3' uridine tails and serve as a binding site for a custom RT primer (see RT step below). To 10 µL of poly(A)-enriched RNA, 4 µL of yeast poly(A) polymerase (PAP) reaction buffer (Affymetrix, 74225Y), 1 µL of 10 mM GTP (Epicentre/Illumina, ASF3507), 1 µL of 3.3 mM ITP (Sigma-ALDRICH, I0879/D0758), and 2 µL of nuclease-free water were added (total of 18 µL per reaction). 2 uL of 600 U/µL yeast PAP (Affymetrix, 74225Y) was then added to each reaction. The tubes were incubated at 37°C for 120 minutes, with an additional 2 µL of 600 U/µL yeast PAP added 90 minutes into the reaction. Once the reaction was complete, the samples were combined with 80 µL of nuclease-free water and

100 µL of Phenol:Chloroform: Isoamyl Alcohol (“PCA” at 25:24:1 and pH 6.7) in a 1.5 microcentrifuge tube, mixed by gentle inversion, and spun for at 4°C at ≥ 13,200 RPM for 5 minutes. The aqueous phase was then transferred to a new 1.5 mL microfuge tube with an equal volume of chloroform. The samples were mixed by gentle inversion and spun at 4°C at ≥ 13,200 RPM for 5 minutes. The aqueous phase was again transferred to a new 1.5 mL microfuge tube and mixed with 500 uL of ice-cold 100% EtOH, 10 µL of 3 M NaOAc, and 1 µL of 15 mg/mL GlycoBlue (ThermoFisher Scientific AM9515). The samples were then mixed by vortexing and incubated at –80°C overnight. The next morning, samples were spun at 4°C at ≥ 13,200 RPM for 25 minutes. The pellets were washed with 500 µL of ice-cold 80% EtOH for a total of two times and spun at 4°C at ≥ 13,200 RPM for 25 minutes after each wash. The pellets were then resuspended in 10 µL of nuclease-free water.

cDNA synthesis

The GI-tailed RNA samples were reverse-transcribed using a custom RT primer that selects for GI-tailed RNAs that contain at least three uridines (5’–GCCTTGGCACCCGAGAATTCCACCCCCCCCCAAA –3’). To generate the RNA and primer mix, 5 µL of GI-tailed RNA, 1 µL of 1 µM RT primer, 1 µL of 10 mM dNTPs (Promega U1515), and 6 µL of nuclease-free water were mixed in the same PCR tube. The buffer mix was generated in a separate PCR tube with 4 µL of 5X SuperScript III Reaction Buffer (ThermoFisher Scientific 18080), 1 µL of 100 mM DTT, and 1 µL of RNAseOUT (ThermoFisher Scientific 10777). Master mixes were generated when possible to streamline the sample preparation. The reactions were then heated to 65°C for 5 minutes and then cooled to 50°C for 5 minutes on a thermocycler. After, the RNA and primer mix and the buffer mix were combined, and 1 µL of 200 U/µL SuperScript III reverse transcriptase (ThermoFisher Scientific 18080) was added. The RT reactions were incubated at 50°C for 60 minutes, after which the enzyme was heat-inactivated at 85°C for 5 minutes. The reactions were then cooled to 4°C, and 1 µL RNase H (ThermoFisher Scientific 18021) was added and subsequently incubated at 37°C for 20 minutes. 80 µL of nuclease-free water was then added to adjust the total volume to ~100 µL, and the cDNA was purified using the GeneJET PCR Purification kit without isopropanol. The cDNA was eluted twice by adding 32 µL of nuclease-free water directly to the column membrane and incubating the column for 2 minutes each time before spinning at ≥ 13,200 RPM for 2 minutes.

Second Strand Synthesis

To 60 µL of purified cDNA, 12 µL of nuclease-free water, 10 µL of 10X Klenow Buffer (prepared in-house; 500 mM Tris–HCl pH 7.5, 100 mM $MgCl_2$, 10 mM DTT, and 0.5 mg/mL BSA), 5 µL 10 mM dNTPs, and 10 µL of 10 µM 2nd strand synthesis primer (5'–GTTCAGAGTTCTACAGTCCGACGATCNNNNNN–3', where $(N)_6$ = random hexamers) were added. 3 µL of 5 U/µL Exo-Klenow fragment DNA polymerase I (ThermoFisher Scientific AM2008) was then added, and the reaction incubated at 37°C for 30 minutes. The reactions were then cooled to room temperature and cleaned using well-mixed and room temperature AMPure RNAClean XP beads. 100 µL beads (1:1 ratio of beads to reaction) were added to each sample with thorough pipette mixing (10 times) and vortexing, and the reactions were then incubated for 15 minutes at room temperature. After, the samples were placed on a magnetic stand for 5 minutes. The supernatant was removed, and 400 µL of freshly-made room temperature 80% EtOH was added to the beads and allowed to sit for 1 min. This was repeated for a total of two washes. After the EtOH was removed and beads allowed to air-dry for 5 minutes at room temperature, 100 µL of nuclease-free water was added to the samples and immediately mixed by pipetting (10 times) and vortex. The samples were then incubated at room temperature for 2 minutes and placed on the magnetic stand for 5 minutes. 98 µL of the eluate was then moved to a fresh 1.5 mL microfuge tube. The AMPure RNAClean XP bead wash was repeated, but the samples eluted with 50 µL of nuclease-free water the second time.

PCR-amplification

50 µL of the eluate was mixed with 83.3 µL of 2X GoTaq Green Master Mix (Promega M712), 6.7 µL of 10 µM 3' Barcoded PCR Primer (5'–CAAGCAGAAGACGGCATACGAGATNNNNNNGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA –3', where $(N)_6$ = a unique barcode for each sample) and 6.7 µL of 10 µM 5' PCR primer (5'– AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGA –3'), 20 µL of nuclease-free water for PCR amplification and barcoding. Each sample was then divided into 20 µL into eight separate PCR tubes, and amplified via the thermocycler program below.

Amplification protocol

1. 94°C	2 minutes
2. 94°C	10 s
3. 40°C	2 minutes
4. 72°C	1 minute
5. Go to step 2	1X (once)
6. 94°C	10 s

7. 55°C	30 s
8. 72°C	1 minute
9. Go to step 6	7X (7 times)
10. 94°C	15 s
11. 55°C	30 s
12. 72°C	1 minute
13. Go to step 10	14X (14 times)
14. 72°C	5 minutes
15. 4°C	forever

After the reaction completed, the samples were consolidated back into one 1.5 mL microfuge tube (1 per sample), and cleaned using two tandem AMPure RNAClean XP beads washes (room temperature and well-mixed beads) as described above. However, both washes used a 5:4 volume ratio of reaction to beads, and the first wash eluted the sample in 100 µL of nuclease-free water while the second eluted in 15 µL. The samples were then submitted to the UW-Madison biotechnology center for paired-end sequencing (2X50 bp) on the HiSeq2500 platform. Libraries were loaded at equal concentrations, with PhiX loaded at 30% of the total concentration.

Mapping reads to the genome

FastQ files were processed with a custom sequencing pipeline (1). Reads from libraries that were sequenced on multiple lanes were pooled and processed together.

Processing read counts

Read counts were processed differently to determine either (A) absolute RNA recording values or (B) recording enrichment at a given site.

A) Calculation of recorded reads per million across ten minimum U-tail lengths.

RNAs species that were recorded reproducibly across three replicates were identified with BioVenn (11) (http://www.biovenn.nl/). The Read counts in each replicate were then normalized to the total number of reads (in millions) for that replicate, yielding the recorded reads per million (RRPM) across ten minimum U-tail lengths (1U to 10U) for each recorded RNA species. The RRPM values for each RNA species in each replicate were then averaged to yield the average RRPM (RRPM Avg.) for each RNA species at each of the U-tail length levels. RNAs species were then ranked by U-tail length and reads, and RNAs with the longest U-tail length reads, but less total reads having priority over those with more reads but shorter U-tail length.

B) Deseq2 analyses to determine enrichment at the ER and Mitochondria.

Raw read counts for RNAs detected in each experimental replicate were analyzed with Deseq2 (4). For each individual analysis (ER- or Mitochondria-Localized RNA Recording data), the PUP alone (+PAB) read counts were used as a control. In all our analyses, a $\log_2$ fold change of two-fold or greater (($\log_2$(FC ≥ +1) or ($\log_2$(FC ≤ -1)) and an adjusted *p*-value (*p*-adj) cutoff of < 0.05 was applied. RNAs that met these criteria were considered "enriched (i.e., “ER-enriched”) or depleted (i.e., “ER-depleted” or PUP “alone-enriched”).

Note: To reduce false positives, DeSeq2 applies the median of ratios strategy to normalize the raw recorded read (RR) count values (4, 12). The program also filters out RNA species with lower quality data (low and/or highly variable RRPM values between replicates). For those that pass the filter, the program penalizes RNA species with lower data in the fold-change estimates such that the resulting exaggerated fold-difference values are reduced (4). Further, RNA species with less data (lower RRPM values) are assigned a lower *p*-value, and many are eliminated via *p*-value (< 0.05) and fold-change ($\log_2$ (fold change) > 1)) cutoffs.

K-means clustering to generate tiers and heatmaps

Clustering was done using Cluster 3.0 (13) (C clustering Library 1.52) using the K-means option, as previously reported (1). Five groups (Tiers) were arbitrarily selected, and $\log_2$-transformed data were clustered based on detection efficiency (i.e., L_2FC). Tiers were ranked based on highest to lowest efficiency of detection (i.e., most enriched to least enriched), with priority given to enrichment at the highest U-tail lengths. Heat maps were generated with MatLab.

Doubling rate estimates

Three biological replicates per strain, per experiment, were grown overnight in 5mL of Yeast Extract–Peptone (YP) with adenine (A) and either dextrose (YPAD) or glycerol (YPAG; 3% v/v) at 30°C with shaking. Cells from these cultures were seeded into 50mL of the corresponding media in 250 mL flasks, and grown at 30°C and 180 RPM. For fermentation and respiration experiments, cells were grown to mid-log phase (OD_{660} = 0.5-1), and optical density values were recorded at time of harvest to calculate an approximate doubling rate for each strain. For DTT exposure experiments, cells were first grown to an OD_{660} = ~0.5 in YPAD media, and then DTT was added to a final concentration of 5mM. Cells were grown at 30°C at 180 RPM for 60 mins. After the exposure, OD_{660} measurements were taken, and approximate doubling rates calculated from the difference from time of first exposure to 60 mins after DTT addition. Approximate doubling times were calculated with the following eqaution d=[1/[log(n/k)/log2]*1/T], where d = approximate doubling time, K = quantity of yeast (measured as OD_{660}/mL) inoculated at start of growth, T = time lapsed between seeding and harvesting the yeast, and n = quantity of yeast (measured as

OD_{660}/mL) at time of harvest. All yeast were allowed to grow to approximately similar densities (OD_{660}/mL).

Data Mining

Proximity-specific ribosome profiling. Proximity-specific ribosome profiling data were mined published experiments (7, 8). mRNAs with greater than or equal to two-fold or higher ribosome-protected fragment reads relative to the input were determined to be "enriched". Since we omitted did not use cycloheximide in our experiments, we selected Sec63p (8) (1 min biotin pulse) and Om45p (7) (2 min biotin pulse) mediated profiling data that omitted the translation inhibitor.

Mitochondrial proteome mRNAs. Mitochondria-copurified proteins from two experiments (14, 15) were consolidated into a single list termed, "mitochondrial proteins".

RBP targets. RNA that interact with RNA-binding proteins were mined from published RNA tagging (1) and RIP-Chip (10) data. RNA tagging data were reprocessed using the DESeq2 (4) approach described here for ER-Pup and Mito-PUP, also using PUP alone (+PAB) as a control set. RNAs that were tagged $\log_2(\Delta$ Tagged Reads) $\geq 1))$ and significance (p-adj) < 0.05 relative to PUP alone (+PAB) were considered enriched. These analyses were done across ten minimum U-tag lengths, and the RNAs were ranked by highest to lowest U-tag length enrichment. For RIP-Chip, RNAs that had a Significance Analysis of Microarrays (SAM) q-value < 10% were considered targets. The classes of Puf3p mRNA targets were retrieved from published biochemical experiments (16).

APEX-Seq. APEX-Seq data (9) was used for these analyses. Similar cutoffs applied to our data ($\log_2$(Fold Change) $\geq +1$, p-adj < 0.05) were applied to ER (ERM) and mitochondrial (OMM) APEX-Seq data to facilitate comparison.

Dual-recorded RNA conservation. Human homologs of dual-recorded yeast genes were retrieved using YeastMine (5, 6) (https://yeastmine.yeastgenome.org/yeastmine/begin.do).

3' UTR motif enrichment analyses

Command line MEME (17) Version 5.0.4 was used for all analyses. The order (Rank) of all yeast 3' UTRs from (1) was randomized using the excel rand (=RAND()) function, and the resulting list was used as background (-neg) for all MEME analyses. Motifs that were enriched in the 3 UTRs of recorded RNAs (ranked by longest U-tail length) were done using the differential enrichment (-objfun de) function, and MEME was prometed to return the top ten motifs (-nmotifs

10) that range from 8-10 nts in length (-minw 8 -maxw 12) with all lengths between those limits included in the scan (-allw). The rank of the 3' UTR within the list of recorded RNAs or the list of 3' UTRs with randomized rank was taken into account (-norand) in the analyses.
Puf3p binding element incidence in the 3' UTRs of RNAs was determined using a custom Perl script (18).

Tools used

Gene ontology (GO). All GO analyses were done using yeast mine lists (5, 6) (https://yeastmine.yeastgenome.org/yeastmine/bag.do). The analyses used the default background, and considered enrichments with a maximum *p*-value of 0.05 after Holm-Bonferroni correction.

Venn Diagrams. Venn diagrams were generated using BioVenn (11) (http://www.biovenn.nl).

Hypergeometric Distribution Analyses. Hypergeometric distribution calculations were done with the online calculator available from the Graeber Lab (https://systems.crump.ucla.edu/hypergeometric/). The total number of yeast transcripts used was 6,712 RNAs, as defined by RNA-seq (1).

Cumulative Fraction Plots. Cumulative fraction plots were done in either RStudio (Fig. 3 D and E, and 4F) (as previously reported (1)) or Excel (this report, Fig. 5D).

Tab file conversion to Fasta. Tab files that contained 3' UTR sequences were converted to FASTA format using the HIV sequence database Format Converter (https://www.hiv.lanl.gov/content/sequence/FORMAT_CONVERSION/form.html)

Confocal Microscopy Strains were grown to mid-log (OD_{660} 0.5-1) phase in 25 mL cultures at 30 °C in a horizontal shaker at 180 RPM. Cells were grown in synthetic complete (19) (SC) Low fluorescence (LOFLO) media (20).

Yeast cells were immobilized on Concanavalin A (ConA Sigma: 11028-71-0) coated coverslips for imaging with a modified version of a published imaging protocol (21).

Coverslip Preparation

Coverslips were first incubated in a methanol/hydrochloric acid (1:1) solution in 50 mL conical tube in a fume hood overnight. The next day, coverslips were rinsed 3Xs with deionized water, and then placed on one edge inside 65 °C oven until fully dry (~ 30 mins). The dry coverslips were then cooled to RT, and 400 uL of 2 mg/mL concanavalin A (ConA) was spread evenly in the center. After drying for 60 mins at RT, the coverslips were then tilted on their side to remove excess ConA. The coverslips were then covered and dried at RT overnight.

Fixation

Mid-log phase cultures were spun down, and resuspended in 500 uL of fresh SC LOFLO media. 100 uL of cells were then placed evenly on the center of the ConA-coated coverslip and allowed to bind for 30 mins in a 30 °C incubator without shaking. Once immobilized on the coverslip, the cells were fixed by submerging the coverslip in SC LOFLO (5% Formaldehyde, Fisher: BP531-500) for 20 minutes. After, the coverslip was washed with 3Xs with fresh SC LOFLO media, and the coverslip was then placed on a slide containing SC LOFLO.

Confocal settings

Cells were imaged on a Leica TCS SP8 confocal microscope using LAS X 3.1.1.15751. The microscope is equipped with a Photomultiplier (PMT) and Hybrid detectors (HyD). A 63x 1.4NA HC Plan Apochromat oil immersion objective was used with a 3.01 zoom and standard scanner with 400Hz scanning speed. Z-stacks with a 0.3 uM step size were collected, and yeGFP (495nm - 530nm) and mCherry (600nm - 650nm) were sequentially imaged.

Fluorescence quantitation

Images were processed using FIJI(22)/ImageJ(23) Version: 2.0.0-rc-69/1.52i. Noise reduction was done on the images using the “Despeckle” function, and the contrast adjusted for each individual image. Fluorescence intensities (gray values) quantified using the measure (Ctrl+M) function along a straight line that was drawn across individual cells. The measurements were then normalized for each individual sample to get the Normalized Gray Value at each individual measurement. This was done with the following equation, Normalized Gray Value at a specific point on the line = (Raw intensity value at a specific point on the line – Minimum of all intensities on the line)/(Maximum of all intensities on the line – Minimum of all intensities on the line). The normalized gray values for each sample were plotted on a line graph using Microsoft Excel.

Yeast Strains (BY4742 background)

Strain	Description	Construct	Auxotrophic Marker
yHCMM1	Wild type	BY4742	-
yHCMM2	PUP (+PAB)	PSEC63_yeGFP_ScPAB14RRMs_CePUP-2_URA3_TADH1 inserted into URA3 locus	URA
yHCMM3	PUP (-PAB)	PSEC63_yeGFP_CePUP-2_URA3_TADH1 inserted into URA3 locus	URA
yHCMM4	ER-PUP	yeGFP_ScPAB14RRMs_CePUP-2_URA3 inserted downstream of SEC63 CDS	URA
yHCMM5	Mito-PUP	yeGFP_ScPAB14RRMs_CePUP-2_URA3 inserted downstream of OM45 CDS	URA
yHCMM6	ER-PUP & ER marker	mCherry_LEU2 inserted downstream of SEC61 CDS in ER-PUP strain	URA, LEU
yHCMM7	Mito-PUP & Mito marker	mCherry_LEU2 inserted downstream of TOM70 CDS in Mito-PUP strain	URA, LEU

Plasmids

Plasmid	Construct	Auxotrophic Marker
pHCMM1	PTEF1_yeGFP_*Sc*PAB14RRMs_*Ce*PUP-2_URA3_TADH1	URA, LEU
pHCMM2	PSEC63_yeGFP_*Sc*PAB14RRMs_*Ce*PUP-2_URA3_TADH1	URA, LEU
pHCMM3	PSEC63_yeGFP_CePUP-2_URA3_TADH1	URA, LEU
pHCMM4	PCYC1_SEC61_mCherry_LEU2_TCYC1	LEU

The plasmids will be available through addgene, and the yeast strains distributed.

Primers

Makes	Input	Primers	Primer Sequence
yHCMM2	pHCMM2	primHCMM1	ATTTATGGTGAAGGATAAGTTTTGACCATCAAAG AAGGTTAATGTGGCTGTGGTTTCAGGGTCCATA AAGCTGGTGCTCTCATGCTTCTT
		primHCMM2	GGCCATGAAGCTTTTTCTTTCCAATTTTTTTTTTT TCGTCATTATAGAAATCATTACGACCGAGATTCC CGGGCGGTAGAGGTGTGGTCA
yHCMM3	pHCMM3	primHCMM1	ATTTATGGTGAAGGATAAGTTTTGACCATCAAAG AAGGTTAATGTGGCTGTGGTTTCAGGGTCCATA AAGCTGGTGCTCTCATGCTTCTT
		primHCMM2	GGCCATGAAGCTTTTTCTTTCCAATTTTTTTTTTT TCGTCATTATAGAAATCATTACGACCGAGATTCC CGGGCGGTAGAGGTGTGGTCA
yHCMM4	pHCMM1	primHCMM3	AGTGATGCTAGCGATTATACTGATATCGATACG GATACAGAAGCTGAAGATGATGAATCACCAGAA ATGTCTAAAGGTGAAGAATTATTC
		primHCMM4	TTTTTATAAAGATGAAATATATACGTCTAAGAGCT AAAATGAAAAACTATACTAATCACTTATATCTACG GTTTCTTTGAAATTTTTTTG
yHCMM5	pHCMM1	primHCMM5	GCCAAGAACAAATTCAAAGAATGGAATGATAAG GGTGATGGTAAATTCTGGAGCTCGAAAAAGGAC ATGTCTAAAGGTGAAGAATTATTC
		primHCMM6	ATGAGAAACATGTGAATATGTATATATGTTATGC GGGAACCAACCCTTTACAATTAGCTATCTAACTA CGGTTTCTTTGAAATTTTTTTG
yHCMM6	pHCMM4	primHCMM7	ATGTCCTCCAACCGTGTTC
		primHCMM8	AAAAATAGAGGGAGGGGTGTGGCTAAATGCGAT TTTTTTTTTCTTTGGATATTATTTTCATTTTATATA ACTGTGGGAATACTCAGGTAT
yHCMM7	pHCMM4	primHCMM9	ACTTTAGCTAAATTACGCGAACAGGGTTTAATGG GATCCCCCGGGATGGTGAGCAAGGGC
		primHCMM10	TGTCTTCTCCTAAAAGTTTTTAAGTTTATGTTTAC TGTAACTGTGGGAATACTCAGGTAT

SI REFERENCES

1. Lapointe CP, et al. Protein-RNA networks revealed through covalent RNA marks. Nat Methods. 2015;12(12):1163-70.
2. Finke K, et al. A second trimeric complex containing homologs of the Sec61p complex functions in protein transport across the ER membrane of S. cerevisiae. The EMBO journal. 1996;15(7):1482-94.
3. Huh WK, et al. Global analysis of protein localization in budding yeast. Nature. 2003;425(6959):686-91.
4. Love MI, et al. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.
5. Hong EL, et al. Gene Ontology annotations at SGD: new data sources and annotation methods. Nucleic acids research. 2008;36(Database issue):D577-D81.
6. Balakrishnan R, et al. YeastMine--an integrated data warehouse for Saccharomyces cerevisiae data as a multipurpose tool-kit. Database : the journal of biological databases and curation. 2012;2012:bar062-bar.
7. Williams CC, et al. Targeting and plasticity of mitochondrial proteins revealed by proximity-specific ribosome profiling. Science. 2014;346(6210):748-51.
8. Jan CH, et al. Principles of ER cotranslational translocation revealed by proximity-specific ribosome profiling. Science. 2014;346(6210):1257521.
9. Fazal FM, et al. Atlas of Subcellular RNA Localization Revealed by APEX-Seq. Cell. 2019;178(2):473-90.e26.
10. Hogan DJ, et al. Diverse RNA-binding proteins interact with functionally related sets of RNAs, suggesting an extensive regulatory system. PLoS Biol. 2008;6(10):e255.
11. Hulsen T, et al. BioVenn - a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. BMC Genomics. 2008;9:488.
12. Anders S, et al. Differential expression analysis for sequence count data. Genome Biology. 2010;11(10):R106.
13. de Hoon MJ, et al. Open source clustering software. Bioinformatics. 2004;20(9):1453-4.
14. Reinders J, et al. Toward the complete yeast mitochondrial proteome: multidimensional separation techniques for mitochondrial proteomics. J Proteome Res. 2006;5(7):1543-54.
15. Morgenstern M, et al. Definition of a High-Confidence Mitochondrial Proteome at Quantitative Scale. Cell Rep. 2017;19(13):2836-52.
16. Saint-Georges Y, et al. Yeast mitochondrial biogenesis: a role for the PUF RNA-binding protein Puf3p in mRNA localization. PLoS One. 2008;3(6):e2293.
17. Bailey TL, et al. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol. 1994;2:28-36.
18. Lapointe CP, et al. Architecture and dynamics of overlapped RNA regulatory networks. RNA (New York, NY). 2017;23(11):1636-47.
19. Kaiser C. Methods in yeast genetics : a Cold Spring Harbor Laboratory course manual: 1994 edition. Cold Spring Harbor, N.Y. : Cold Spring Harbor Laboratory Press, 1994.; 1994.

20. Sheff MA, et al. Optimized cassettes for fluorescent protein tagging in Saccharomyces cerevisiae. Yeast. 2004;21(8):661-70.
21. Kaplan C, et al. Optimized sample preparation for single-molecule localization-based superresolution microscopy in yeast. Nat Protoc. 2015;10(7):1007-21.
22. Schindelin J, et al. Fiji: an open-source platform for biological-image analysis. Nat Methods. 2012;9(7):676-82.
23. Schneider CA, et al. NIH Image to ImageJ: 25 years of image analysis. Nature Methods. 2012;9:671.