



Selective Translation of Low Abundance and Upregulated Transcripts in *Halobacterium salinarum*

Adrián López García de Lomana,^a Ulrike Kusebauch,^a Arjun V. Raman,^a Min Pan,^a Serdar Turkarslan,^a Alan P. R. Lorenzetti,^{a,b} Robert L. Moritz,^a Nitin S. Baliga^{a,c,d,e,f}

^aInstitute for Systems Biology, Seattle, Washington, USA

^bMicrobial Systems Biology Lab, Universidade de São Paulo, Ribeirão Preto, Brazil

^cDepartment of Biology, University of Washington, Seattle, Washington, USA

^dDepartment of Microbiology, University of Washington, Seattle, Washington, USA

^eMolecular and Cellular Biology Program, University of Washington, Seattle, Washington, USA

^fLawrence Berkeley National Lab, Berkeley, California, USA

Adrián López García de Lomana, Ulrike Kusebauch, and Arjun V. Raman contributed equally to this work. Author order reflects relative effort towards completion of the manuscript.

ABSTRACT When organisms encounter an unfavorable environment, they transition to a physiologically distinct, quiescent state wherein abundant transcripts from the previous active growth state continue to persist, albeit their active transcription is downregulated. In order to generate proteins for the new quiescent physiological state, we hypothesized that the translation machinery must selectively translate upregulated transcripts in an intracellular milieu crowded with considerably higher abundance transcripts from the previous active growth state. Here, we have analyzed genome-wide changes in the transcriptome (RNA sequencing [RNA-seq]), changes in translational regulation and efficiency by ribosome profiling across all transcripts (ribosome profiling [Ribo-seq]), and protein level changes in assembled ribosomal proteins (sequential window acquisition of all theoretical mass spectra [SWATH-MS]) to investigate the interplay of transcriptional and translational regulation in *Halobacterium salinarum* as it transitions from active growth to quiescence. We have discovered that interplay of regulatory processes at different levels of information processing generates condition-specific ribosomal complexes to translate preferentially pools of low abundance and upregulated transcripts. Through analysis of the gene regulatory network architecture of *H. salinarum*, *Escherichia coli*, and *Saccharomyces cerevisiae*, we demonstrate that this conditional, modular organization of regulatory programs governing translational systems is a generalized feature across all domains of life.


IMPORTANCE Our findings demonstrate conclusively that low abundance and upregulated transcripts are preferentially translated, potentially by environment-specific translation systems with distinct ribosomal protein composition. We show that a complex interplay of transcriptional and posttranscriptional regulation underlies the conditional and modular regulatory programs that generate ribosomes of distinct protein composition. The modular regulation of ribosomal proteins with other transcription, translation, and metabolic genes is generalizable to bacterial and eukaryotic microbes. These findings are relevant to how microorganisms adapt to unfavorable environments when they transition from active growth to quiescence by generating proteins from upregulated transcripts that are in considerably lower abundance relative to transcripts associated with the previous physiological state. Selective translation of transcripts by distinct ribosomes could form the basis for adaptive evolution to new environments through a modular regulation of the translational systems.

Citation López García de Lomana A, Kusebauch U, Raman AV, Pan M, Turkarslan S, Lorenzetti APR, Moritz RL, Baliga NS. 2020. Selective translation of low abundance and upregulated transcripts in *Halobacterium salinarum*. mSystems 5:e00329-20. <https://doi.org/10.1128/mSystems.00329-20>.

Editor Michael Rust, Institute for Genomics & Systems Biology

Copyright © 2020 López García de Lomana et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Nitin S. Baliga, nbaliga@systemsbiology.org.

 Selective translation of low-abundance and upregulated transcripts in *Halobacterium salinarum*

Received 22 April 2020

Accepted 8 July 2020

Published 28 July 2020

KEYWORDS translational regulation, selective translation, transcription-translation interplay, ribosome heterogeneity, transcriptomics, ribosome profiling, proteomics, archaea

The concept that functional heterogeneity results from variations in the translational machinery is gaining renewed support (1). Variations in ribosomal proteins (RPs), ribosomal RNAs (rRNAs), transfer RNAs (tRNAs), and translation factors are four molecular axes that provide support for translational regulation (2, 3). Translational regulation plays an important role in fundamental biological processes like vertebrate development, where ribosomes with specific subunit composition, i.e., presence or absence of ribosomal protein RPL10A/uL1 and RPS25/eS25, preferentially translate functionally distinct pools of mRNAs in murine stem cells (4). Also in vertebrate development, ribosome-mediated translational specificity occurs through direct interaction between RP L38 and structural RNA elements resembling internal ribosome entry sites in the 5' untranslated region (UTR) of select Hox mRNAs (5, 6). Furthermore, selective translation modulates stress response and various complex human disease phenotypes (7, 8). Variations in RP stoichiometry with functional implications have been reported for *Saccharomyces cerevisiae* (9), such that RP Asc1/RACK1 is required for efficient translation of short mRNAs (10), and Rps26-depleted ribosomes support stress responses (11, 12). Moreover, a ribosome code has been proposed to explain the divergence in phenotypic outcome when individual paralogs of duplicated RP genes are deleted (13). Not only variations in RP stoichiometry (14) but also 16S rRNA variation in *Escherichia coli* has been associated with functional differences that can regulate stress response via RpoS and RelA regulons (15). While the ultimate functional consequences of ribosome diversity require careful experimental validation (16) and meticulous interpretation in light of alternative regulatory mechanisms (17), these studies argue against the notion that the ribosome is a rote translation machine, but rather a context-sensitive regulatory control element steering conditional protein synthesis (14, 18). Compared to the organisms above, fewer studies have been performed in the archaea domain. Nevertheless, translation initiation is suggested to utilize a completely novel mechanism in haloarchaea (19). In the archaeon *Sulfolobus acidocaldarius*, RP L7Ae binds to select coding and noncoding RNAs, including its own mRNA, to regulate translation (20).

In this article, we explore the interplay of translational and transcriptional regulation in driving microbial transitions to quiescent states upon encountering unfavorable environments. Upon exposure to a stressful environment, microorganisms elicit protective and acclimation responses that are often associated with a dormant or quiescent phenotypic state (21). We previously used a systems biology approach to investigate transcript and protein level changes in *Halobacterium salinarum* when this halophilic archaeon underwent transition from active to quiescent growth states in batch cultures (22) and in response to a controlled switch from favorable oxic to unfavorable anoxic conditions (23). We discovered that upon encountering nutrient depletion or anoxic conditions, a large fraction of genes (51% in batch cultures [24] and 9% in oxygen transitions [23]) in the genome were differentially regulated. Notably, the drop in ATP level partially explained why downregulated transcripts from the active growth state continued to persist in the stationary phase and under anoxic conditions. Two observations highlighted the differences in translational regulation across transitions between active growth and quiescent states. First, even though the upregulated transcripts in the quiescent state were orders of magnitude lower in abundance relative to active growth state transcripts, their protein levels increased. Second, upon encountering favorable conditions, protein levels of active growth-associated genes increased within minutes, well before their upregulation at the transcriptional level (23). Therefore, we hypothesized that genes encoding proteins of the translation system (e.g., RPs, translation factors, RNases, and RNA-modifying enzymes) are regulated to

produce a diverse population of ribosomes that selectively translate physiological state-specific transcripts relevant for a given environmental condition.

To test this hypothesis, we performed sequential window acquisition of all theoretical mass spectra (SWATH-MS) proteomics in concert with RNA sequencing and ribosome profiling in the halophilic archaeon *H. salinarum*. This analysis revealed preferential translation of upregulated and low abundance transcripts. In an attempt to uncover the underlying mechanism for this phenomenon, we discovered that transcriptional, translational, and posttranslational regulatory mechanisms act on RP genes across the growth phase to effect conditional changes in their abundance and stoichiometric association with the assembled ribosome. We further analyzed previously developed environment and gene regulatory influence network (EGRIN) models (25) for signatures of heterogeneity in the whole translation system. We discovered that transcriptional regulation of RPs is fractured into multiple mostly discrete but partially overlapping conditionally coregulated gene modules (corems). Subsets of translation system genes associated with tRNA charging (aminoacyl-tRNA synthetases), translation factors (e.g., initiation, elongation, release, and recycling factors), and RNA transcription and processing (e.g., RNA polymerase, RNases, and RNA modification enzymes) are split among conditionally coregulated corems. Notably, we observed similar modular regulation of ribosomal proteins in *E. coli* and yeast. This modular and environment-specific regulation of ribosomal proteins might have emerged to favor evolvability (26) and functional specialization of ribosomes (27). In sum, our data support the hypothesis that environment-specific ribosome composition and coupled transcription-translation in prokaryotes selectively bias translation of low abundance and upregulated transcripts to produce proteins needed for the environment-relevant physiological state.

RESULTS

Interplay of transcriptional and translational regulation across growth phase-associated physiological state transitions. When *H. salinarum* cells transition from lag to stationary growth phase, they travel across multiple regulatory configurations or states involving gene expression changes in more than two thirds of all genes in the genome (24). In order to investigate the interplay of transcriptional and translational regulation at the whole-genome scale for the growth-associated physiological state transition in *H. salinarum*, we quantified for each transcript the relative change in its abundance and ribosomal footprints across all phases of growth in batch culture. Specifically, we probed four time points representative of different growth phases, i.e., early exponential (time point 1 [TP1]), mid-exponential (TP2), late exponential (TP3), and stationary (TP4). For each sampled time point, we quantified the transcriptome state using RNA sequencing (RNA-seq) and ribosome translational activity using ribosome profiling (Ribo-seq) (see Materials and Methods for details). We used a filter on magnitude (absolute \log_2 fold change [FC] > 1) and significance (DESeq2-adjusted $P < 0.05$) to assess both transcript abundance and ribosomal footprint changes across growth phase (for a detailed list of transcript quantification, see Table S1, tabs 1 to 12, in the supplemental material). Of the total set of 2,663 annotated genes in *H. salinarum*, 304 were not regulated at the transcriptional or translational level (yellow dots in Fig. 1A). Among the 1,103 genes that were subjected to regulation, changes in ribosomal footprints per transcript for 875 genes (79%) are proportional to changes in transcript abundance (black dots in Fig. 1A). Among this set, upregulated genes were significantly enriched with functions associated with gas vesicle organization (GO:0031412) and cell motility (GO:0048870), protein phosphorylation (GO:0006468), the iron-sulfur cluster (GO:0016226), reactive oxygen species metabolism (GO:0072593), and response to stress (GO:0006950). Functions enriched within downregulated genes included nitrogen and phosphorous metabolism (GO:0006807 and GO:0006793), carbohydrate catabolism (GO:0016052), pyruvate biosynthesis (GO:0042866) and energy production via proton-exporting ATPase activity (GO:0036442).

Interestingly, ribosomal footprint changes in 228 genes (21%) could not be explained by transcript level change, i.e., ribosomal footprints were greater or less than

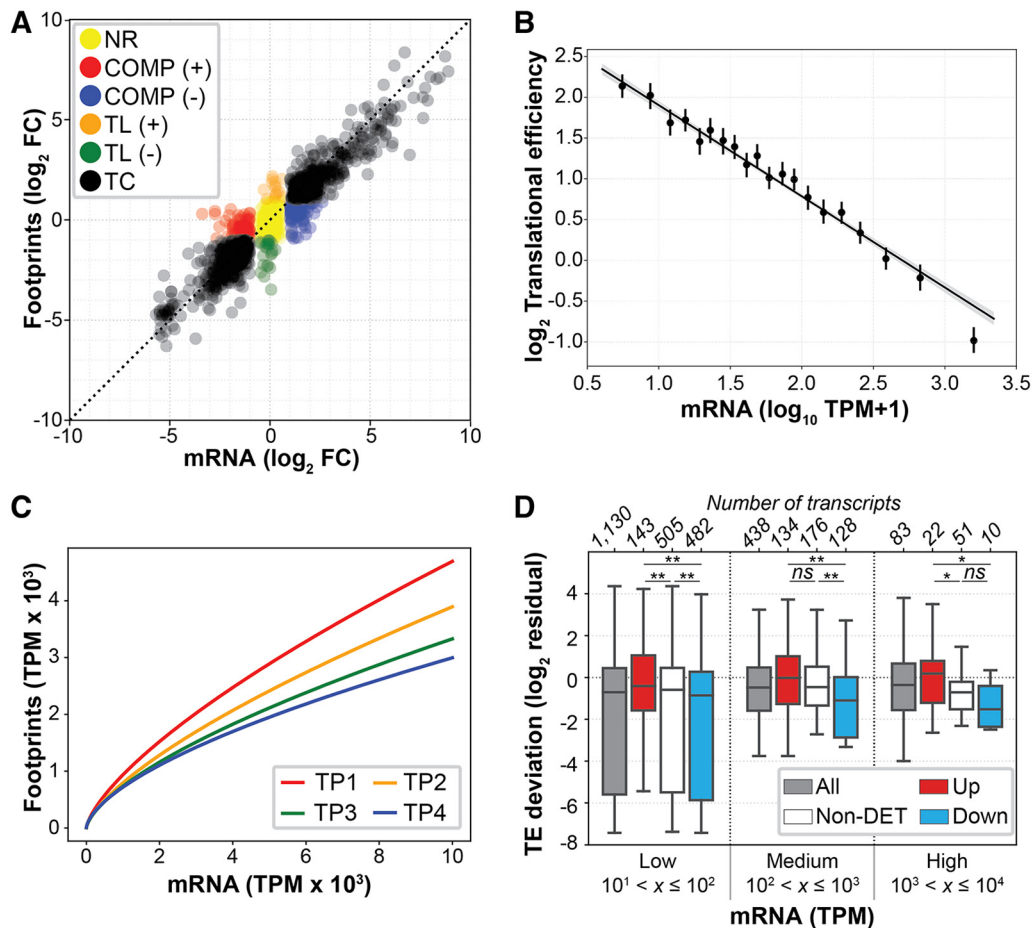


FIG 1 Coupled transcription-translation regulation across growth phase. (A) Scatterplot of relative changes in transcript abundance (\log_2 FC mRNA) and ribosomal footprints (\log_2 FC footprints) in stationary phase (TP4) with respect to early exponential phase (TP1). Black dots (TC; $n = 875$) represent genes regulated at the transcriptional level only. Red ($n = 79$) and blue ($n = 110$) dots represent genes under positive (+) and negative (-) compensatory mechanisms, respectively (COMP). Orange ($n = 15$) and green ($n = 24$) dots represent genes translationally regulated only (TL). Yellow dots represent genes that are not transcriptionally or translationally regulated (NR; $n = 304$). (B) Linear regression of transcript abundance (x axis; \log_{10} TPM + 1) and TE (y axis; \log_2 Ribo-seq/RNA-seq ratio). Slope, $a = -1.10$; correlation coefficient $R = -0.52$; $P < 10^{-132}$. The gray area represents the 95% confidence interval. (C) Regression analysis of predicted ribosomal footprints from transcript expression at different growth phases. (D) Deviation distributions from the expected TE given expression (y axis) in the context of transcriptional regulation across growth phase (stationary versus early exponential; TP4 versus TP1). Transcriptionally upregulated genes are shown in red, downregulated genes in blue, nondifferentially expressed (non-DET) genes in white, and all genes in gray. The horizontal axis shows expression levels (x): low, $10 < x \leq 100$; medium, $100 < x \leq 1,000$; high, $1,000 < x \leq 10,000$. The number of transcripts of each boxplot (n) is shown at the top. Asterisks indicate significance: ns, nonsignificant; *, $P < 0.05$; **, $P < 0.01$.

expected relative to change in transcript level. This observation is in agreement with an earlier report which showed that more than 20% of the *H. salinarum* transcriptome exhibits nonaverage translational efficiency (TE) (quantified as the \log_2 ratio of ribosomal footprints over transcript abundance [28]) in stationary versus exponential phase (29). Among this set of translationally regulated genes, 189 genes displayed compensatory mechanisms (30). First, a set of 110 genes were upregulated at the transcriptional level, which was not reflected in a corresponding increase in their ribosomal footprints (blue dots in Fig. 1A). This set of genes are involved in phospholipid and polysaccharide metabolism (GO:0006644 and GO:0000271), transcription initiation (GO:0006352), and response to stimulus (GO:0050896). Inversely, 79 genes had a significant decrease in transcript levels, although their ribosomal loading remained constant (red dots in Fig. 1A). This set of genes encode functions of DNA replication and repair (GO:0006261 and GO:0006284), modified amino acid biosynthesis (GO:0042398), and electron transfer activity (GO:0009055).

Orthogonally, 39 genes changed only at the translational level, i.e., while their mRNA levels did not change significantly, their ribosomal footprint abundances did (orange and green dots in Fig. 1A). Among the translationally upregulated genes, we noted *VNG2625C*, which encodes a PrsW family protease, and *nrmA*, which encodes a protein that is a bifunctional oligoribonuclease/PAP phosphatase which regulates degradation of nanoRNAs in bacteria (31). The translationally downregulated genes were enriched in redox energy metabolism functions, specifically components of the electron transport chain, including NADH-dependent flavin oxidoreductase gene *VNG0933G*, cytochrome *c* biogenesis protein gene *VNG0150H*, and nicotinate-nucleotide pyrophosphorylase gene *VNG1884G*, in addition to the PadR family transcriptional repressor *VNG7102*. A detailed list of all enriched functions is provided in Table S1, tabs 13 to 20; additionally, a visualization summary is provided as Fig. S1B to H in the supplemental material.

These observations are consistent with the known physiological shift of *H. salinarum* when it transitions from early exponential to stationary phase of growth (24). In summary, while 79% of all regulated genes have consistent changes at the mRNA level and ribosomal footprints ($n = 875$), there is significant evidence for interplay of transcriptional and translational regulation for at least 228 genes (21%).

Selective translation of low abundance and upregulated transcripts. Next, we investigated the potential for ribosomes to regulate physiological transitions through different phases of active growth and into a quiescent state in stationary phase. In particular, we asked whether and how the translation machinery selectively translates low abundance and upregulated mRNAs that are required for homeostasis in each growth phase. By comparing transcript abundance to TE, we discovered that TE negatively correlates with mRNA expression (slope = -1.10 ; $R = -0.52$; $P < 10^{-132}$; Fig. 1B), irrespective of transcript length or transcript half-life (data obtained from reference 32; see Fig. S2 for more details). This negative relationship exists at all physiological states across growth phase, becoming stronger in stationary phase (Fig. 1C). This finding suggests that ribosomes associate more efficiently with low abundance transcripts. This finding was further supported by the observation that the 15 transcripts that are exclusively upregulated at the translational level [TL (+) in Fig. 1A] are low in abundance (Fig. S3). Furthermore, we discovered across all levels of expression that transcriptionally upregulated genes were associated with higher TE with respect to transcriptionally downregulated genes (Mann-Whitney U test, $P < 0.05$; Fig. 1D). This observation supports the notion that transcription and translation are coupled in archaea and is consistent with our discovery of selective translation of upregulated transcripts across growth-associated physiological transitions in *H. salinarum*. Through the analysis of corresponding changes in transcript levels and ribosome occupancy in *E. coli* (33), we have discovered compelling evidence that bacteria also preferentially translate low abundance transcripts. However, we did not see evidence for this phenomenon in yeast (34), suggesting that this phenomenon might be exclusive to prokaryotes (Fig. S4).

RP abundance and composition within ribosomes across growth phase-associated physiological states. The preferential translation of low abundance and upregulated transcripts, as well as translational regulation of 228 genes (21% of all regulated genes), motivated further investigation into a mechanistic explanation for these phenomena. It has been demonstrated in other organisms that differential RP stoichiometry plays a major role in driving physiological modulation during cell growth (9). Therefore, we assessed protein abundance and composition of assembled ribosomes using quantitative proteomics, specifically sequential window acquisition of all theoretical mass spectra (SWATH-TM; see Materials and Methods and Table S1, tabs 21 and 22, for details). We observed a progressive decrease of RP abundance as growth phase advanced toward stationary phase, when a median \log_2 FC = -1.04 indicates a general repression of translation in the cell (Fig. 2A). Notably, relative abundance changes in some RPs deviated across time from the overall trend, suggesting that

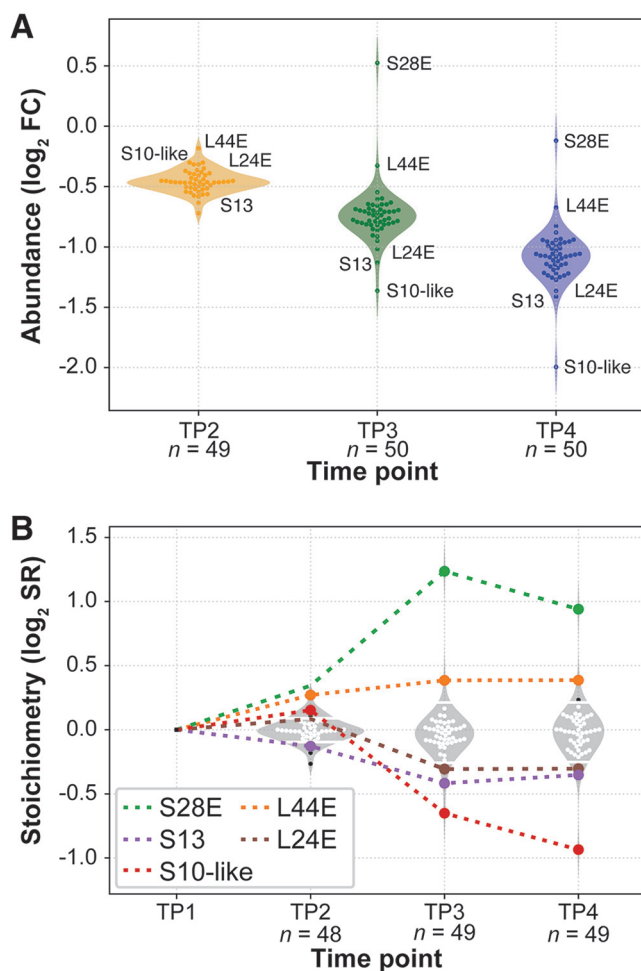


FIG 2 Ribosome abundance and composition shifts across growth-related physiological states. (A) Relative RP abundance changes (\log_2 FC) with respect to early exponential phase (TP1). Protein detection ranged from $n = 49$ to $n = 50$ RPs across samples. (B) Ribosome composition (\log_2 RP stoichiometry ratio) changes across growth phase. Small white dots represent RP stoichiometry values, and small black dots represent values outside the 95% confidence interval (95% CI). Large colored dots highlight RPs with two or more deviation events outside the 95% CI threshold. Dashed colored lines assist picturing the trend across time. White horizontal bars inside violin plots indicate 95% CIs.

ribosomes with distinct RP compositions were active in each of the different growth-associated physiological states.

As expected, ribosome composition was generally conserved but had few notable exceptions. Relative stoichiometry ratio (SR) of five RPs consistently deviated from the general trend across multiple time points, with most significant stoichiometric deviations of approximately twofold increase and decrease in S28E and S10-like proteins, respectively (Fig. 2B). Within statistical significance, we observed higher ribosomal associations of S28E (\log_2 SR = 0.94) and L44E (\log_2 SR = 0.39). Both RPs with higher stoichiometry belong to a three-gene operon together with *ndk* (which encodes a nucleoside diphosphate kinase), a conserved gene across all domains of life with pleiotropic effects in a wide range of functions (35). The RP stoichiometry of S28E and L44E increased despite significant downregulation of the operon at the transcriptional level (5,483 to 342 transcripts per million [TPM]; \log_2 FC = -4.00 ; $P < 1 \times 10^{-7}$). The differential association of the two RPs in the ribosome has implications on regulation of translation, given that S28E is located at the mRNA exit site in eukaryotes (36) and L44E is a conserved component of the E-tRNA site (37) of the eukaryotic and archaeal ribosome where it interacts with initiation factor eEF3 (38). Similarly, ribosomes in stationary phase have lower stoichiometry for three RPs: S10-like, S13, and L24E with

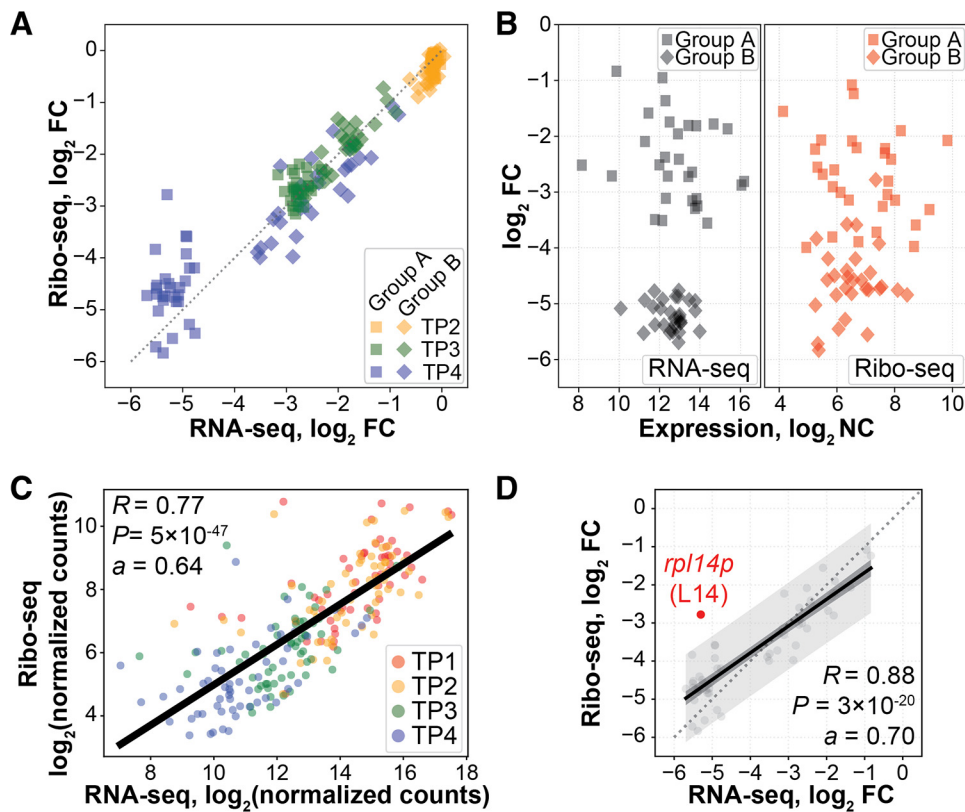


FIG 3 RP genes are transcriptionally and translationally regulated. (A) Scatterplot of relative abundance changes with respect to TP1. Each dot represents median \log_2 FC over three biological replicates for a given RP gene. Colors map to time points. (B) Scatterplot of expression and relative change of RP gene abundance (RNA-seq; left) and ribosomal footprints (Ribo-seq; right) at time point TP4 compared to TP1. (C) Scatterplot of absolute transcript and footprint abundances measured across growth phase. Each dot represents the median \log_2 normalized counts over three biological replicates for a given RP transcript. Colors map to time points. R , correlation coefficient; a , slope. (D) Scatterplot and regression model of RP transcript and footprint abundance changes in time point TP4 compared to TP1. The black line represents regression model, the dark gray area represents regression confidence interval, and the light gray area represents the regression prediction interval. The red dot highlights *rpl14p* outside the prediction interval. R , correlation coefficient; a , slope.

ratios of \log_2 SR = -0.94 , \log_2 SR = -0.35 , and \log_2 SR = -0.30 , respectively. At least two of these RPs have also been critically implicated in regulatory functions: while not much is known about S10-like protein, S13 is a conserved protein across archaea, bacteria, and eukaryotes that controls mRNA-tRNA complex translocation (39–41), and L24E binds to translation initiation factor IF6, which is conserved in archaea and eukaryotes (42). Thus, protein composition and abundance of the pool of assembled ribosomes progressively change across different stages of growth. Importantly, five RPs occur in different stoichiometry within assembled ribosomes in stationary phase compared to early exponential phase. We predict that these five RPs play an important role in growth phase-dependent regulation of protein synthesis in *H. salinarum*.

Transcriptional and translational regulation of RP genes. We investigated whether transcriptional and translational regulatory mechanisms account for the differential RP stoichiometry we observed. Interestingly, changes in RP transcript levels appeared to be mediated by at least two distinct processes in stationary phase with one group of 29 transcripts (group A) with median stationary phase downregulation of \log_2 FC = -2.04 and a second group of 29 transcripts (group B) with \log_2 FC = -4.23 (blue dots in Fig. 3A). While multiple mechanisms contribute to transcript level changes (new transcription, RNA stability, targeted degradation, etc.), this bipartite transcriptional regulatory program is most likely an outcome of distinct transcriptional regulation of RP genes in group A and group B and not operon structure (Table S1, tabs 23 to 25, and Fig. S5 and S6).

To understand the implications of the bipartite transcriptional regulation, we analyzed transcript abundances, ribosomal footprints, and protein levels of RPs across different stages of growth. The bipartite grouping was apparent only in relative expression and not at the level of absolute expression or ribosomal footprints (Fig. 3B). There was significant correlation between change in abundance of RP transcripts and ribosomal footprints (slope = 0.64; $R = 0.77$; $P < 10^{-46}$; Fig. 3C). RP transcript abundance decreased as growth phase advanced—from 4,719 median normalized counts in early exponential (TP1) to 366 in stationary phase (TP4). Ribosomal footprints followed a similar trend, from 1,761 in TP1 to 217 in TP4. In fact, an ordinary least-squares regression model demonstrated that all footprint changes were within the expected prediction interval based on transcript level changes (slope = 0.70; $R = 0.88$; $P < 10^{-20}$; Fig. 3D). For only one instance, *rpl14p*, we observed ribosomal footprints significantly more abundant than expected by chance. While the *rpl14p* transcript level was downregulated, its ribosomal footprints did not change proportionally. Given *rpl14p* transcript levels and downregulation, we would expect a 26-fold downregulation (\log_2 FC = -4.70) of its footprint levels. However, *rpl14p* exhibits only a sevenfold downregulation (\log_2 FC = -2.78). Unexpectedly, we did not observe a significant deviation of L14 at the level of protein abundance and ribosomal stoichiometry, suggesting that deviation of changes in ribosomal footprints for *rpl14p* is likely a false-positive result, especially given that it is in the middle of a large 20-gene operon. Consistently for all RPs upon transition to stationary phase, downregulation was much more pronounced at the transcriptional (\log_2 FC = -3.68) and ribosomal footprint (\log_2 FC = -3.02) levels, relative to protein level (\log_2 FC = -1.04), possibly reflecting the well-known fact that proteins are more stable than mRNAs. Taken together, these observations demonstrate that relative changes in transcript abundance and TE of RP genes do not manifest at the protein level, thus maintaining the overall conserved stoichiometry of RPs within ribosomes.

Modular programs govern conditional regulation of RP transcription. Given the extensive growth phase-dependent mRNA changes in RP genes, we explored a whole-genome gene regulatory network of *H. salinarum* EGRIN model to determine whether regulation of translation system genes was governed by context-specific transcriptional programs. In brief, the EGRIN model of *H. salinarum* was constructed in two steps (43). First, the cMonkey algorithm (44, 45) was used to discover biclusters, which are sets of conditionally coregulated genes that share conserved gene regulatory elements (GREs) in their promoters. Second, regulators for each core were inferred using Inferelator (46, 47), which explains and predicts relative changes in expression levels of genes within each bicluster as a weighted sum of corresponding or preceding changes in transcriptional and environmental factors. We further developed an EGRIN2 model (25) which consists of an ensemble of EGRIN models—each constructed with a different data subset from a compendium of 1,495 transcriptome profiles from diverse environmental conditions. EGRIN2 models high-confidence associations among genes, environments, GREs, and regulators based on frequency of their cooccurrence within biclusters across all EGRIN models, resulting in corems. Corems are modular entities in EGRIN2 that capture with high confidence specific environmental context and regulatory mechanisms for coregulation of genes, and therefore, EGRIN2 represents an ideal framework to investigate condition-specific regulation of the translation machinery.

With the exception of four RP genes (*rps19e*, *rps27ae*, *rps24e*, and *rps12P*) that were not grouped into any core, 54 out of 58 annotated RP genes segregated into four classes based on applying hierarchical clustering on membership to 72 identified corems (Fig. 4A). These classes (corem membership information is detailed in Table S1, tabs 26 to 28), which we have labeled “class I-IV,” were also somewhat distinguished by operon and genome architecture, as would be expected from the inclusion of sequence features in the biclustering algorithm. For example, a large operon encoding 22 RPs, the RNase P protein subunit, the SecY translocation protein, and a putative RNA methyltransferase, was fully represented in two corems. In all other corems, this operon was

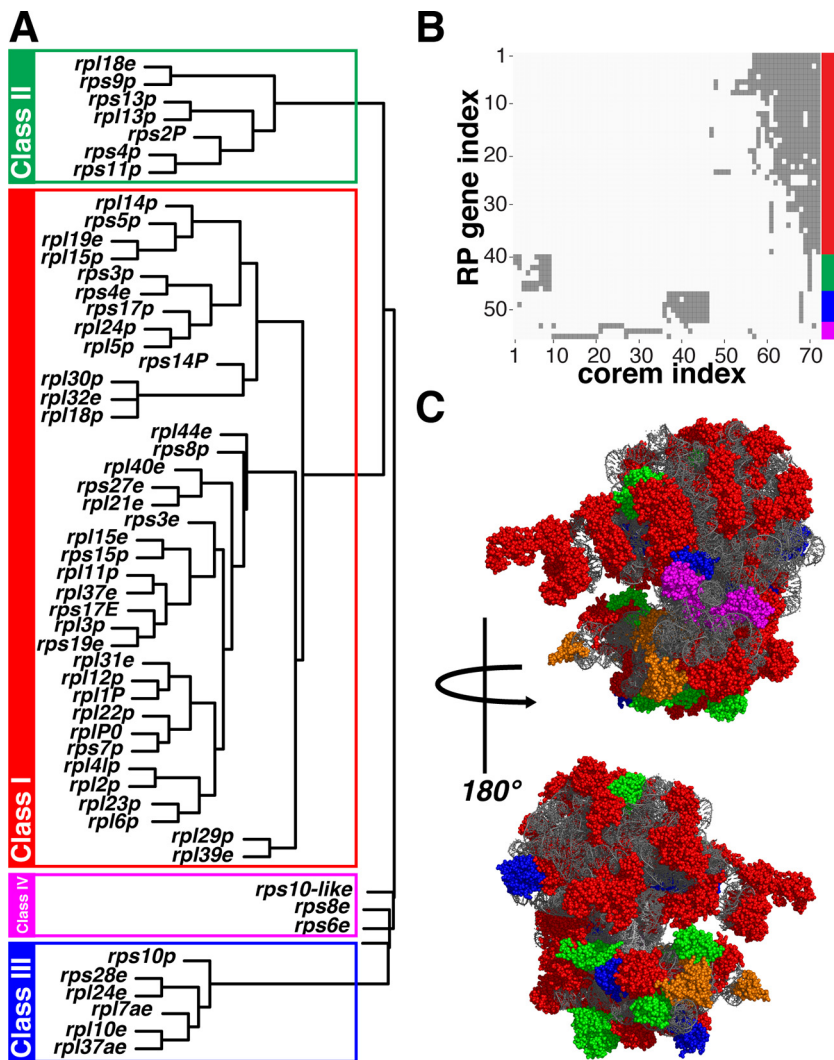


FIG 4 RP genes are organized into distinct functional coregulated classes. (A) Bootstrapped hierarchical clustering of RP genes based on core membership identified four classes, including three principal classes plus three outlier genes, joined together as class IV. Classes are boxed and colored with red, green, and blue, and outlier genes in magenta. (B) RP genes are depicted on the y axis versus corems on the x axis. Dark gray squares indicate the presence of a particular gene in a given core. The RP genes are arranged and colored by class on the right side. Corems comprise both neighboring genes in operons but also distal genes. (C) RPs are depicted in the ribosomal 3D structure following the color scheme of the functional classes as shown in panel A. Functional classes of RPs do not follow a restricted pattern of physical interactions, indicative of functional specialization of the ribosome due to coregulation in different environments, rather than coexpression derived from physical interactions at the protein level. Gray sections represent rRNA molecules. Subunits excluded (S12P, S19E, S24E, and S27AE) from the clustering analysis because they were not present in any core are depicted in orange.

fragmented, meaning that subsets of genes within the operon were conditionally split and differentially coregulated as transcript isoforms (48). All corems containing genes of this operon included different combinations of an additional 16 RP genes scattered throughout the genome (Fig. 4B). Further, these four classes extended to other translational proteins (Fig. S7).

In the largest group, class I, 38 RP genes—12 belonging to the small subunit and 26 to the large subunit—are broadly coregulated across 26 corems, albeit with substantial differences among the individual corems. Class II, consisting of seven RP genes (five small subunit and two large subunit) was coregulated in three corems, fractured in seven more corems, and altogether coregulated with additional genes. In the genome, the seven RP genes are physically located in two consecutive operons also containing

three RNA polymerase subunits. The functional coherence of this class is further emphasized by six out of the seven genes (*rps2P*, *rps4p*, *rps9p*, *rps11p*, *rps13p*, and *rpl13p*) being universal RP genes, with the lone exception of *rpl18e*. Class III consists of six RP genes (two small, four large subunits) across 12 corems, seven of which contained all six RP genes. Three of the RP genes are in an operon, while the other three are separated from the operon and each other on the chromosome. In contrast to class II, five RP genes (*rps28e*, *rpl7ae*, *rpl37ae*, *rpl10e*, and *rpl24e*) are specific to archaea and eukaryotes, with the lone exception being *rps10p*. Classes II and III are each coregulated with class I in single, discrete corems, further suggesting modular regulation of the ribosome as a whole. Three of the remaining RP genes (*rps6e*, *rps8e*, and *rps10*-like) show sparse association with each other and previous classes and were associated together as class IV. Two genes from the S10 ribosomal family, *rps10* and *rps10-like*, encode two ostensibly similar RPs that are divergently regulated. Further, *rps6e* encodes an RP that is located near the A-site of the ribosome, where it can interact with mRNA structures to regulate translation (Fig. 4C). Together, these findings suggest that specific conditional regulation of these proteins may lead to distinct ribosomes with functional specialization.

Evidence and mechanisms of condition-specific regulation of RPs. Once we determined that RPs associate with corems containing genes with other cellular functions, we hypothesized that *H. salinarum* might conditionally regulate RP composition of ribosomes in different environmental conditions. We explored evidence for context-specific regulation of RPs by analyzing expression coherence of genes within and across corems of the four classes under a comprehensive set of environmental conditions. We observed that genes within corems of the same class are coregulated across many conditions (Fig. 5A), but there is variability across classes (Fig. 5B). Notably, we discovered that variability across corems from different classes are condition specific. Specifically, for each broad category of conditions—growth in batch culture, shifts between high and low oxygen, response to different metals, etc.—we computed expression similarity between ribosomal corem classes, defined as the proportion of corem comparisons that had no significant gene expression differences (Fig. 5C). We conclude that the degree of variability in coregulation of ribosomal genes across the four classes of corems is strongly dependent on environmental context.

We hypothesize that differential coregulation of RPs across environmental conditions should be apparent in distinct gene regulatory elements defining their promoter architecture. Specifically, we hypothesized that identified RP corem classes should have distinct GREs defining their promoter architecture. The EGRIN2 model predicts specific mechanisms for transcriptional regulation for every gene in the genome; for instance, it predicts that *rps13p* (*VNG1132G*) is regulated by at least three transcription factors via binding to three distinct ~6- to 20-nucleotide GREs in the promoter of this gene (Fig. 5D and F). Furthermore, we demonstrated previously that EGRIN2 also accurately predicts which subset of GREs in the promoters of genes in corems are responsible for their environment-specific coregulation (25). Using this capability of EGRIN2, we investigated whether the observed variability in coexpression of RP genes across the corems of different classes was a consequence of different GRE composition. We performed hierarchical clustering on the composition of GREs implicated in coregulation of genes within corems enriched in RP genes (Fig. 5E). Distinct combinations of seven GREs are implicated in differential coregulation of ribosomal genes across the corems with at least one third of RP genes, supporting the hypothesis that distinct transcriptional regulatory mechanisms are responsible for the condition-specific variation in modular coregulation of RPs with one another and 614 other genes in the genome. We conclude from this analysis of the EGRIN2 model that variation in gene expression corresponds to environment-dependent modular regulation of ribosomal genes.

Coregulation of translational complexes in *E. coli* and *S. cerevisiae*. We investigated the generality of our findings from *H. salinarum* by analyzing the structure of conditional coregulation of RP genes in *E. coli* and *S. cerevisiae*. For these two organ-

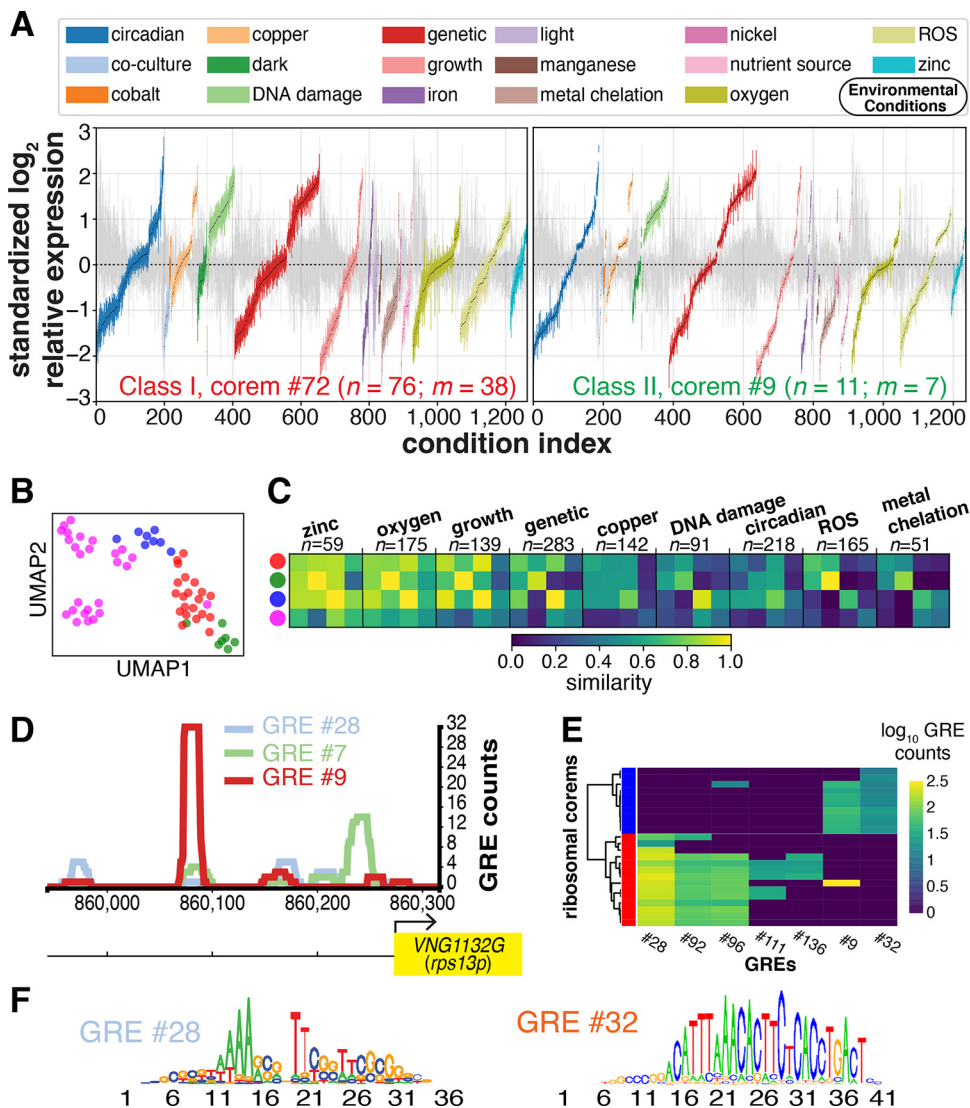


FIG 5 RP genes are coregulated in a condition-specific manner. (A) Expression distribution for two representative examples of the 72 ribosomal corems identified. Vertical bars represent interquartile range (third quartile [Q_3] to first quartile [Q_1]) of expression across genes. n indicates the number of genes in each corem, whereas m refers to the subset that encode RPs. Colors correspond to different environmental conditions (see key). Conditions are ranked based on median expression. Gray background bars correspond to average expression distribution of 10,000 permutations of randomly selected gene sets of the same size as the corem. Note the larger interquartile range of background distributions. (B) Corem similarity based on gene expression. We used uniform manifold approximation and projection (UMAP [77]) to visualize 62 ribosomal corems in a two-dimensional space from an original high-dimensional space of 1,495 environmental conditions. Each dot represents a corem. Color maps to functional classes. (C) Similarity matrix between ribosomal corem classes broken into nine different broad condition categories (n , number of experimental conditions). Similarity values across diagonals may not reach one, as we defined class similarity as the proportion of class corems with no significant expression differences. (D) Promoter architecture of *rps13p* (*VNG1132G*) deciphered by MAST (78) alignments of *cis*-regulatory motifs from each GRE, in this case GRE #7, #9, and #28 discovered by MEME (79) in gene promoters of EGRIN biclusters that include *rps13p*. The heights of the histograms are proportional to the frequency of GRE alignments to the *VNG1132G* promoter. (E) Hierarchical clustering of corems enriched in RP genes based on relative importance of each GRE (log₁₀ GRE counts) coregulating corem members. Red and blue bars correspond to class I and III, respectively. (F) Motif sequence logos for two representative GREs.

isms, the EGRIN models (25, 49) were mined in the same manner as for *H. salinarum*, and the resulting corems containing RPs, other translation factors, and the transcription apparatus were dissected. Analysis of the EGRIN models of *E. coli* and *S. cerevisiae* also showed that RP genes are organized into groups of corems, none of which encompasses all the genes in all circumstances. Rather, the set of RP genes is fractured into

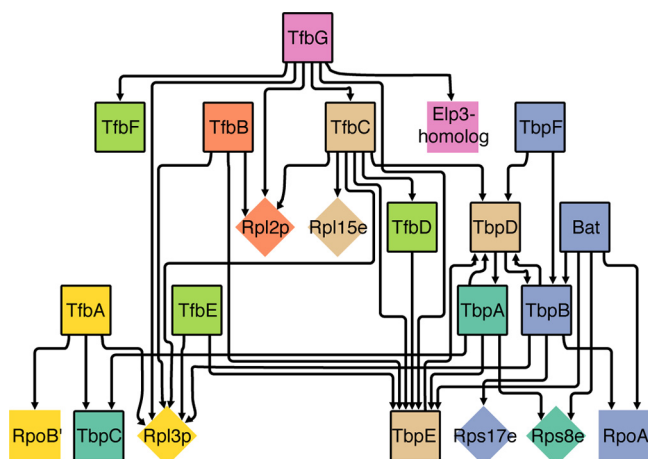


FIG 6 RPs physically interact with transcription complex components. Diamonds represent RPs; squares represent transcription complex components. Tagged proteins used as bait in the immunoprecipitation experiment are highlighted by a black border. Arrowheads link bait to coimmunoprecipitated proteins. We labeled each of the seven modules obtained by the Newman-Girvan clustering algorithm using a different color.

multiple mostly discrete, but partially overlapping, corems (Fig. S8). This indicates that ribosome composition regulation in all three organisms is less of a singular entity, and more of a mosaic patchwork. Additionally, subsets of translation system genes associated with tRNA charging and translation factors (including aminoacyl-tRNA synthetases, initiation, elongation, and release factors), as well as RNA handling (RNases and RNA modification enzymes) are split among the corems. This pattern was apparent in all three organisms tested, suggesting a conserved pattern in all domains of life. Since the conservation in regulatory network architecture stems from gene expression, in addition to sequence features, it is not simply a genome sequence comparison, but it represents a correspondence in active physiology of organisms in relation to particular environmental conditions. Thus, these findings are consistent with the idea that functional specialization drives modularity in biological systems. Such modular regulation has evolutionary implications, as modularity would facilitate the evolvability of the translational complexes (27).

Physical protein-protein interactions support coupled transcription-translation.

Physical interaction of the RNA polymerase (RNAP) with the 30S ribosomal subunit in prokaryotes (50–52) suggests that actively transcribed genes also actively recruit ribosomes for coupled translation. We investigated the evidence for a similar phenomenon in archaea by analyzing a protein-protein interaction map of *H. salinarum* constructed through immunoprecipitation of 14 protein A-tagged transcription complex components (22). In brief, 13 general transcription factors (GTFs)—six TATA-binding proteins (TBP) and seven transcription factor B proteins (TFB)—and bacterioopsin activator (Bat) were epitope tagged with protein A, and used as bait to immunoprecipitate *H. salinarum* transcriptional complexes in 14 independent experiments, performed in duplicate. We constructed a network of 128 proteins (as nodes) and 228 interactions, i.e., unidirectional edges from tagged baits to coimmunoprecipitated proteins that rendered seven modules (see Materials and Methods). This network presents 13 physical interactions between eight components of the transcriptional machinery and five ribosomal proteins (Fig. 6). Interestingly, while TFBS associated exclusively with the ribosome large subunit, TBPs preferentially interacted with the small ribosomal subunit. We further investigated whether coimmunoprecipitated RPs form an interacting interface in the ribosome. We explored their physical location in the ribosome three-dimensional (3D) structure. We found that while three of the five ribosomal proteins are scattered across the ribosome surface, L2 and L15E are particularly close to each other (Fig. S9). Nevertheless, functional implications of this observation require further inves-

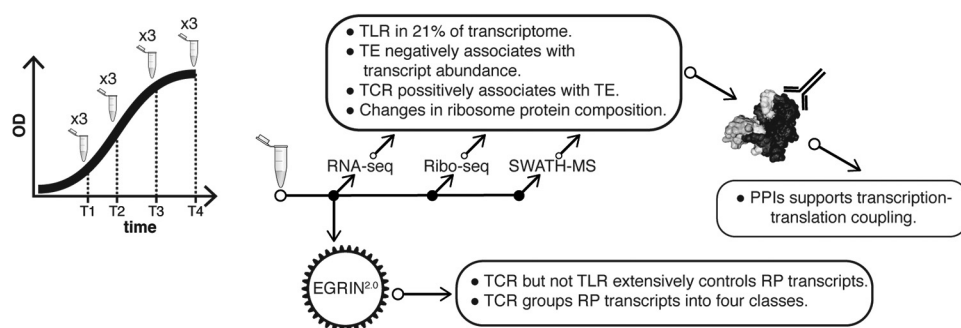


FIG 7 Synthesis figure. Growth phase *H. salinarum* transcriptome abundance, ribosomal footprints, and proteome quantification and analysis identified key points of translational regulation. TLR, translational regulation; TE, translational efficiency; TCR, transcriptional regulation; PPI, protein-protein interaction; OD, optical density; RP, ribosomal protein.

tigation. Physical interactions between transcription and translation complexes have been previously implicated as the mechanism by which transcription and translation are coupled in prokaryotes, and here it gives a mechanistic hypothesis for why actively transcribed genes are preferentially translated. Moreover, this is consistent with the evidence we have provided that low abundance and upregulated transcripts are preferentially translated in both *E. coli* and *H. salinarum*, but not in yeast, because of physical separation of the two processes (Fig. S4).

DISCUSSION

Here, we have demonstrated that preferential translation of low abundance and upregulated transcripts influences growth-associated physiological state shifts in *H. salinarum*. Characterizing the precise mechanism underlying this phenomenon has important implications for understanding how cells transition to a physiological state appropriate for a resource-limited environment (anoxia, nutrient starvation, etc.), which requires preferential and efficient translation of low abundance transcripts (Fig. 7).

Prior studies showed that highly expressed transcripts were upregulated and functionally required for processes such as aerobic respiration, ATP synthesis, tricarboxylic acid (TCA) cycle, transcription and translation, during oxic growth of *H. salinarum*—when the oxygen level dropped, the haloarchaeon adopted a quiescent state and these transcripts were downregulated, but they continued to persist in high abundance, even though they were functionally irrelevant for anaerobic physiology (23). Importantly, protein levels of these downregulated transcripts also decreased during anoxia. In other words, even though these transcripts were present in high abundance, they were not actively translated in an anoxic environment. Further, *H. salinarum* reinitiated translation of these persistent highly abundant transcripts almost concurrently with increase in oxygen level, and well before their transcriptional upregulation (23). The current study proposes a mechanistic explanation for these classic observations.

Bernstein et al. (53) have demonstrated in *E. coli* that transcription initiation may be the dominant factor in determining mRNA steady-state levels in the cell, while mRNA decay might serve as a mechanism to respond rapidly to environmental changes. We hypothesize that most low abundance transcripts are transcribed constitutively at a low rate that is proportional to or less than their degradation rate, and therefore, they are associated with higher TE relative to highly abundant transcripts, which are expressed at a high level only in environments where their functions are relevant. Although outside the scope of this study, this hypothesis can be experimentally tested with GRO-seq and pulse-chase experiments (54, 55). Nonetheless, these observations are also consistent with both our finding that all transcriptionally upregulated genes have higher TE relative to downregulated transcripts, and the Schmid et al. finding that highly abundant transcripts are not translated in an unfavorable environment when

they are transcriptionally downregulated (23). Alternate mechanisms can also explain why some upregulated transcripts are selectively translated in a growth phase-dependent manner. For examples, posttranslational modifications such as differential ubiquitination in eukaryotes can alter the stability of proteins to regulate ribosomal function such as by stalling assembled ribosomes and blocking access to initiation factors (56). While archaea do not possess the classical ubiquitination pathway, they do utilize hypusination to stop translation and growth via initiation factor aIF5A (57).

At a mechanistic level, the preferential translation of low abundance and upregulated transcript can be explained by the well-known phenomenon of coupled transcription and translation in prokaryotes (58, 59). At the molecular level, results presented in this study (Fig. 6) and previous reports for bacteria (50–52) and archaea (60) have demonstrated that the transcription machinery facilitates the recruitment of translation factors and the ribosome through physical protein-protein interactions. While this mechanism of transcription-translation coupling is pertinent to 875 *H. salinarum* genes that are regulated just at the transcriptional level, it is noteworthy that at least 228 genes (21% of all regulated genes) are subject to translational regulation during physiological state transitions. We observed two orthogonal modes of translational regulation: (i) changes in ribosomal footprints in fixed-abundance transcripts, and (ii) compensatory mechanisms, where ribosomal footprints remained unchanged in spite of transcript abundance changes. These translationally regulated genes encode a wide range of critical functions that include amino acid and lipid metabolism, DNA replication and repair, transcription regulation and energy homeostasis. This result has major implications on understanding physiological state transitions in archaea, as it has been already noted in human disease physiology (7, 61).

To that end, we discovered that regulation of translation systems is heterogeneous, and modular. Notably, these ribosomal modules (corems) also include 561 additional genes of diverse functions, including transcription, metabolism, signal transduction, and transmembrane transport, suggesting that regulation of components of the translation system is coordinated with the expression of diverse functions. This modular regulation of the translational machinery could provide a basis for specialization and adaptive evolution. There is extensive evidence that specialization leads to the emergence of modularity in biological systems, including in metabolic and transcriptional regulatory networks. There are at least two reasons why modularity facilitates the ability of an organism to generate adaptive heritable variation, i.e., evolvability. First, an organism can select variations inside one module, without perturbing other modules. Second, there is also evidence that modules can be repurposed or merged to generate novel functions (26, 62, 63). Despite the complex modular regulation of RP genes, we observed the expected and coordinated decrease in transcript and protein abundance of RPs when cells shifted to stationary phase. Concordantly, the majority of RPs in the ribosome-enriched fraction also decreased in a coherent manner, with very similar changes in relative abundance. These observations suggest that regulation of RP genes at the transcriptional and translational level has evolved to maintain stoichiometry of protein subunits within the assembled ribosome. However, ribosomal associations of five RPs significantly deviated upon transition to stationary phase, most likely driven by a posttranscriptional mechanism, and indicative of a pool of ribosomes with different RP stoichiometry that is responsible for variation in TE across physiological states (4, 18, 64). While further experimental validation is needed to demonstrate that ribosomes of distinct composition selectively translate specific sets of transcripts, our analysis of transcriptional regulatory networks of *E. coli* and yeast suggests that modular regulation and coordination of ribosomal proteins with other cellular functions are generalizable phenomena that underlie environment-specific adaptation and specialization of all organisms.

MATERIALS AND METHODS

Cell culture and sampling. Wild-type *Halobacterium salinarum* NRC-1 was cultured in a liquid nutrient-rich complex medium (CM) (250 g/liter NaCl, 20 g/liter MgSO₄·7H₂O, 3 g/liter sodium citrate, 2

g/liter KCl, and 10 g/liter peptone [Oxoid, United Kingdom] made with distilled water). Cultures were inoculated to a starting optical density at 600 nm (OD_{600}) of 0.02 with starter culture with an OD_{600} of 0.5 which was derived from a single colony. Cultures were grown in un baffled flasks in which 40% of the flask volume is occupied by the culture. Cultures were grown at 37°C, shaken at 220 rpm, and illuminated at $\sim 20 \mu\text{mol}/\text{m}^2/\text{s}$ in Innova9400 incubators (New Brunswick). Triplicate cultures were grown, and samples were harvested at four time points. The four time points were selected to represent the early exponential phase (OD_{600} of 0.2; 14.3 h), mid-exponential growth (OD_{600} of 0.5; 21.5 h), late exponential phase (OD_{600} of 0.8; 28.8 h), and stationary phase (40.8 h). The final time point was selected to be 12 h past the late exponential phase, since OD_{600} readings are not representative of cell growth in *H. salinarum* in stationary phase (24). At each time point, whole cells were collected by centrifugation for analysis by RNA sequencing, ribosome profiling, and mass spectrometry (MS) proteomics.

RNA-seq and ribosome profiling analysis. Cells were pelleted by centrifugation ($8,000 \times g$, 2 min, 4°C), resuspended in a buffer containing 3.4 M KCl, 100 mM MgCl_2 , and 10 mM Tris-HCl at pH 7.4, sonicated at 4°C to lyse cells (amplitude 50%, pulse 30 s on and 15 s off, repeated 6 times), and centrifuged again at $14,000 \times g$ for 10 min at 4°C to remove cell debris. Supernatants were collected and treated with RQ1 DNase (Promega), followed by centrifugation ($14,000 \times g$, 10 min, 4°C). Ribosome-bound RNA was generated by treating the lysate with RNase I and quenching the reaction with Superase-In RNase inhibitor (see Fig. S1A in the supplemental material). Macromolecular complexes were collected by spin column isolation (MicroSpin S-400 HR; GE), elution was performed by centrifugation ($600 \times g$, 2 min, room temperature), and samples were snap-frozen in liquid nitrogen and stored at -80°C . The elution sample was split into two aliquots, one for ribosome footprint sequencing and one for proteome analysis. For transcriptome sequencing, total RNA was collected from the cell lysate using TRIzol-chloroform extraction and elution with water. A total of 24 barcoded libraries were prepared for sequencing; 12 using the TruSeq Stranded mRNA HT library prep kit for mRNA, and 12 using the NEBNext Small RNA Library Prep Set from Illumina for the ribosome-bound fragments. Libraries were pooled, denatured, and diluted according to the NextSeq 500 protocol. Single-end sequencing of libraries was performed on the Illumina NextSeq 500 platform using two high-output flow cells with 75-bp read lengths. Adapter sequences were trimmed using Trimmomatic (65). Transcript abundance and ribosomal footprint quantification in the form of transcripts per million (TPM) was performed using kallisto (66) against a reference transcriptome of 2,665 open reading frames (ORFs). Differential gene expression analysis was performed using DESeq2 (67) (after HTSeq [68] and STAR [69]).

Assembled ribosome protein analysis. As described above, macromolecular complex isolation spin column (MicroSpin S-400 HR; GE), elution samples (enriched fractions) were obtained, snap-frozen in liquid nitrogen, and stored at -80°C . The protein content of the samples was determined by bicinchoninic acid assay (Thermo-Fisher). Proteins were reduced (5 mM dithiothreitol, 45 min, 37°C), alkylated (14 mM iodoacetamide, 30 min, room temperature, darkness), and digested with trypsin (1:50 enzyme/substrate ratio, 37°C, 16 h). Samples were desalted with tC18 SepPak cartridges (Waters). Samples were analyzed with a TripleTOF 5600+ system equipped with a Nanospray-III source (Sciex) and an Eksigent Eksport nanoLC 425 with cHiPLC system in trap-elute mode (Sciex). Peptides were separated with a gradient from 3% to 33% 0.1% formic acid in acetonitrile (vol/vol) in 120 min. Data were collected in MS/MS^{ALL} SWATH acquisition mode using 100 variable acquisition windows. Data were analyzed with the OneOmics SWATH Proteomics Toolkit (Sciex) within the BaseSpace cloud computing environment (Illumina). An ion library was generated from *H. salinarum* grown to mid-exponential and stationary phase acquired in shotgun mode (information-dependent acquisition scanning of mass spectrometry performed in tandem [IDA-MS/MS]) with the TripleTOF 5600+ system. A confidence filter of $\geq 75\%$ (statistically significant differentially expressed proteins) was applied to report protein expression changes.

Gene ontology analysis and visualization. Gene ontology (GO) annotations for each *H. salinarum* gene were obtained from MicrobesOnline (70). We used the Bioconductor package topGO (71) to discover significantly enriched GO terms in gene sets of interest. We used REVIGO (72) to summarize and visualize enriched GO terms.

Cluster analysis. Corems were identified based on an extensive pipeline that was previously published (25). To group corems by similarity of gene content, hierarchical agglomerative clustering of genes based on presence or absence in a corem was performed in R. The method used to create the distance matrix therefore was binary, and the clustering algorithm was average similarity. The clusters were bootstrapped using the package pvclust 10,000 times, and maximal clusters with $>95\%$ significant *P* values were selected, resulting in four robust classes. Genes that were not present in any corem were excluded from the analysis. In order to compare two given EGRIN2 corems, we computed both Spearman's rank correlation coefficient (SRCC) and Kolmogorov-Smirnov test (KST) on their expression signatures over all conditions. If correlation was positive and significant ($\text{SRCC} > 0.4$ and $P < 0.05$) and KST was nonsignificant ($P > 0.05$), we considered that the corems had no significant expression differences.

Structural modeling. To obtain a visual sense of the RP classes, subunits were analyzed in PyMol. The *Haloarcula marismortui* large ribosomal subunit (PDB 4V9F) was aligned to the large ribosomal subunit of the archaeon *Pyrococcus furiosus* (PDB 4V6U). The *P. furiosus* large subunit was then hidden, while the small subunit was kept. rRNA structures were colored gray, and the RPs were colored according to their class or lack thereof.

Protein-protein interaction network analysis. Protein interaction data were retrieved from Supplementary Material Data Set 3 from Facciotti et al. (22). We removed duplicate entries. We retrieved protein annotation from NCBI Assembly (ASM680v1; RefSeq annotation), MicrobesOnline (73), and

the *Halobacterium salinarum* NRC-1 SBEAMS database (<https://baliga.systemsbiology.net/projects/halobacterium-species-nrc-1-genome>). We used the Newman-Girvan algorithm (74) implemented in clusterMaker2 (75) for Cytoscape (76) version 3.7.2 to call network modules. To highlight interactions between general transcription factors and ribosome proteins, we hid all the nodes and edges not connected to them, and applied Cytoscape yFiles Hierarchic Layout. We minimally shifted the position of a few nodes to improve network legibility.

Data availability. RNA sequencing data have been deposited into NCBI SRA under BioProject number PRJNA413990. Mass spectrometry data are available in the PeptideAtlas data repository: <http://www.peptideatlas.org/PASS/PASS01559>. All code implementation, including sequence quantification and EGRIN model analyses, is available at the GitHub repository: <https://github.com/adelomana/30sols>.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, TIF file, 2.5 MB.

FIG S2, TIF file, 2.2 MB.

FIG S3, TIF file, 2.7 MB.

FIG S4, TIF file, 1.6 MB.

FIG S5, TIF file, 2.7 MB.

FIG S6, TIF file, 2.6 MB.

FIG S7, TIF file, 2.6 MB.

FIG S8, TIF file, 2.8 MB.

FIG S9, TIF file, 2.6 MB.

TABLE S1, XLSX file, 1.4 MB.

ACKNOWLEDGMENTS

This project was supported by the U.S. National Science Foundation (grants MSB-1237267, DB-1262637, DB-1565166, MCB-1330912, and MCB-1616955) and the U.S. National Institutes of Health (grants 2P50GM076547, R01GM087221, R01AI128215, and R01AI141953). A.P.R.L. was supported by grants #2017/03052-2 and #2019/13440-5 from the São Paulo Research Foundation (FAPESP).

We thank Aaron Hudson, Mark Cafazzo, and Christie Hunter (Sciex) for access to a 5600+ SWATH-MS-enabled TripleTOF MS instrument. We thank Wei-Ju Wu for technical assistance with EGRIN models.

A.V.R., U.K., R.L.M., and N.S.B. conceived and planned the experiments. A.L.G.D.L. performed data analysis, results interpretation, and figure construction with respect to all but Fig. 4 and 6. A.V.R. performed data analysis, results interpretation, and figure construction with respect to Fig. 4. A.P.R.L. performed data analysis, results interpretation, and figure construction with respect to Fig. 6. S.T. analyzed data with respect to the EGRIN model. A.V.R., M.P., and U.K. performed the experiments. R.L.M. and N.S.B. supervised the project. A.V.R., A.L.G.D.L., and N.S.B. drafted and wrote the majority of the manuscript. All authors discussed the results and edited the manuscript.

REFERENCES

- Genuth NR, Barna M. 2018. Heterogeneity and specialized functions of translation machinery: from genes to organisms. *Nat Rev Genet* 19: 431–452. <https://doi.org/10.1038/s41576-018-0008-z>.
- Sauert M, Temmel H, Moll I. 2015. Heterogeneity of the translational machinery: variations on a common theme. *Biochimie* 114:39–47. <https://doi.org/10.1016/j.biochi.2014.12.011>.
- Venezia ND, Vincent A, Marcel V, Catez F, Diaz JJ. 2019. Emerging role of eukaryote ribosomes in translational control. *Int J Mol Sci* 20:1226. <https://doi.org/10.3390/ijms20051226>.
- Shi Z, Fujii K, Kovary KM, Genuth NR, Röst HL, Teruel MN, Barna M. 2017. Heterogeneous ribosomes preferentially translate distinct subpools of mRNAs genome-wide. *Mol Cell* 67:71–83.e7. <https://doi.org/10.1016/j.molcel.2017.05.021>.
- Kondrashov N, Pusic A, Stumpf CR, Shimizu K, Hsieh AC, Xue S, Ishijima J, Shiroishi T, Barna M. 2011. Ribosome-mediated specificity in Hox mRNA translation and vertebrate tissue patterning. *Cell* 145:383–397. <https://doi.org/10.1016/j.cell.2011.03.028>.
- Xue S, Tian S, Fujii K, Kladwang W, Das R, Barna M. 2015. RNA regulons in Hox 5' UTRs confer ribosome specificity to gene regulation. *Nature* 517:33–38. <https://doi.org/10.1038/nature14010>.
- Schafer S, Adami E, Heinig M, Rodrigues KEC, Kreuchwig F, Silhavy J, van Heesch S, Simaite D, Rajewsky N, Cuppen E, Pravenec M, Vingron M, Cook SA, Hubner N. 2015. Translational regulation shapes the molecular landscape of complex disease phenotypes. *Nat Commun* 6:7200. <https://doi.org/10.1038/ncomms8200>.
- Zhang X, Gao X, Coots RA, Conn CS, Liu B, Qian S-B. 2015. Translational control of the cytosolic stress response by mitochondrial ribosomal protein L18. *Nat Struct Mol Biol* 22:404–410. <https://doi.org/10.1038/nsmb.3010>.
- Slavov N, Semrau S, Airoidi E, Budnik B, van Oudenaarden A. 2015. Differential stoichiometry among core ribosomal proteins. *Cell Rep* 13: 865–873. <https://doi.org/10.1016/j.celrep.2015.09.056>.
- Thompson MK, Rojas-Duran MF, Gangaramani P, Gilbert WV. 2016. The ribosomal protein Asc1/RACK1 is required for efficient translation of short mRNAs. *Elife* 5:e11154. <https://doi.org/10.7554/eLife.11154>.
- Ferretti MB, Ghalei H, Ward EA, Potts EL, Karbstein K. 2017. Rps26 directs

- mRNA-specific translation by recognition of Kozak sequence elements. *Nat Struct Mol Biol* 24:700–707. <https://doi.org/10.1038/nsmb.3442>.
12. Ferretti MB, Barre JL, Karbstein K. 2018. Translational reprogramming provides a blueprint for cellular adaptation. *Cell Chem Biol* 25:1372–1379.e3. <https://doi.org/10.1016/j.chembiol.2018.08.003>.
 13. Komili S, Farny NG, Roth FP, Silver PA. 2007. Functional specificity among ribosomal proteins regulates gene expression. *Cell* 131:557–571. <https://doi.org/10.1016/j.cell.2007.08.037>.
 14. Byrgazov K, Vesper O, Moll I. 2013. Ribosome heterogeneity: another level of complexity in bacterial translation regulation. *Curr Opin Microbiol* 16:133–139. <https://doi.org/10.1016/j.mib.2013.01.009>.
 15. Kurylo CM, Parks MM, Juette MF, Zinshteyn B, Altman RB, Thibado JK, Vincent CT, Blanchard SC. 2018. Endogenous rRNA sequence variation can regulate stress response gene expression and phenotype. *Cell Rep* 25:236–248.e6. <https://doi.org/10.1016/j.celrep.2018.08.093>.
 16. Culviner PH, Laub MT. 2018. Global analysis of the *E. coli* toxin MazF reveals widespread cleavage of mRNA and the inhibition of rRNA maturation and ribosome biogenesis. *Mol Cell* 70:868–880.e10. <https://doi.org/10.1016/j.molcel.2018.04.026>.
 17. Ferretti MB, Karbstein K. 2019. Does functional specialization of ribosomes really exist? *RNA* 25:521–538. <https://doi.org/10.1261/rna.069823.118>.
 18. Xue S, Barna M. 2012. Specialized ribosomes: a new frontier in gene regulation and organismal biology. *Nat Rev Mol Cell Biol* 13:355–369. <https://doi.org/10.1038/nrm3359>.
 19. Hering O, Brenneis M, Beer J, Suess B, Soppa J. 2009. A novel mechanism for translation initiation operates in haloarchaea. *Mol Microbiol* 71:1451–1463. <https://doi.org/10.1111/j.1365-2958.2009.06615.x>.
 20. Daume M, Uhl M, Backofen R, Randau L. 2017. RIP-Seq suggests translational regulation by L7Ae in *Archaea*. *mBio* 8:e00730-17. <https://doi.org/10.1128/mBio.00730-17>.
 21. Storz G, Hengge R (ed). 2011. Bacterial stress responses, 2nd ed. American Society for Microbiology, Washington, DC.
 22. Facciotti MT, Reiss DJ, Pan M, Kaur A, Vuthoori M, Bonneau R, Shannon P, Srivastava A, Donohoe SM, Hood LE, Baliga NS. 2007. General transcription factor specified global gene regulation in archaea. *Proc Natl Acad Sci U S A* 104:4630–4635. <https://doi.org/10.1073/pnas.0611663104>.
 23. Schmid AK, Reiss DJ, Kaur A, Pan M, King N, Van PT, Hohmann L, Martin DB, Baliga NS. 2007. The anatomy of microbial cell state transitions in response to oxygen. *Genome Res* 17:1399–1413. <https://doi.org/10.1101/gr.6728007>.
 24. Facciotti MT, Pang WL, Lo F, Whitehead K, Koide T, Masumura K, Pan M, Kaur A, Larsen DJ, Reiss DJ, Hoang L, Kalisiak E, Northen T, Trauger SA, Siuzdak G, Baliga NS. 2010. Large scale physiological readjustment during growth enables rapid, comprehensive and inexpensive systems analysis. *BMC Syst Biol* 4:64. <https://doi.org/10.1186/1752-0509-4-64>.
 25. Brooks AN, Reiss DJ, Allard A, Wu W-J, Salvanha DM, Plaisier CL, Chandrasekaran S, Pan M, Kaur A, Baliga NS. 2014. A system-level model for the microbial regulatory genome. *Mol Syst Biol* 10:740. <https://doi.org/10.15252/msb.20145160>.
 26. Wagner GP, Altenberg L. 1996. Complex adaptations and the evolution of evolvability. *Evolution* 50:967–976. <https://doi.org/10.2307/2410639>.
 27. Espinosa-Soto C, Wagner A. 2010. Specialization can drive the evolution of modularity. *PLoS Comput Biol* 6:e1000719. <https://doi.org/10.1371/journal.pcbi.1000719>.
 28. Ingolia NT, Lareau LF, Weissman JS. 2011. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* 147:789–802. <https://doi.org/10.1016/j.cell.2011.10.002>.
 29. Lange C, Zaigler A, Hammelmann M, Twellmeyer J, Raddatz G, Schuster SC, Oesterheld T, Soppa J. 2007. Genome-wide analysis of growth phase-dependent translational and transcriptional regulation in halophilic archaea. *BMC Genomics* 8:415. <https://doi.org/10.1186/1471-2164-8-415>.
 30. McManus CJ, May GE, Spealman P, Shteyman A. 2014. Ribosome profiling reveals post-transcriptional buffering of divergent gene expression in yeast. *Genome Res* 24:422–430. <https://doi.org/10.1101/gr.164996.113>.
 31. Schmier BJ, Nelersa CM, Malhotra A. 2017. Structural basis for the bidirectional activity of *Bacillus* nanoRNase NrnA. *Sci Rep* 7:11085. <https://doi.org/10.1038/s41598-017-09403-x>.
 32. Hundt S, Zaigler A, Lange C, Soppa J, Klug G. 2007. Global analysis of mRNA decay in *Halobacterium salinarum* NRC-1 at single-gene resolution using DNA microarrays. *J Bacteriol* 189:6936–6944. <https://doi.org/10.1128/JB.00559-07>.
 33. Li GW, Burkhardt D, Gross C, Weissman JS. 2014. Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell* 157:624–635. <https://doi.org/10.1016/j.cell.2014.02.033>.
 34. Kasari V, Margus T, Atkinson GC, Johansson MJO, Hauryluk V. 2019. Ribosome profiling analysis of eEF3-depleted *Saccharomyces cerevisiae*. *Sci Rep* 9:3037. <https://doi.org/10.1038/s41598-019-39403-y>.
 35. Yu H, Rao X, Zhang K. 2017. Nucleoside diphosphate kinase (Ndk): a pleiotropic effector manipulating bacterial virulence and adaptive responses. *Microbiol Res* 205:125–134. <https://doi.org/10.1016/j.micres.2017.09.001>.
 36. Malygin AA, Karpova GG. 2010. Site-specific cleavage of the 40S ribosomal subunit reveals eukaryote-specific ribosomal protein S28 in the subunit head. *FEBS Lett* 584:4396–4400. <https://doi.org/10.1016/j.febslet.2010.10.013>.
 37. Schroeder SJ, Blaha G, Tirado-Rives J, Steitz TA, Moore PB. 2007. The structures of antibiotics bound to the E site region of the 50S ribosomal subunit of *Haloarcula marismortui*: 13-deoxytetracycline and girodazole. *J Mol Biol* 367:1471–1479. <https://doi.org/10.1016/j.jmb.2007.01.081>.
 38. Armache J-P, Jarasch A, Anger AM, Villa E, Becker T, Bhushan S, Jossinet F, Habeck M, Dindar G, Franckenberg S, Marquez V, Mielke T, Thomm M, Berninghausen O, Beatrix B, Söding J, Westhof E, Wilson DN, Beckmann R. 2010. Localization of eukaryote-specific ribosomal proteins in a 5.5-Å cryo-EM map of the 80S eukaryotic ribosome. *Proc Natl Acad Sci U S A* 107:19754–19759. <https://doi.org/10.1073/pnas.1010005107>.
 39. Kukras AR, Southworth DR, Brunelle JL, Culver GM, Green R. 2003. Ribosomal proteins S12 and S13 function as control elements for translocation of the mRNA:tRNA complex. *Mol Cell* 12:321–328. [https://doi.org/10.1016/s1097-2765\(03\)00275-2](https://doi.org/10.1016/s1097-2765(03)00275-2).
 40. Schuwirth BS, Borovinskaya MA, Hau CW, Zhang W, Vila-Sanjurjo A, Holton JM, Cate J. 2005. Structures of the bacterial ribosome at 3.5 Å resolution. *Science* 310:827–834. <https://doi.org/10.1126/science.1117230>.
 41. Zhang W, Dunkle JA, Cate J. 2009. Structures of the ribosome in intermediate states of ratcheting. *Science* 325:1014–1017. <https://doi.org/10.1126/science.1175275>.
 42. Greber BJ, Boehringer D, Godinic-Mikulcic V, Crnkovic A, Ibba M, Weygand-Durasevic I, Ban N. 2012. Cryo-EM structure of the archaeal 50S ribosomal subunit in complex with initiation factor 6 and implications for ribosome evolution. *J Mol Biol* 418:145–160. <https://doi.org/10.1016/j.jmb.2012.01.018>.
 43. Bonneau R, Facciotti MT, Reiss DJ, Schmid AK, Pan M, Kaur A, Thorsson V, Shannon P, Johnson MH, Bare JC, Longabaugh W, Vuthoori M, Whitehead K, Madar A, Suzuki L, Mori T, Chang D-E, Diruggiero J, Johnson CH, Hood L, Baliga NS. 2007. A predictive model for transcriptional control of physiology in a free living cell. *Cell* 131:1354–1365. <https://doi.org/10.1016/j.cell.2007.10.053>.
 44. Reiss DJ, Baliga NS, Bonneau R. 2006. Integrated biclustering of heterogeneous genome-wide datasets for the inference of global regulatory networks. *BMC Bioinformatics* 7:280. <https://doi.org/10.1186/1471-2105-7-280>.
 45. Reiss DJ, Plaisier CL, Wu W-J, Baliga NS. 2015. cMonkey2: automated, systematic, integrated detection of co-regulated gene modules for any organism. *Nucleic Acids Res* 43:e87. <https://doi.org/10.1093/nar/gkv300>.
 46. Bonneau R, Reiss DJ, Shannon P, Facciotti M, Hood L, Baliga NS, Thorsson V. 2006. The Inferelator: an algorithm for learning parsimonious regulatory networks from systems-biology data sets de novo. *Genome Biol* 7:R36. <https://doi.org/10.1186/gb-2006-7-5-r36>.
 47. Madar A, Greenfield A, Ostrer H, Vanden-Eijnden E, Bonneau R. 2009. The Inferelator 2.0: a scalable framework for reconstruction of dynamic regulatory network models. *Conf Proc IEEE Eng Med Biol Soc* 2009:5448–5451. <https://doi.org/10.1109/IEMBS.2009.5334018>.
 48. Koide T, Reiss DJ, Bare JC, Pang WL, Facciotti MT, Schmid AK, Pan M, Marzolf B, Van PT, Lo F-Y, Pratap A, Deutsch EW, Peterson A, Martin D, Baliga NS. 2009. Prevalence of transcription promoters within archaeal operons and coding sequences. *Mol Syst Biol* 5:285. <https://doi.org/10.1038/msb.2009.42>.
 49. Danziger SA, Ratushny AV, Smith JJ, Saleem RA, Wan Y, Arens CE, Armstrong AM, Sitko K, Chen W-M, Chiang J-H, Reiss DJ, Baliga NS, Aitchison JD. 2014. Molecular mechanisms of system responses to novel stimuli are predictable from public data. *Nucleic Acids Res* 42:1442–1460. <https://doi.org/10.1093/nar/gkt938>.
 50. Demo G, Rasouly A, Vasilyev N, Svetlov V, Loveland AB, Diaz-Avalos R, Grigorieff N, Nudler E, Korostelev AA. 2017. Structure of RNA polymerase

- bound to ribosomal 30S subunit. *Elife* 6:e28560. <https://doi.org/10.7554/eLife.28560>.
51. Kohler R, Mooney RA, Mills DJ, Landick R, Cramer P. 2017. Architecture of a transcribing-translating expressome. *Science* 356:194–197. <https://doi.org/10.1126/science.aal3059>.
 52. Fan H, Conn AB, Williams PB, Diggs S, Hahm J, Gamper HB, Hou Y-M, O'Leary SE, Wang Y, Blaha GM. 2017. Transcription-translation coupling: direct interactions of RNA polymerase with ribosomes and ribosomal subunits. *Nucleic Acids Res* 45:11043–11055. <https://doi.org/10.1093/nar/gkx719>.
 53. Bernstein JA, Khodursky AB, Lin PH, Lin-Chao S, Cohen SN. 2002. Global analysis of mRNA decay and abundance in *Escherichia coli* at single-gene resolution using two-color fluorescent DNA microarrays. *Proc Natl Acad Sci U S A* 99:9697–9702. <https://doi.org/10.1073/pnas.112318199>.
 54. Lopes R, Agami R, Korkmaz G. 2017. GRO-seq, a tool for identification of transcripts regulating gene expression. *Methods Mol Biol* 1543:43–55. https://doi.org/10.1007/978-1-4939-6716-2_3.
 55. Lugowski A, Nicholson B, Rissland OS. 2018. Determining mRNA half-lives on a transcriptome-wide scale. *Methods* 137:90–98. <https://doi.org/10.1016/j.ymeth.2017.12.006>.
 56. Back S, Gorman AW, Vogel C, Silva GM. 2019. Site-specific K63 ubiquitination provides insights into translation regulation under stress. *J Proteome Res* 18:309–318. <https://doi.org/10.1021/acs.jproteome.8b00623>.
 57. Jansson BPM, Malandrin L, Johansson HE. 2000. Cell cycle arrest in archaea by the hypusination inhibitor N1-guanyl-1,7-diaminoheptane. *J Bacteriol* 182:1158–1161. <https://doi.org/10.1128/jb.182.4.1158-1161.2000>.
 58. Byrne R, Levin JG, Bladen HA, Nirenberg MW. 1964. The in vitro formation of a DNA-ribosome complex. *Proc Natl Acad Sci U S A* 52:140–148. <https://doi.org/10.1073/pnas.52.1.140>.
 59. Miller OL, Hamkalo BA, Thomas CA. 1970. Visualization of bacterial genes in action. *Science* 169:392–395. <https://doi.org/10.1126/science.169.3943.392>.
 60. French SL, Santangelo TJ, Beyer AL, Reeve JN. 2007. Transcription and translation are coupled in Archaea. *Mol Biol Evol* 24:893–895. <https://doi.org/10.1093/molbev/msm007>.
 61. Gonzalez C, Sims JS, Hornstein N, Mela A, Garcia F, Lei L, Gass DA, Amendolara B, Bruce JN, Canoll P, Sims PA. 2014. Ribosome profiling reveals a cell-type-specific translational landscape in brain tumors. *J Neurosci* 34:10924–10936. <https://doi.org/10.1523/JNEUROSCI.0084-14.2014>.
 62. Kirschner M, Gerhart J. 1998. Evolvability. *Proc Natl Acad Sci U S A* 95:8420–8427. <https://doi.org/10.1073/pnas.95.15.8420>.
 63. Schlosser G, Wagner GP. 2004. Modularity in development and evolution. University of Chicago Press, Chicago, IL.
 64. Dinman JD. 2016. Pathways to specialized ribosomes: the Brussels Lecture. *J Mol Biol* 428:2186–2194. <https://doi.org/10.1016/j.jmb.2015.12.021>.
 65. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
 66. Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* 34:525–527. <https://doi.org/10.1038/nbt.3519>.
 67. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15:550. <https://doi.org/10.1186/s13059-014-0550-8>.
 68. Anders S, Pyl PT, Huber W. 2015. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31:166–169. <https://doi.org/10.1093/bioinformatics/btu638>.
 69. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29:15–21. <https://doi.org/10.1093/bioinformatics/bts635>.
 70. Alm EJ, Huang KH, Price MN, Koche RP, Keller K, Dubchak IL, Arkin AP. 2005. The MicrobesOnline Web site for comparative genomics. *Genome Res* 15:1015–1022. <https://doi.org/10.1101/gr.3844805>.
 71. Alexa A, Rahnenfuhrer J, Lengauer T. 2006. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics* 22:1600–1607. <https://doi.org/10.1093/bioinformatics/btl140>.
 72. Supek F, Bošnjak M, Škunca N, Šmuc T. 2011. REVIGO summarizes and visualizes long lists of Gene Ontology terms. *PLoS One* 6:e21800. <https://doi.org/10.1371/journal.pone.0021800>.
 73. Dehal PS, Joachimiak MP, Price MN, Bates JT, Baumohl JK, Chivian D, Friedland GD, Huang KH, Keller K, Novichkov PS, Dubchak IL, Alm EJ, Arkin AP. 2009. MicrobesOnline: an integrated portal for comparative and functional genomics. *Nucleic Acids Res* 38:D396–D400. <https://doi.org/10.1093/nar/gkp919>.
 74. Newman MEJ, Girvan M. 2004. Finding and evaluating community structure in networks. *Phys Rev E Stat Nonlin Soft Matter Phys* 69:026113. <https://doi.org/10.1103/PhysRevE.69.026113>.
 75. Morris JH, Apeltsin L, Newman AM, Baumbach J, Wittkop T, Su G, Bader GD, Ferrin TE. 2011. clusterMaker: a multi-algorithm clustering plugin for Cytoscape. *BMC Bioinformatics* 12:436. <https://doi.org/10.1186/1471-2105-12-436>.
 76. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498–2504. <https://doi.org/10.1101/gr.1239303>.
 77. Becht E, McInnes L, Healy J, Dutertre C-A, Kwok IWH, Ng LG, Ginhoux F, Newell EW. 2019. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol* 37:38–44. <https://doi.org/10.1038/nbt.4314>.
 78. Bailey TL, Gribskov M. 1998. Combining evidence using p-values: application to sequence homology searches. *Bioinformatics* 14:48–54. <https://doi.org/10.1093/bioinformatics/14.1.48>.
 79. Bailey TL, Elkan C. 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc Int Conf Intell Syst Mol Biol* 2:28–36.