Supporting Information

Rezgui et al. 10.1073/pnas.1300781110

SI Materials and Methods

Stable Isotope-Labeling by Amino Acids in Cell Culture Labeling. Stable isotope-labeling by amino acids in cell culture (SILAC) strains were exponentially grown in SD-K-R [6.7 mg/mL yeast nitrogen base without amino acids and $(NH_4)_2SO_4$ (Becton Dickinson), 0.68 mg/mL Triple Dropout CSM Mix -K, -R, -A (Formedium), 2% (wt/vol) glucose, 0.01% (wt/vol) Adenine, 0.2 mg/mL Pro] supplemented with 0.02 mg/mL Arg, 0.03 mg/mL Lys or 0.02 mg/mL $^{13}C_6$,$^{15}N_4$-Arg (Sigma), and 0.03 mg/mL $^{13}C_6$,$^{15}N_2$-Lys (Sigma). Equal ODs of light and heavy culture were mixed and subsequently processed together.

Protein Extraction and Digestion for MS. Cells were lysed with 1.85 M NaOH, 7.6% (vol/vol) β-mercaptoethanol, then proteins precipitated with 50% (wt/vol) TCA and washed twice with acetone. Pellets were resuspended in urea-buffer [8 M urea, 50 mM NH_4HCO_3, 0.5% RapiGest (Waters)] or a SDS-buffer [5% (wt/vol) SDS, 50 mM NH_4HCO_3]. Subsequently, proteins were reduced with 10 mM DTT at 56 °C for 30 min and carboxyamidomethylated with 25 mM iodoacetamide at room temperature for 30 min. Proteins in urea-buffer were diluted ten times and digested overnight with sequencing grade-modified Trypsin (Promega). Proteins in SDS-buffer were digested using the FASP method (1) and peptides were purified on C18 spin columns (The Nest Group).

Strong Cation Exchange Fractionation. Strong cation exchange fractionation was performed on microspin PolySULFOETHYL Aspartamide strong cation exchange columns (The Nest Group). Peptides were eluted with a six-step NaCl fractionation (50 mM, 100 mM, 150 mM, 200 mM, 400 mM, and 800 mM). Flow-through and fractions were cleaned-up on C18 MicroSpin columns (The Nest Group) and dried in a vacuum dryer.

Peptide Purification and Iso-Electric Focusing. Purified peptides were fractionated by iso-electric focusing on the OffGel Fractionator (Agilent; G3100AA) according to the manual of the High Res Kit, pH 3–10 (Agilent; 5188–6424), except the strips were exchanged by either Immobiline DryStrip pH 3–11 NL, 24 cm (GE Healthcare; 17-6003-77) or pH 3–11 NL, 13 cm (GE Healthcare; 17-6003-75), and ampholytes were substituted by 2% (vol/vol) IPG Buffer pH 3–11 NL (GE Healthcare; 17-6004-40). Peptides were focused into 24 fractions for 50 kVh at a maximum current of 50 μA, maximum voltage of 8,000 V, and maximum runtime of 100 h. Each fraction was acidified with 1% (vol/vol) CF_3COOH, purified on C18 MicroSpin columns (The Nest Group) and dried in a vacuum dryer.

LC-MS-MS/MS. Dried peptides were resuspended in 0.1% CF_3COOH for the LC-MS-MS/MS analysis. Split-free Easy nLC chromatography system (Proxeon) was used for the online reverse phase (C18 silica) liquid chromatography. Fused silica columns of 20-cm length (PicoFrit columns, PF-360-75-10-*N*-5; New Objective) were packed with C18 silica beads (Magic C18, 200 Å, 3 μm; Michrom Bioresources). A 250 nL/min gradient of buffer B [0.08% (vol/vol) HCOOH, 90% (vol/vol) CH_3CN] in buffer A [0.1% (vol/vol) HCOOH, 2% (vol/vol) CH_3CN] ranging from 2 to 35% (vol/vol) over 170 min was used to resolve peptides. The chromatography setup was directly coupled to the mass spectrometer (LTQ-Orbitrap Velos; Thermo Finnigan) configured for top-15 data dependent acquisition (DDA) by collision-induced fragmentation or top-8 DDA for the higher-energy collisional dissociation. FT-MS resolution was set at 60,000.

Protein Identification and Quantitation. RAW data files were converted to the mzXML format (2) and searched against the *Saccharomyces* Genome Database protein database using X! Tandem (3) with the *K*-score plug-in (4), OMSSA (5), Mascot (Matrix Science), and UW SEQUEST. Search parameters used were carboxyamidomethylation (57.022 Da) of Cys as static modification, $^{13}C_6$,$^{15}N_2$-Lys (8.01419892 Da), $^{13}C_6$,$^{15}N_4$-Arg (10.008252778 Da), and oxidation of Met (15.99491463 Da) as variable modifications, semitryptic digestion with a maximum of two missed cleavages, 25 ppm and 0.4-Da error tolerances for MS/MS and MS, respectively. Peptide probabilities were evaluated with PeptideProphet (6). iProphet was used to integrate the results from the four peptide search engines (7) and ProteinProphet (8) was used to estimate protein probabilities. Protein abundance ratios were computed as L/H (light/heavy) using XPRESS (9). Finally proteins were filtered for 1% false-discovery rate (FDR).

Data Normalization and Statistical Analysis of Differential Abundance. Results were stored in an in-house–developed database. From there, protein abundance ratios were imported into R (v2.11.0) (10). Proteins quantified in only one biological replicate and with less than two peptides were filtered out. Log_2 of ratios were median normalized using the preprocessCore library (v1.8) from the Bioconductor project (v2.5) (11). Statistical analysis of the differential abundance of proteins was done with bayes moderated *t* test using the LIMMA package (v3.8.2) of the Bioconductor project (v2.8) (12). Proteins ratios were filtered at the FDR threshold of 5% (or adjusted $P = 0.05$).

Random Forest Analysis. Significantly changing proteins were split into two classes and the abundance of codons that best predicted class membership was extracted by machine learning in R using the random forest implementation of the party package (v0.9-99992). The abundance of all codons, minus the stop codons in each gene was used as variables. The overall number of trees in the forest was set to 1,000 and the number of randomly preselected predictor variables for each split was set to the square root of the variables.

Quantitative PCR. Total RNA was extracted from exponentially growing cells with hot phenol as described previously (13). DNA was digested on a RNeasy column with DNase I according to the manufacturer's protocol (Qiagen). RNA was reverse-transcribed with SuperScript II reverse transcriptase (Invitrogen). Quantitative PCR (qPCR) was performed using the LightCycler FastStart DNA Master SYBR Green I kit (Roche) and the data analyzed using the LightCycler Quick System 350S (Roche). Quantification was performed using the relative standard curve method, and the transcriptional level of *PGK1* mRNA was used for normalization. The primers used are listed in Dataset S1, Table S7. For the experiments in Figs. S3*B* and S4*B*, RNA was extracted from the exponentially growing cells using the Ribo-pure yeast kit (Life Technologies) and converted to cDNA using the High Capacity cDNA Reverse Transcription Kit with RNase Inhibitor (Life Technologies), according to the manufacturer's instructions. qPCR was performed using the predesigned TaqMan gene expression assays on Applied Biosystems 7900HT Fast Real-Time PCR System.

Polysome Profiles. In yeast cells growing exponentially in YPD, translation was stopped with 100 μg/mL cycloheximide on ice for 10 min. Washed cells were resuspended in lysis buffer (10 mM Tris pH 7.5, 100 mM NaCl, 30 mM $MgCl_2$, 100 μg/mL cycloheximide, 200 μg/mL Heparin, 1:700 DEPC). Yeast cells

were lysed on ice by bead beating and cell extracts loaded on 6–45% (wt/vol) sucrose gradient prepared in 10 mM Tris pH 7.5, 100 mM NaCl, 30 mM $MgCl_2$, and ultracentrifuged at 260,000 × *g* for 2.45 h at 4 °C in a SW41 swing bucket rotor. Ribosome content was measured and recorded with a teledyne ISCO UA-6 detector. Polysome profiles were scanned and quantified with ImageJ.

Cloning of the Dual-Fluorescent Translation Reporter. Quadruple cyan fluorescent protein (CFP) under *GAL1* promoter and with a *CYC1* terminator was cloned into pRS305 with EcoRI/PmeI. Quadruple venus was cloned with SalI/NotI restriction site and SphI/SpeI sites were added after the starting ATG to allow insertion of codon-traps. Codon-traps insertion generated loss of the SpeI site. Fluorescent reporter was integrated into the *ura3-1* locus in W303 background with the AD-GEV Gal4 variant integrated in the *trp1-1* locus.

Expression of Fluorescent Reporter and Cell Imaging. Expression of the translation reporter was induced in cells at OD_{600} 0.1 grown in SD-complete medium with 50 nM β-estradiol (Sigma E1024) for 3 h (6 h for the time-course) and translation was blocked with 100 μg/mL cycloheximide. Cells were imaged after 1 h to allow maturation of the fluorophores. Cells were sonicated and loaded on well-slides coated with Con A. Differential interference contrast, CFP, and yellow fluorescent protein (YFP) images were taken at every position with an inverted epi-fluorescence microscope (Ti-Eclipse, Nikon) controlled by micromanager with a 40× oil objective in an incubation chamber at 30 °C. YFP exposure time was adjusted dependent on the codon-trap. $(AAA)_{10}$, $(GAA)_{10}$, $(CAA)_{10}$, and $(AAG)_{10}$ traps were imaged with 50-ms, 200-ms, 60-ms, and 600-ms exposure, respectively, in wild-type and *urm1*Δ cells. CFP images were all taken with 400-ms exposure. Single cells were segmented and analyzed with YeastQuant software (14), YFP/CFP ratios were measured in single cells and analyzed with MATLAB.

[^{35}S] Metabolic Labeling. Equal amounts of exponentially growing cells were pulsed for 15 min with [^{35}S]Met and [^{35}S]Cys (Hartmann analytics SCIS-103) and chased for 5 min with cold amino acids. Control cells were treated with 100 μg/mL cycloheximide before labeling. Extracts of the labeled cells were analyzed with a liquid scintillation counter.

Cycloheximide Chase. Exponentially growing TAP-tagged strains in SD-complete were treated with 100 μg/mL cycloheximide and samples of equal volume taken at different times after translation block. Cells extracts were probed with 1:2,500 PAP (Sigma; P1291) and 1:2,500 anti-Pgk1 (Invitrogen; 459250).

Total tRNA Preparation. Total tRNA from yeast was prepared by modified procedure of ref. 15, as follows: Total tRNA was extracted from 18 L of exponentially growing cells with 35% (vol/vol) acid phenol (Roti-Aqua-Phenol) for 3 h at room temperature. The aqueous phase was collected by centrifugation at 2,880 × *g* for 20 min. The extraction was repeated by adding equal amount of water. RNA was precipitated with 2% (wt/vol) potassium acetate (pH 5.0) and 2 volumes of ethanol for 2 h at −20 °C. Pellet was resuspended in 1 M NaCl and mixed vigorously for 2.5 h at 4 °C. The supernatant was collected, and extraction was repeated with 1 M NaCl at −20 °C. Crude tRNA was precipitated with ethanol. The pellet was dissolved in 0.3 M sodium acetate (pH 7.0) and 0.4 volume of isopropanol was added drop-wise at room temperature. Supernatant was collected by centrifugation at 2,880 × *g* for 20 min, and extraction was repeated with 0.2 volume of isopropanol. To the supernatant 0.4 volume of isopronanol was added and incubated for 2 h at –20 °C. The tRNA pellet was dissolved in water and purified with anion exchange column (DE52; Whatman) with a 0–1 M NaCl gradient. The fractions were pooled and precipitated with ethanol. The final tRNA pellet was dissolved in water.

Cloning and Purification of the tRNA Synthetases for Lysine and Phenylalanine. Phenylalanine tRNA synthetase was purified as described in ref. 16. The *Saccharomyces cerevisiae* gene for Lysine tRNA synthetase was cloned with NdeI/BamHI into a modified version of pPROEx vector (Invitrogen), in which the EheI site is replaced by a NdeI site. Synthetase expression was induced with isfopropyl-β-D-thiogalactopyranoside in *Escherichia coli* BL21 at 20 °C. The protein was purified by Ni^{2+} affinity chromatography and cleaved by tobacco etch virus protease (Invitrogen). Purified synthetase was stored in 20 mM Hepes pH7.5, 50 mM NaCl, 1 mM EDTA, 1 mM DTT, 10% (vol/vol) glycerol at −80 °C.

Aminoacylation and Purification of [^{14}C]Lys-tRNA^{Lys} and [^{14}C]Phe-tRNA^{Phe}. Fifty-three A_{260} units total tRNAs were aminoacylated using 10 μM purified lysine tRNA synthetase (LysRS) or 0.5% (vol/vol) of purified phenylalanine tRNA synthetase (PheRS) in aminoacylation buffer (50 mM Hepes pH 7.5, 70 mM NH_4Cl, 30 mM KCl, 11 mM $MgCl_2$, 3 mM ATP, 2 mM β-mercaptoethanol) supplemented with 20 μM [^{14}C]Lys or [^{14}C]Phe, respectively. After incubation for 45 min at 37 °C, His_6-EF-Tu·GTP was added for 1 min at room temperature. EF-Tu·GTP·[^{14}C]Lys-tRNA^{Lys} or EF-Tu·GTP·[^{14}C]Phe-tRNA^{Phe} were purified using Ni^{2+} column (Protino Ni-IDA), proteins were removed with phenol, and tRNAs were precipitated with 2% (wt/vol) potassium acetate (pH 5.0) and cold ethanol.

Biochemical and Kinetic Assays. For 70S initiation complexes, ribosomes (1.6 μM) were incubated with a fourfold excess of mRNA (GGCAAGGAGGUAAAUA AUG AAA UUC GUU AC or GGCAAGGAGGUAAAUA AUG UUC AAA GUU AC, from IBA; codon occupying the A-site is underlined) in the presence of 2.4 μM initiation factors IF1, IF2, IF3, 3.2 μM f[^{3}H]Met-tRNA^{fMet}, and 1 mM GTP in buffer A (50 mM Tris·HCl pH 7.5, 70 mM NH_4Cl, 30 mM KCl, 7 mM $MgCl_2$) for 30 min at 37 °C. Ternary complex, EF-Tu·GTP·[^{14}C]Lys-tRNA^{Lys} or EF-Tu·GTP·[^{14}C]Phe-tRNA^{Phe} were prepared by incubating 3 μM EF-Tu with 1 mM GTP, 3 mM phosphoenol pyruvate, 0.1 mg/mL pyruvate kinase for 15 min at 37 °C, followed by addition of 1 μM [^{14}C]Lys-tRNA^{Lys} or [^{14}C]Phe-tRNA^{Phe} for 1 min at room temperature. Ternary complex was added to the initiation complex and incubated for 10 min at room temperature. The amount of [^{14}C]Lys and f[^{3}H]Met; [^{14}C]Phe and f[^{3}H]Met bound to ribosomes was determined by nitrocellulose filtration.

For dissociation of fMetLys-tRNA^{Lys} from the A site, 4 μM ribosomes, 6 μM initiation factors (IF1, IF2, IF3), and 12 μM [^{14}C]Lys-tRNA^{Lys} was used for initiation complex and ternary complex, respectively. Ternary complex was incubated with initiation complex for 1 min at room temperature to form a pretranslocation complex. Then, the Mg^{2+} concentration of the pretranslocation complex was adjusted to 21 mM to prevent premature drop-off of fMetLys-tRNA^{Lys} from the A site and the complex was kept on ice. Pretranslocation complexes were purified by size-exclusion chromatography (BioSuite 450 HR; Waters) in buffer A with 21 mM $MgCl_2$, shock-frozen, and stored at −80 °C in small aliquots. Dissociation of fMetLys-tRNA^{Lys} from the A site was followed after lowering Mg^{2+} concentration to 7 mM at 37 °C, and the amount of peptidyl-tRNA bound to the A site at different time points was determined by nitrocellulose filtration. Evaluation was done as in ref. 17.

To measure the time courses of peptide bond formation, quench-flow assays were performed at 24 °C in a KinTek RQF-3 apparatus. Initiation complex (2 μM) was rapidly mixed with ternary complex (0.6 μM). After different incubation times, reactions were stopped with KOH (0.5 M). RNA was hydrolyzed for 30 min at 37 °C, neutralized, and dipeptides were analyzed by RP-HPLC (18).

1. Wiśniewski JR, Zougman A, Nagaraj N, Mann M (2009) Universal sample preparation method for proteome analysis. *Nat Methods* 6(5):359–362.
2. Pedrioli PGA, et al. (2004) A common open representation of mass spectrometry data and its application to proteomics research. *Nat Biotechnol* 22(11):1459–1466.
3. Craig R, Beavis RC (2004) TANDEM: Matching proteins with tandem mass spectra. *Bioinformatics* 20(9):1466–1467.
4. MacLean B, Eng JK, Beavis RC, McIntosh M (2006) General framework for developing and evaluating database scoring algorithms using the TANDEM search engine. *Bioinformatics* 22(22):2830–2832.
5. Geer LY, et al. (2004) Open mass spectrometry search algorithm. *J Proteome Res* 3(5):958–964.
6. Nesvizhskii AI, Keller A, Kolker E, Aebersold R (2003) A statistical model for identifying proteins by tandem mass spectrometry. *Anal Chem* 75(17):4646–4658.
7. Shteynberg D, et al. (2011) iProphet: Multi-level integrative analysis of shotgun proteomic data improves peptide and protein identification rates and error estimates. *Mol Cell Proteomics* 10(12):M111.007690.
8. Keller A, Nesvizhskii AI, Kolker E, Aebersold R (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal Chem* 74(20):5383–5392.
9. Han DK, Eng J, Zhou H, Aebersold R (2001) Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. *Nat Biotechnol* 19(10):946–951.
10. R Core Team (2013) *R: A Language and Environment for Statistical Computing*. Available at http://www.R-project.org. Accessed June 26, 2013.
11. Bolstad BM, Irizarry RA, Astrand M, Speed TP (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 19(2):185–193.
12. Smyth GK (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 3:Article3.
13. Collart MA, Oliviero S (2001) Preparation of yeast RNA. *Curr Protoc Mol Biol* Chapter 13:Unit13.12.
14. Pelet S, Dechant R, Lee SS, van Drogen F, Peter M (2012) An integrated image analysis platform to quantify signal transduction in single cells. *Integr Biol (Camb)* 4(10): 1274–1282.
15. Holley RW (1967) Isolation of sRNA from intact yeast cells. *Methods Enzymol* 12: 596–598.
16. Mittelstaet J, Konevega AL, Rodnina MV (2011) Distortion of tRNA upon near-cognate codon recognition on the ribosome. *J Biol Chem* 286(10):8158–8164.
17. Konevega AL, et al. (2004) Purine bases at position 37 of tRNA stabilize codon-anticodon interaction in the ribosomal A site by stacking and Mg2+-dependent interactions. *RNA* 10(1):90–101.
18. Katunin VI, Muth GW, Strobel SA, Wintermeyer W, Rodnina MV (2002) Important contribution to catalysis of peptide bond formation by a single ionizing group within the ribosome. *Mol Cell* 10(2):339–346.

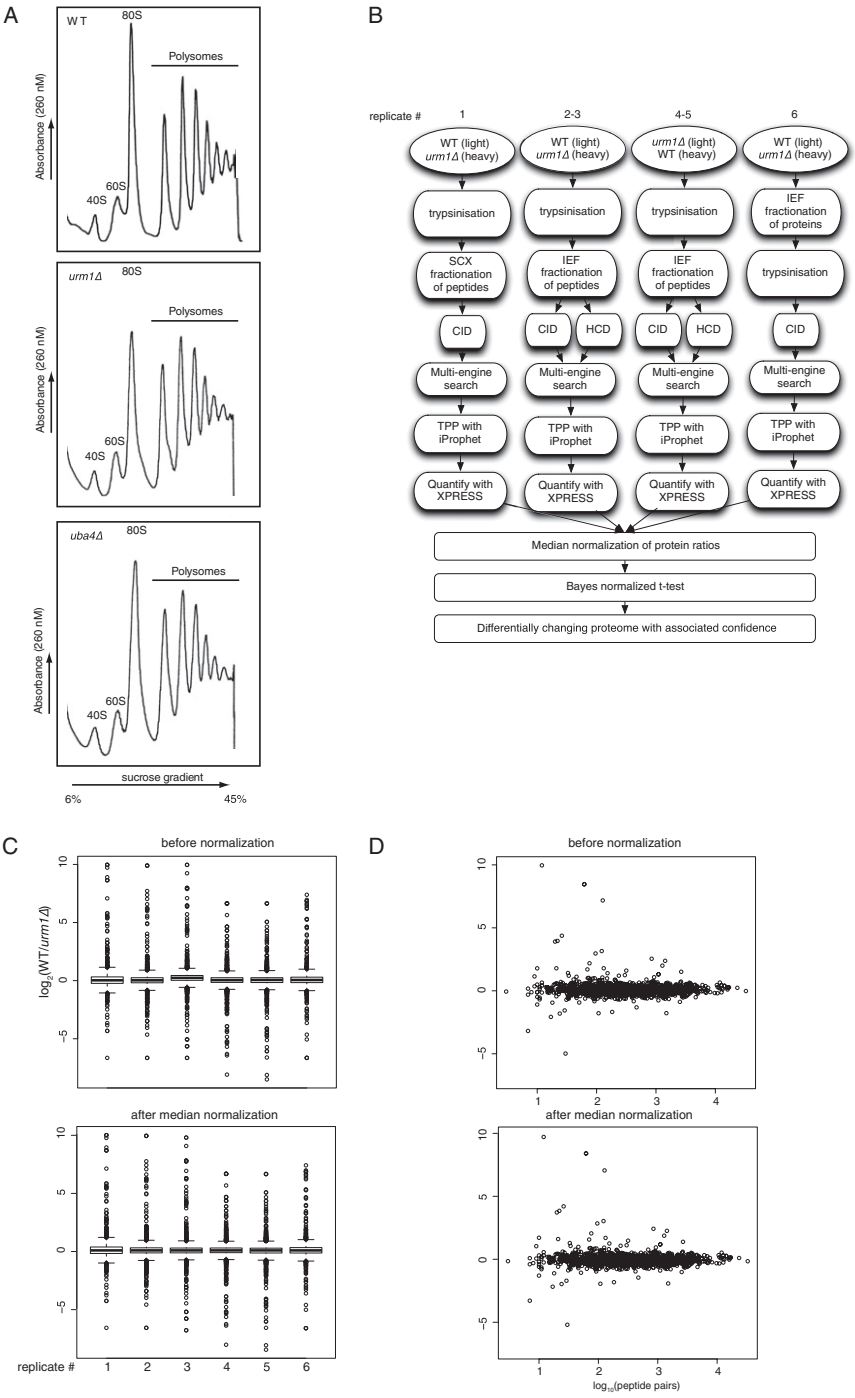

Fig. S1. Complement to Figs. 1 and 2. (*A*) Polysome profiles of wild-type, *urm1*Δ, and *uba4*Δ cells separated on a 6–45% (wt/vol) sucrose gradient. (*B*) Scheme of the workflow used to detect statistically significant differentially expressed proteins using SILAC-based mass spectrometry. (*C*) Box-and-whisker plots showing the distribution of protein abundance ratios before and after median normalization. The median ratio is shown as a thick black line surrounded by a box representing the inter quartile range containing 50% of the data. Whiskers extend up to two SDs from the median. (*D*) MA plot of the protein-abundance ratios before and after median normalization. For each protein the mean of WT/*urm1*Δ ratio over the six replicates was calculated and plotted as $\log_2$(WT/*urm1*Δ) against $\log_{10}$ of number of SILAC pairs used for quantification of each protein.

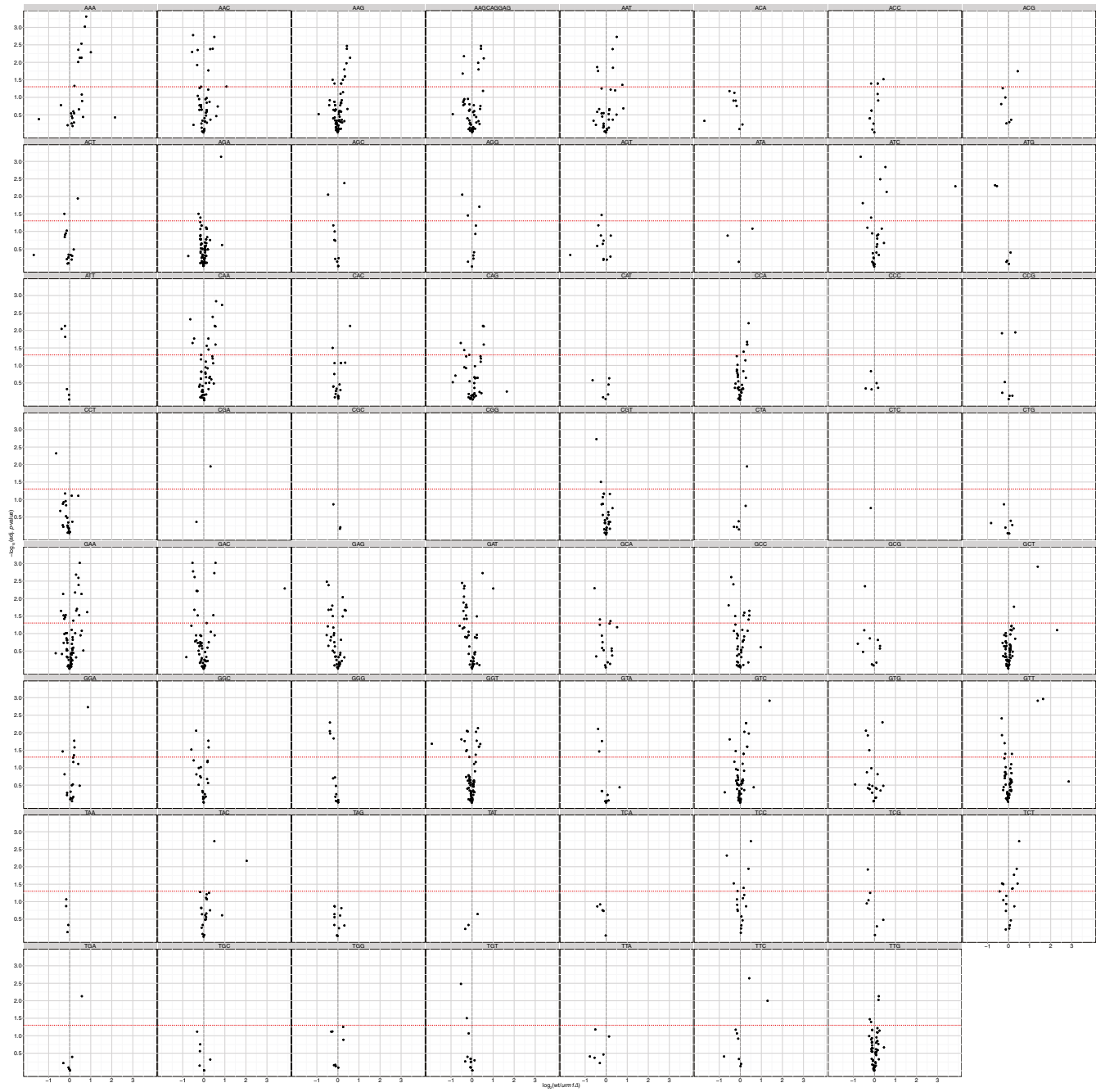

Fig. S2. Complement to Figs. 2 and 3. Volcano plots showing the protein abundance ratios, measured by SILAC-based quantitative proteomics, with statistical significance, computed using Bayes moderated *t* test, of the top 1% yeast genes with the highest frequency of the indicated codons. The dotted red line indicates the 5% FDR, threshold for statistically significant changes in protein abundance. The gray dotted line indicates a wild-type/*urm1*Δ of 1.

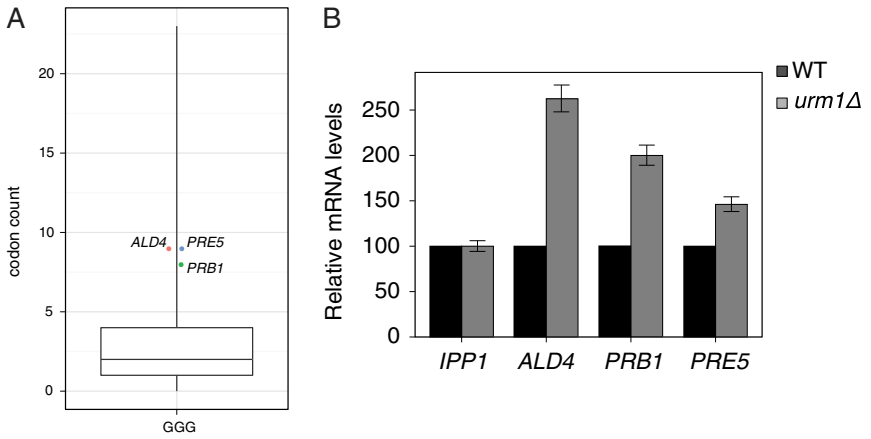

Fig. S3. Complement to Fig. 2. The variable importance of GGG identified by the random forest analysis is caused by differential transcription rather than differential translation. (*A*) Box-and-whisker plot showing the distribution of GGG codons across the yeast genome. The median is represented by the thick black line surrounded by a box representing the interquartile range containing 50% of the data. Whiskers represent the minimum and maximum values. Highlighted are three genes that were significantly up-regulated in the *urm1Δ* quantitative proteomics analysis. (*B*) The mRNA levels of the three genes highlighted in *A* were measured by quantitative PCR. *IPP1* is used as negative control. Data show the mean ± SEM from three experiments.

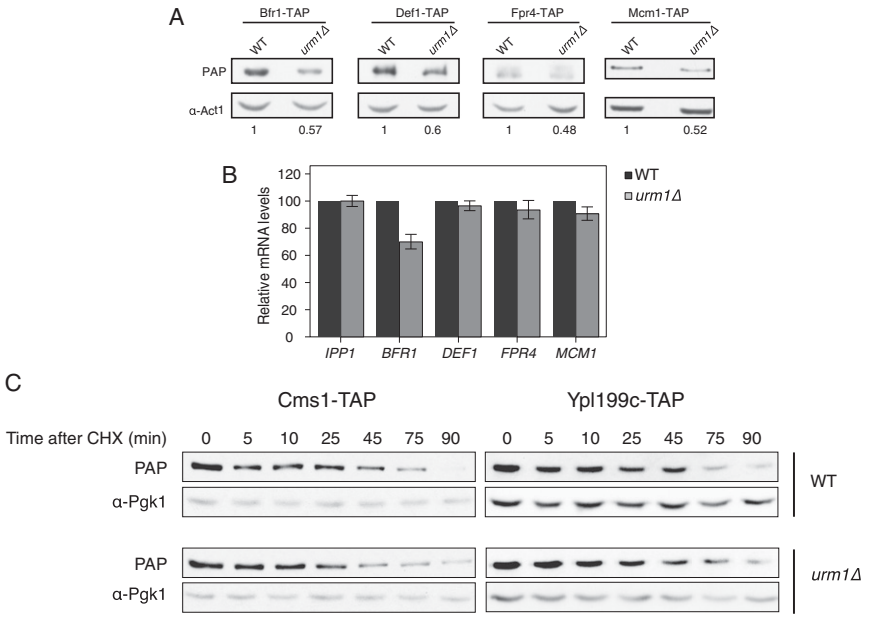

Fig. S4. Complement to Fig. 3. Translation of AAA-, CAA-, and GAA-rich genes is regulated by *URM1*. (*A*) Western blot of TAP-tagged candidate genes enriched for AAA, CAA, and GAA codons from wild-type and *urm1*Δ cells using PAP antibody or anti-Act1 as loading control. The quantification indicates the protein abundance ratio relative to wild-type averaged from three independent experiments. (*B*) The mRNA levels of the candidates were quantified by qPCR. *IPP1* is used as negative control. Data show the mean ± SEM over three experiments. (*C*) Protein stability in wild-type and *urm1*Δ cells was followed after translation block by 200 μg/μL cycloheximide (CHX). Equal volumes of culture were harvested at different time points after CHX block and protein levels probed with PAP and anti-Pgk1.

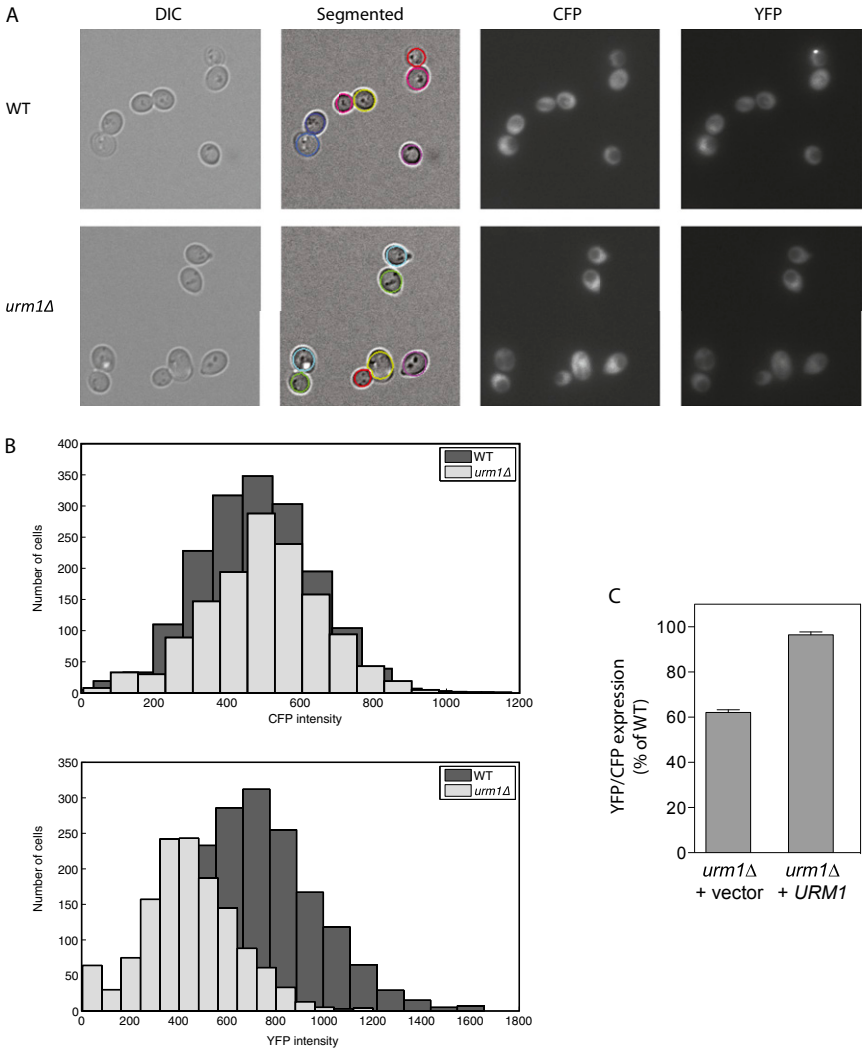

Fig. S5. Complement to Fig. 4. *URM1* is required for efficient translation of codon-enriched translation reporters. (*A*) Representative images (40× magnification) showing the expression of the codon-specific reporter in wild-type and *urm1*Δ cells. Expression levels of YFP and CFP reporters were measured in WT or *urm1*Δ cells after 3-h induction. YFP/CFP ratio was subsequently analyzed in single cells with the YeastQuant program. (*B*) Distribution of control CFP intensity signal (*Upper*) and YFP intensity signal (*Lower*) in wild-type and *urm1*Δ cells expressing a $(CAA)_{10}$-YFP reporter. (*C*) Expression of the translation reporter after 3-h induction with $(CAA)_{10}$-codon trap was compared in *urm1*Δ cell with an empty vector or a vector expressing *URM1*. Data show the mean YFP/CFP ratio ± SEM from at least 1,000 cells plotted as percent of wild-type control.

Dataset S1. List of proteins, genes, and reagents used in this study

Dataset S1

The dataset consists of seven tables, as follows: Table S1: List of the significantly differentially regulated proteins in *urm1*Δ; Table S2: Gene ontology and Munich Information Center for Protein Sequences (MIPS) functional classes enrichment analysis of the significantly down-regulated proteins in *urm1*Δ; Table S3: Gene ontology and MIPS functional classes enrichment analysis of the significantly up-regulated proteins in *urm1*Δ; Table S4: Results from the *elp3*Δ vs. *urm1*Δ quantitative proteomics screen; Table S5: AAA, CAA, and GAA codon frequency distribution in the yeast genome; Table S6: Gene ontology and MIPS functional classes enrichment analysis for the most AAA, CAA, and GAA codon rich genes in the yeast genome; Table S7: List of reagents used in this study.