nature structural & molecular biology

Rps26 directs mRNA-specific translation by recognition of Kozak sequence elements

Max B Ferretti[1,2], Homa Ghalei[1], Ethan A Ward[1,4], Elizabeth L Potts[1,3,4] & Katrin Karbstein[1,2]

We describe a novel approach to separate two ribosome populations from the same cells and use this method in combination with RNA-seq to identify mRNAs bound to *Saccharomyces cerevisiae* ribosomes with and without Rps26, a protein linked to the pathogenesis of Diamond–Blackfan anemia (DBA). These analyses reveal that Rps26 contributes to mRNA-specific translation by recognition of the Kozak sequence in well-translated mRNAs and that Rps26-deficient ribosomes preferentially translate mRNA from select stress-response pathways. Surprisingly, exposure of yeast to these stresses leads to the formation of Rps26-deficient ribosomes and to the increased translation of their target mRNAs. These results describe a novel paradigm: the production of specialized ribosomes, which play physiological roles in augmenting the well-characterized transcriptional stress response with a heretofore unknown translational response, thereby creating a feed-forward loop in gene expression. Moreover, the simultaneous gain-of-function and loss-of-function phenotypes from Rps26-deficient ribosomes can explain the pathogenesis of DBA.

Translational control of gene expression is integral to the maintenance of protein homeostasis[1–4]. Recent findings from ribosomal profiling studies show that different mRNAs are recruited to ribosomes with vastly differing efficiencies[5,6]. Furthermore, translational efficiency for any given mRNA can vary under different cellular conditions[5,7–9]. Although the molecular mechanisms underlying these differences remain poorly understood, one element that is known to affect the translation efficiency of an mRNA is the sequence immediately upstream of the start codon. Originally described by Kozak, nucleotide changes in this sequence affect protein production by an order of magnitude[10–12]. Recent studies have confirmed the importance of a ten-nucleotide window upstream of the start codon[13–15], and structural studies have identified a contact between eIF2α and the highly conserved nucleotide at the −3 position (with AUG in positions 1–3)[16]. How the other upstream bases are recognized is not well understood, partly because the mRNA in these studies did not conform to the Kozak consensus sequence, and partly because a 21–amino acid C-terminal extension of Rps26, which is located nearby, was not resolved in the structure.

Haploinsufficiency of ribosomal proteins underlies a number of diseases, such as DBA, congenital asplenia, and 5q– syndrome[17–20]. These diseases exhibit a range of tissue-specific symptoms but share a seemingly paradoxical phenotype: proliferative and growth deficiencies (often coupled with developmental defects) paired with a greatly increased risk of cancer[20–22]. Whereas the defects in rapidly proliferating tissues are perhaps expected from insufficiency in ribosomal proteins, a lack of ribosomes would predict a resistance and not a susceptibility to cancer. Rps26 is the third most commonly mutated protein in DBA[23], and its location in the mRNA exit channel[16], as well as crosslinking data[24], predicts contacts with the mRNA upstream of the start codon during translation initiation.

Here we report the use of a purification method that enables the separation of Rps26-depleted (ΔRps26) and Rps26-containing (+Rps26) ribosomes from yeast cells, followed by RNA-seq to identify the mRNAs bound to both ribosome pools. The data show that each type of ribosome bound a specific subset of mRNAs, and that mRNAs bound to ΔRps26 ribosomes were translated less efficiently in wild-type (WT) cells than in those enriched in +Rps26 ribosomes. Furthermore, mRNAs enriched in ΔRps26 ribosomes lacked the sequence conservation upstream of the Kozak sequence. Luciferase reporter assays confirmed that Rps26 was required for preferential translation of mRNAs with an adenosine at position −4 and furthermore demonstrated that mRNAs with a guanosine at position −4 (−4G) were preferentially translated by ΔRps26 ribosomes. Pathway analysis showed clustering of mRNAs bound to ΔRps26 ribosomes in the Hog1 and Rim101 pathways, which regulate the response to high-salt and high-pH stress conditions, respectively. Correspondingly, Rps26 depletion led to constitutive activation of these pathways and therefore to increased resistance to high salt and high pH. Finally, we show that upon exposure to high salt or high pH, cells generated ΔRps26 ribosomes, thereby allowing for preferential translation of mRNAs with mutations in the −4 position of the Kozak sequence. These data reveal the molecular basis for recognition of the Kozak sequence and suggest that perturbed protein homeostasis could play a role in the pathogenesis of DBA. Moreover, they demonstrate that the well-characterized transcriptional response to stress induced by high-salt and high-pH conditions is augmented

[1]Department of Integrative Structural and Computational Biology, The Scripps Research Institute, Jupiter, Florida, USA. [2]The Doctoral Program in Chemical and Biological Sciences, The Scripps Research Institute, Jupiter, Florida, USA. [3]Harriet L. Wilkes Honors College, Florida Atlantic University, Boca Raton, Florida, USA. [4]Present addresses: Henry Samueli School of Engineering and Applied Science, University of California, Los Angeles, California, USA (E.A.W.); The University of Florida, Gainesville, Florida, USA (E.L.P.). Correspondence should be addressed to K.K. (kkarbst@scripps.edu).

Received 14 October 2016; accepted 27 June 2017; published online 31 July 2017; doi:10.1038/nsmb.3442

by a hitherto unknown translational response that involves changes in the composition of ribosomes to enable preferential translation of a subset of stress-related mRNAs. Thus, Rps26-deficient ribosomes have physiological roles.

RESULTS

ΔRps26 and +Rps26 ribosomes bind different mRNAs

Structural and crosslinking data demonstrate that Rps26 is located in the mRNA exit channel, adjacent to the Kozak sequence[16,24]. To assess whether Rps26 influences the repertoire of mRNAs translated by the ribosome, we purified +Rps26 and ΔRps26 ribosomes from a yeast strain in which ΔRps26 ribosomes are selectively TAP tagged (**Fig. 1a**). In this strain, Rps26 and Rps3, a distal protein, are under galactose-inducible control. Furthermore, Rps3-TAP is under the control of the doxycycline (dox)-repressible TET promoter, allowing for separate induction of Rps3-TAP when Rps26 and Rps3 are repressed. Importantly, Rps3-TAP fully complements the absence of Rps3 (ref. 25), and ribosomes containing Rps3-TAP are recruited into polysomes akin to those with untagged Rps3 (**Supplementary Fig. 1a–d**). After incubation for 4 h in glucose, cells were harvested and lysed in the presence of cycloheximide to maintain mRNA–ribosome interactions, and then ΔRps26 ribosomes were separated from +Rps26 ribosomes by affinity purification on IgG beads (**Fig. 1b**). To exclude free mRNAs, we loaded the flowthrough fraction onto a sucrose gradient and collected only the ribosome-bound fractions. SDS-PAGE analysis and immunoblotting of the eluate (containing ΔRps26 ribosomes) and the flowthrough (containing +Rps26 ribosomes) showed that this method efficiently separated the two ribosome populations (**Fig. 1b**). Furthermore, northern blotting and Bioanalyzer results demonstrated that the ΔRps26 ribosomes contained mature 18S rRNA (**Fig. 1c** and **Supplementary Fig. 1f,g**). RNAs were isolated from these ribosome pools, sequenced on an Illumina NexSeq 500, and analyzed with the DESeq2 algorithm. Notably, because the two ribosome pools were isolated from the same cells, and therefore encountered the same mRNAs, the RNA-seq data from both pools could be compared directly.

RNA-seq analysis showed that 88–95% of the reads from the +Rps26 sample and 89–95% of the reads from the ΔRps26 sample mapped to the yeast genome. Further, 90% of the open reading frames (ORFs: 5,219 out of a total of 5,784) had 128 or more reads per experiment, a threshold experimentally determined to reduce the false appearance of gene regulation[5]. Of the remaining 565 ORFs, 274 (49%) were annotated as 'uncharacterized' or 'dubious' and may not encode proteins (**Supplementary Fig. 2c**). Furthermore, data from three independently grown and purified samples showed high correlation between experiments (**Supplementary Fig. 2d**). Thus, RNA-seq enabled the identification of the overwhelming majority of ORFs with high confidence and reproducibility.

Using a conservative cutoff of $P_{adj} < 0.05$, we determined that more than one-quarter of all mRNAs were significantly enriched in either of the ribosome pools. Of those, 13% were enriched in the ΔRps26 pool and 15% were preferentially bound to +Rps26 ribosomes (**Fig. 1d** and **Supplementary Data Set 1**), and all mRNAs were enriched more than 1.5-fold. Importantly, control experiments in which we sequenced Rps3-TAP-bound mRNAs demonstrated that the differences between these samples did not arise from the TAP tag on Rps26-deficient ribosomes (**Supplementary Fig. 1h**).

To validate these data, we determined the effects of Rps26 depletion on polysome recruitment of mRNAs enriched in the two ribosome pools. We depleted Rps26 in a dox-repressible Rps26 strain that accumulates ΔRps26 ribosomes in the presence of dox but maintains a level of the protein similar to that in the parent strain in its absence (**Fig. 2a** and **Supplementary Fig. 3a–d**). We hypothesized that if ribosome pulldowns reflect functional interactions with translating ribosomes, Rps26 depletion should selectively decrease the polysome recruitment

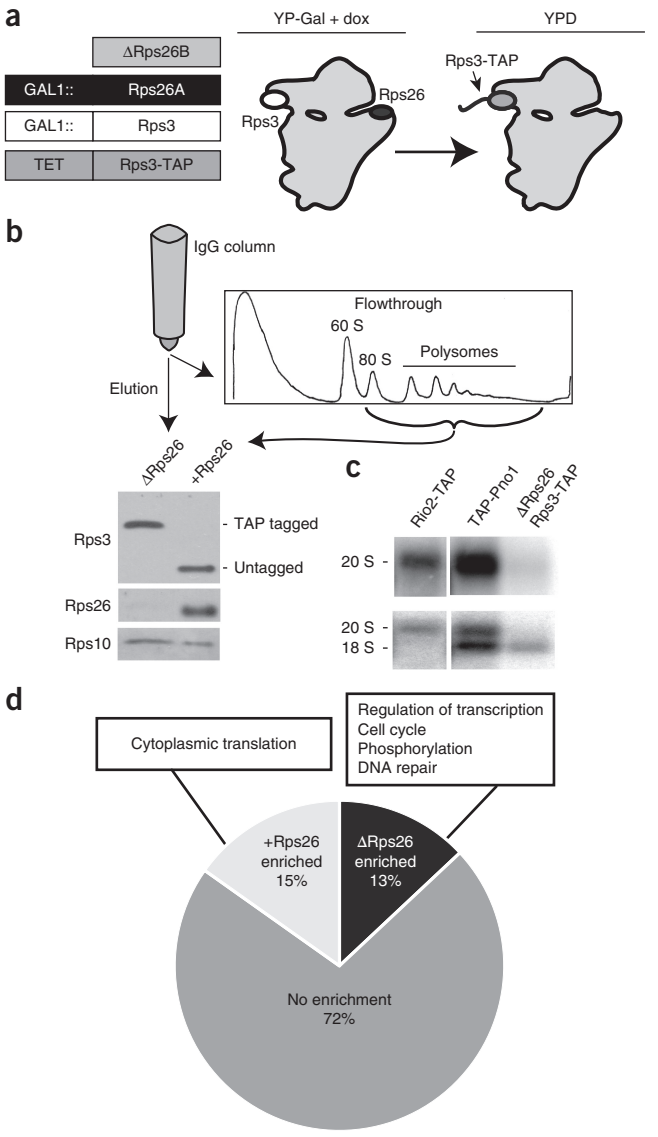

Figure 1 Isolation of ΔRps26 ribosomes and characterization of bound mRNAs. (**a**) The strain used to purify ribosomes lacking Rps26 (ΔRps26). (**b**) Separation of ΔRps26 and +Rps26 ribosomes by affinity purification with Rps3-TAP. The flowthrough was loaded on a sucrose gradient, and only the fractions containing mRNA–ribosome complexes were used for sequencing analysis. A western blot of the elution and flowthrough for (tagged and untagged) Rps3 and Rps26 is shown. Rps10 was used as a loading control. Data are representative of at least five experiments. (**c**) Northern blot analysis of ΔRps26 ribosomes compared to immature ribosomes captured by TAP-tagging of the assembly factors Rio2 and Pno1. Samples were probed against ITS1 (20S rRNA) (top) and 18S rRNA (bottom). Data are representative of three experiments. (**d**) The distribution of mRNAs enriched ($P_{adj} < 0.05$ by DESeq2) in +Rps26 and ΔRps26 ribosomes. mRNAs in each pool: +Rps26, 865; ΔRps26, 741; no enrichment, 4,084. The GO terms with the most genes represented in each pool are also listed (**Supplementary Fig. 1**). Original blot images are in **Supplementary Data Set 3**. Source data for **d** are available online.

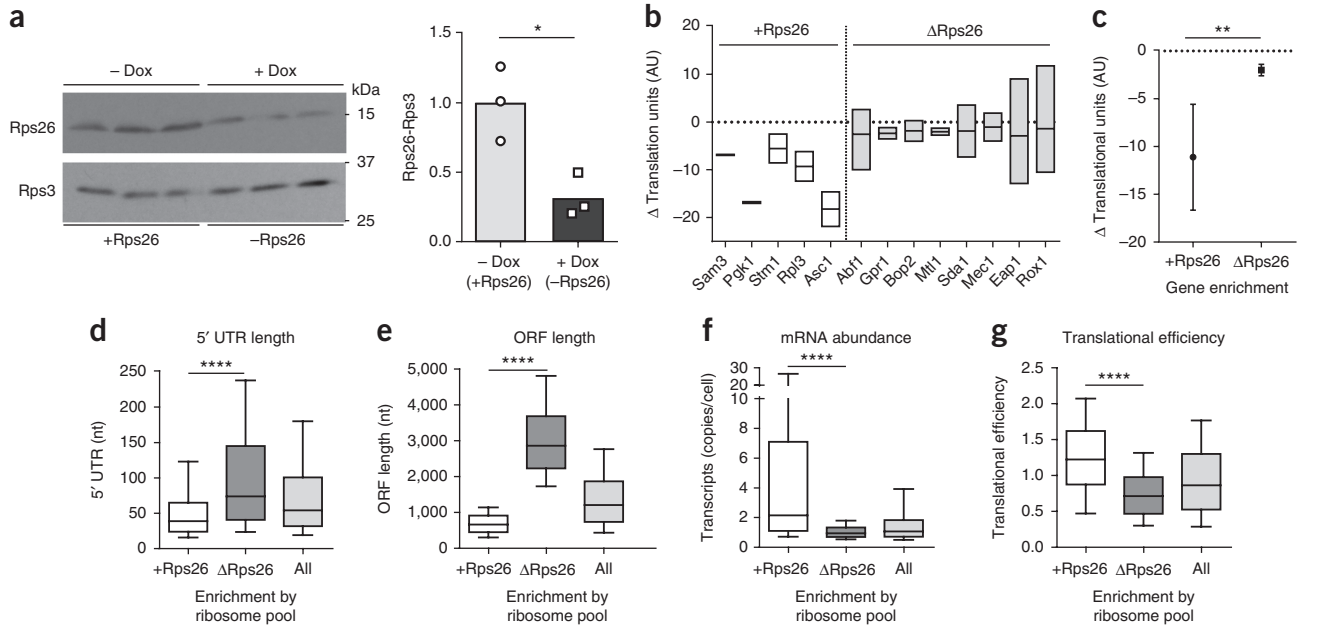

Figure 2 ΔRps26 ribosomes bind a distinct set of poorly translated mRNAs. (**a**) Ribosomes isolated from TET-Rps26 cells exposed to dox were probed for Rps26 occupancy relative to Rps3. Shown are a western blot (left) and quantitation (right) of ribosomes purified in parallel from three separate cultures. Bars represent the mean. $*P < 0.05$, $t = 3.795$, Degrees of freedom (DF) = 4 (Student's t test). Data are representative of four experiments. (**b**) Translational downregulation (measured by bound ribosomes) of mRNAs enriched in +Rps26 or ΔRps26 ribosomes after Rps26 depletion. Three independent experiments were run for Abf1, Eap1 and Rox1, and duplicates were run for all others. Box plots represent the range; the midline indicates the mean. AU, arbitrary units. (**c**) Analysis of all mRNAs in **b**; data are mean ± s.d. $**P = 0.0016$, Kolmogorov–Smirnov test. (**d**–**g**) Metagene analysis using existing data sets[58–61]. $****P < 0.0001$, Kolmogorov–Smirnov test. Number of mRNAs analyzed for 5′ UTR length: +Rps26, 763; ΔRps26, 575; all, 4,363. ORF length: +Rps26, 763; ΔRps26, 575; all, 5,690. Abundance: +Rps26, 804; ΔRps26, 739; all, 5,459. Translational efficiency: +Rps26, 838; ΔRps26, 736; all = 5476. Box edges represent the interquartile range, midlines indicate the median, and lower and upper whiskers extend to the 10th and 90th percentiles, respectively (**Supplementary Figs. 2** and **3**). Original blot images are in **Supplementary Data Set 3**. Source data for graphs in this figure are available online.

of mRNAs enriched in the +Rps26 pool. As previously described[26], we coupled sucrose-gradient fractionation with RT-qPCR to track the distribution and number of ribosomes bound to mRNAs enriched in the +Rps26 versus ΔRps26 pools when Rps26 was replete or depleted (**Supplementary Fig. 3e**–**h**). Quantification of these analyses showed that Rps26 depletion had a smaller effect on polysome recruitment of mRNAs preferentially bound by ΔRps26 ribosomes compared to that of mRNAs enriched on +Rps26 ribosomes (**Fig. 2b**,**c**), thus validating the sequencing data *in vivo*.

Further analysis revealed that the mRNAs enriched on +Rps26 and ΔRps26 ribosomes differed in their physical characteristics. Specifically, mRNAs enriched on ΔRps26 ribosomes had longer 5′ UTRs (untranslated regions) and longer ORFs, and were less abundant than +Rps26-enriched mRNAs or the transcriptome as a whole (**Fig. 2d**–**f**). This finding held true even when mRNAs encoding ribosomal proteins, which tend to be short, were excluded (**Supplementary Fig. 2e**). In contrast, there were no significant differences in the average lengths of the 3′ UTRs or transcript half-lives (**Supplementary Fig. 2f**). Although the molecular basis for these observations is unclear, long 5′ UTRs and ORFs are known to be associated with less efficient translation[5,27,28]. Importantly, mRNAs enriched on ΔRps26 ribosomes were much less efficiently translated than mRNAs enriched on +Rps26 ribosomes and the entire transcriptome. In contrast, +Rps26 ribosomes enriched highly translated mRNAs (**Fig. 2g**), consistent with a role for Rps26 in recognition of the Kozak sequence, which imparts highly efficient translation.

Rps26-based mRNA preference is mediated by Kozak sequence elements

We used Weblogo analysis to interrogate the sequence enrichment of mRNAs bound to ΔRps26 and +Rps26 ribosomes[29]. Crosslinking[24] and structural[16] data indicate that Rps26 is bound in the mRNA exit channel, potentially contacting residues −4, and −7 through −10 (relative to the start codon). Thus, in our analysis of sequence enrichment we focused on the sequence immediately upstream of the start codon. This analysis showed that mRNAs enriched on +Rps26 ribosomes were likely to contain an adenosine at residue −3, as discovered by Kozak[30]. In addition, adenosines were moderately enriched at residues in the −1, −2, −4, −7, −8 and −10 positions (**Fig. 3a**), broadly recapitulating the Kozak consensus sequence for yeast[13,31,32]. This finding is consistent with the notion that +Rps26 ribosomes preferentially translate mRNAs containing adenosines at these positions. In contrast, ΔRps26-enriched mRNAs lacked conservation at any position other than −3 and reflected the entire yeast transcriptome (**Fig. 3b**,**c**). Thus, the selection of mRNAs containing adenosines at positions −1, −2, −4, −7, −8 and −10 relies on Rps26.

To test whether translation of mRNAs with a Kozak consensus sequence upstream of the start codon requires Rps26, we developed a dual luciferase reporter assay. We transformed yeast with plasmids encoding firefly luciferase preceded by the adenosine sequence enriched in +Rps26 ribosomes (A_{10}) and *Renilla* luciferase preceded by various upstream sequences, each containing mutations at different positions of the Kozak sequence (**Supplementary Fig. 4a**); thus, the ratio of *Renilla* to firefly luciferase quantified

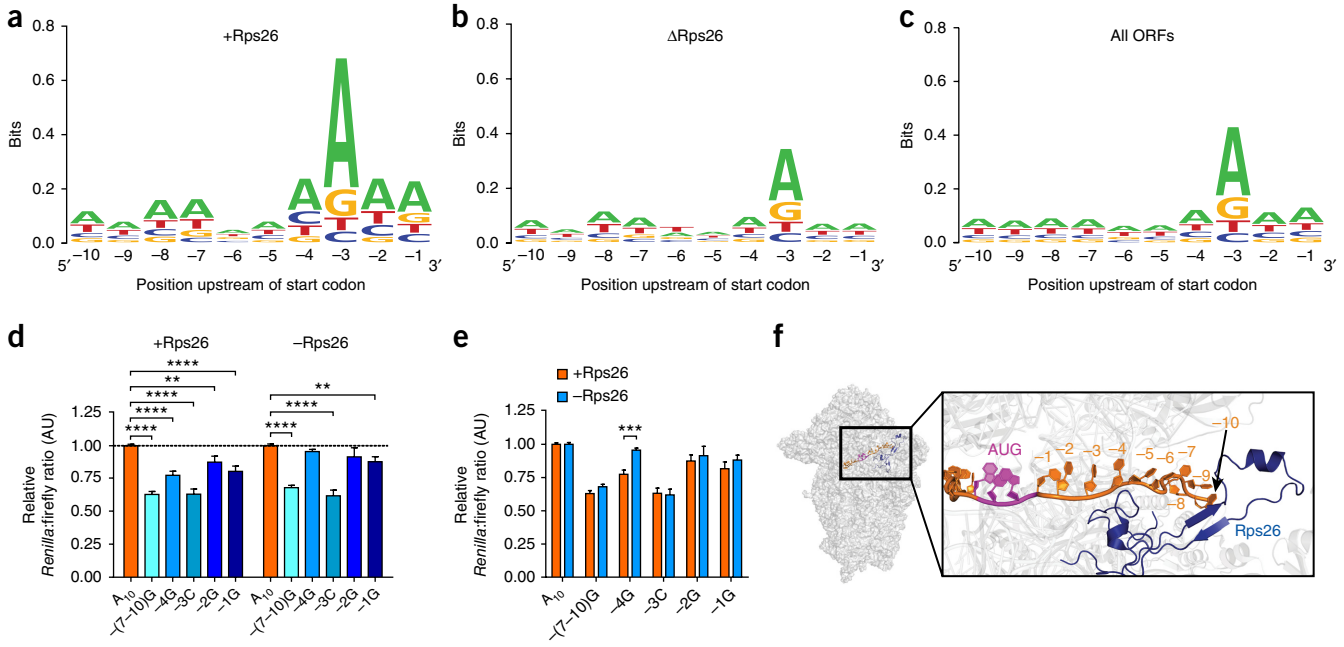

Figure 3 Rps26 promotes translation by recognizing specific residues in the Kozak sequence. (**a**–**c**) Weblogo conservation analysis of mRNAs enriched in +Rps26 ribosomes (**a**; $n = 865$) and ΔRps26 ribosomes (**b**; $n = 741$) versus all mRNAs in our data set (**c**; $n = 5{,}696$). (**d**) Results of luciferase reporter assays from TET-Rps26 cells grown in the absence or presence of dox, to test effects from Kozak-sequence mutations with replete (+) or depleted (–) Rps26. For +Rps26, from left to right starting with –(7–10)G: $t = 10.24$, 6.229, 7.594, 3.193, or 4.84 versus A_{10}; $^{****}P < 0.0001$, $^{**}P = 0.0017$. For –Rps26, from left to right starting with –(7–10)G: $^{****}P < 0.0001$, $t = 8.869$; $P = 0.2228$, $t = 1.223$; $^{****}P < 0.0001$, $t = 8.328$; $P = 0.0603$, $t = 2.179$; $^{**}P = 0.0064$, $t = 3.116$. Statistical values determined by two-way analysis of variance (ANOVA). DF = 184. (**e**) The effect of Rps26 depletion on the recognition of individual point mutations. –Rps26 versus +Rps26 from left to right, starting with –(7–10)G: $P = 0.6562$, $t = 1.193$; $^{***}P = 0.0002$, $t = 4.235$; $P = 0.9712$, $t = 0.2146$; $P = 0.7879$, $t = 0.8371$; $P = 0.5621$, $t = 1.438$ (two-way ANOVA). DF = 184. For **d** and **e**, experiments used the following numbers of independent cultures: –(7–10)G and –4G, $n = 15$; –3C, $n = 7$; –2G, $n = 12$; –1G, $n = 13$. Data shown are mean and s.e.m. (**Supplementary Fig. 4**). (**f**) Rps26 (blue) binds mRNA near the –4 position when the start codon (purple) is in the P site (adapted from PDB 3J81, ref. 16). The C-terminal 21 amino acids of Rps26 are not resolved in this structure. Source data for **d** and **e** are available online.

the effects of a given upstream sequence on translation relative to that of the A_{10} sequence.

In Rps26-replete yeast, mutation of individual adenosines at the –1, –2, –3 and –4 positions, as well as positions –7 to –10 together, reduced translation (**Fig. 3d**, left). These data demonstrate that residues upstream of the start codon impact translation, which is consistent with findings from Kozak's studies in mammalian cells[10–12,33] and more recent work in yeast[13]. In contrast, in Rps26-depleted cells, translation of the –4G and –2G mutants was indistinguishable from that of the A_{10} mRNA (**Fig. 3d**, right). Thus, Rps26 is required for preferential translation of mRNAs with adenosine at the –4 and –2 positions. Further, translation of mRNAs containing –4G was increased in Rps26-deficient versus Rps26-replete yeast (**Fig. 3e**). This result is consistent with the sequencing data, which show that relative to the transcriptome as a whole, ΔRps26 ribosomes specifically enriched mRNAs with a G at the –4 position. This preference came at the loss of mRNAs with an A or C at that position (**Supplementary Fig. 4b**). In contrast, +Rps26 ribosomes enriched for mRNAs with A or C and depleted those with G at the –4 position (**Supplementary Fig. 4b**). This effect was specific to Rps26, as depletion of Rps3 or Rps17 had no such effect on translation of the –4G reporter (**Supplementary Fig. 4d**). Because Rps26 depletion leads to moderate 20S rRNA accumulation[34], we wondered whether the accumulation of 20S rRNA and not the depletion of Rps26 was responsible for the observed defect in recognition of the –4 residue. We therefore analyzed translation of the luciferase reporter in WT yeast and in yeast lacking the 40S assembly factor Ltv1 (ΔLtv1). Similar to the Rps26-depleted strain, the ΔLtv1 strain accumulates moderate amounts of 20S rRNA[35,36]. Importantly, translation of the luciferase reporter was sensitive to the residue at the –4 position in ΔLtv1 yeast (**Supplementary Fig. 4e**), demonstrating that it is depletion of Rps26 and not accumulation of 20S rRNA that is responsible for the observed defect in recognition of the –4 residue.

Importantly, Rps26 directly contacts the –4 position, as well as an rRNA residue directly neighboring an interactor of the –2 residue (**Fig. 3f** and ref. 16). Thus, the sequencing data, reporter assays and structural studies all demonstrate that Rps26 recognizes the –4 and –2 positions of the Kozak sequence, and that this is necessary for preferential translation of mRNAs with a Kozak consensus of A or C at that position. ΔRps26 ribosomes have lost this preference and instead show a slight preference for mRNAs containing –4G.

Accumulation of ΔRps26 ribosomes activates the Hog1 and Rim101 pathways

To assess whether mRNAs enriched in ΔRps26 ribosomes cluster in specific biological pathways in which they could produce specific biological outcomes, we used mRNAs enriched in each ribosome pool as a basis for GO-term enrichment analysis on GeneCodis[37] (**Fig. 1c** and **Supplementary Data Set 2**). We found that +Rps26-ribosome-enriched mRNAs encoded mainly ribosomal proteins and translation factors. This is consistent with previous findings showing that mRNAs encoding ribosomal proteins are among the best translated[5,26]. In contrast, ΔRps26-ribosome-enriched transcripts were associated with highly regulated processes, including transcriptional control, phosphorylation, cell cycle, DNA repair and, notably, well-characterized

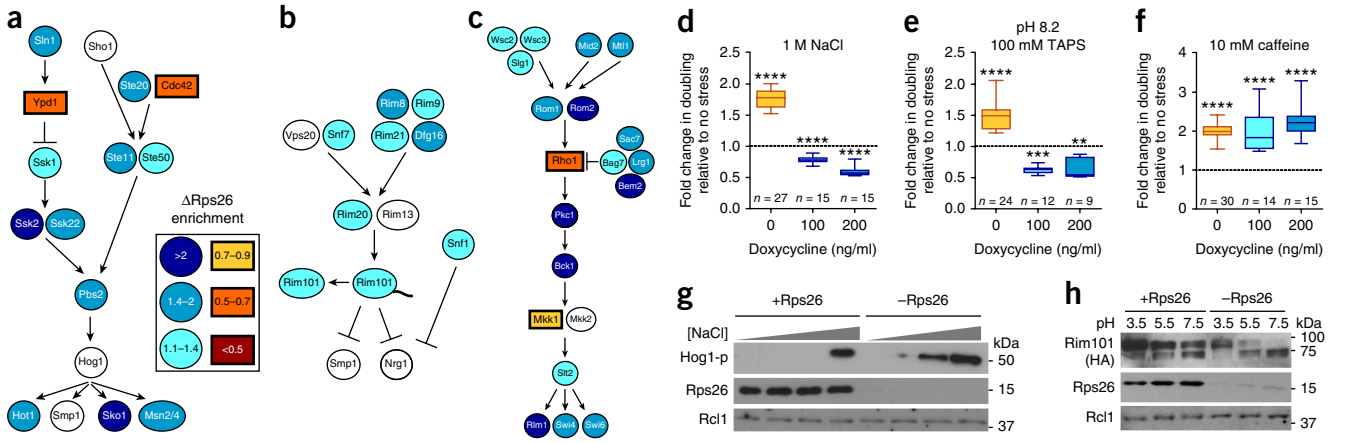

Figure 4 Accumulation of ΔRps26 ribosomes activates specific biological pathways. (**a**–**c**) Enrichment of mRNAs encoding components of the Hog MAPK (**a**), Rim101 (**b**), and cell-wall integrity (**c**) pathways in ΔRps26 or +Rps26 ribosomes. mRNAs enriched in ΔRps26 ribosomes are shown in ovals, mRNAs enriched in +Rps26 ribosomes are in rectangles, colored according to the key; mRNAs not enriched in either type of ribosome are shown in white. (**d**–**f**) Changes in doubling time in cells containing Rps26 (–dox) or deficient in Rps26 (+dox) after exposure to salt (**d**), high pH (**e**) or caffeine (**f**). Box edges indicate the interquartile range, midlines indicate the median, and whiskers extend to the minimum and maximum values. Values were compared to no-stress conditions (fold change = 1). $^{****}P < 0.0001$, $^{***}P = 0.0005$, $^{**}P = 0.0039$ (Wilcoxon signed rank test). Sum of signed ranks (W): **d**, 378, –120, 120; **e**, 300, –78, –45; and **f**, 465, 105, 120, for cells exposed to 0, 100, and 200 ng/ml of dox, respectively. Data are from *n* independently grown cultures, as indicated. (**g**) Hog1 phosphorylation (Hog1-p) in Rps26-containing and Rps26-deficient cells after exposure to increasing concentrations of NaCl (0–500 mM). (**h**) Proteolytic activation of Rim101 at different pH values in Rps26-containing and Rps26-deficient cells. HA, hemagglutinin. Original blot images are in **Supplementary Data Set 3**. Western blots are representative of four independent experiments, and representative data are shown. Source data for **d**–**f** are available online.

stress-responsive signaling pathways including the Hog1 pathway[38], a MAPK cascade that responds to high-salt stress, and the Rim101 pathway[39], which responds to pH stress (**Fig. 4**).

Because Rps26 depletion increased the relative translation of mRNAs enriched in ΔRps26 ribosomes (**Fig. 2b**,**c**), we hypothesized that cells depleted of Rps26 are more resistant to these stresses. To test this prediction, we measured growth rates for Rps26-replete and Rps26-depleted yeast cells in rich media with and without high salt or high pH. Surprisingly, Rps26-deficient cells grew faster in high-salt and high-pH conditions than in rich media, whereas Rps26-replete cells were sensitive to both salt and pH (**Fig. 4d**,**e**). Additionally, Rps26-deficient yeast displayed increased phosphorylation of the Hog1 MAPK at lower concentrations of NaCl (**Fig. 4g**) and C-terminal cleavage of the Rim101 protein, a hallmark of the high-pH-response pathway, even in mildly acidic conditions (**Fig. 4h**). Thus, both the Hog1 MAPK and the Rim101 pathways are induced when Rps26 is depleted. This phenotype was pathway selective, as Rps26-deficient yeast were not resistant to caffeine, which activates a distinct MAPK cascade to produce cell-wall stress[40] (**Fig. 4c**,**f**). Further, high-salt and high-pH conditions produced a growth defect in yeast strains depleted of two other late-binding ribosomal proteins, Rps3 and Rps17 (**Supplementary Fig. 5a**–**g**), demonstrating the specific role of Rps26 deficiency in resistance to salt and pH stress. Finally, ΔLtv1 cells also did not show stress resistance, demonstrating that stress resistance does not arise from the accumulation of 20S pre-rRNA in these cells (**Supplementary Fig. 5h**,**i**).

Yeast form ΔRps26 ribosomes in response to osmotic and high-pH stress

Given the high-salt and high-pH stress phenotypes observed in Rps26-deficient yeast, we reasoned that the formation of ΔRps26 ribosomes might be part of the cellular response to these stresses. To test this, we grew WT yeast in rich media containing high salt concentrations, high pH or caffeine. We then purified ribosomes from these stressed cells and from non-stressed control cells, and analyzed Rps26 protein levels by western blotting. We also assessed the levels of five additional proteins from the small ribosomal subunit (Rps0, Rps3, Rps5, Rps8 and Asc1) as controls. Notably, when normalized to any of these proteins, ribosomes from cells grown under high-salt or high-pH conditions contained less Rps26 than ribosomes from unstressed cells, whereas the levels of Rps26 in cells exposed to caffeine were similar to those in non-stressed control cells (**Fig. 5a**–**c**). These data indicate that the formation of Rps26-deficient ribosomes is part of the response to high-salt and high-pH conditions.

We next validated the *in vivo* translational effects of this physiological depletion of Rps26 using the dual luciferase assay. We have shown that translation of the –4G *Renilla* reporter is increased in Rps26-deficient yeast, whereas it is indistinguishable from that of the A_{10} *Renilla* reporter (**Fig. 3d**, right). We used this functional signature of ΔRps26 ribosomes as a readout for the formation and functionality of ΔRps26 ribosomes under stress. We exposed WT cells transformed with the A_{10} or the –4G *Renilla* reporters to salt, high pH or caffeine stress. As expected from the accumulation of ΔRps26 ribosomes in the presence of high salt concentrations or high pH, translation of the –4G reporter recovered to the level of that of the A_{10} construct when WT cells were exposed to salt or high pH stress but not when exposed to caffeine (**Fig. 5d**–**g**). These data strongly suggest that the ΔRps26 ribosomes that form when yeast are exposed to high salt or high pH translate –4G-containing mRNAs. Thus, in response to high-salt and high-pH stress, cells produce ΔRps26 ribosomes, thereby augmenting the well-characterized transcriptional response to these stresses with preferential translation of mRNAs from these stress-response pathways by ΔRps26 ribosomes.

DISCUSSION

Rps26 enforces the translational program encoded by the Kozak sequence

The data herein demonstrate key roles for Rps26 in recognizing adenosines at positions –2 and –4 of the Kozak sequence, leading to

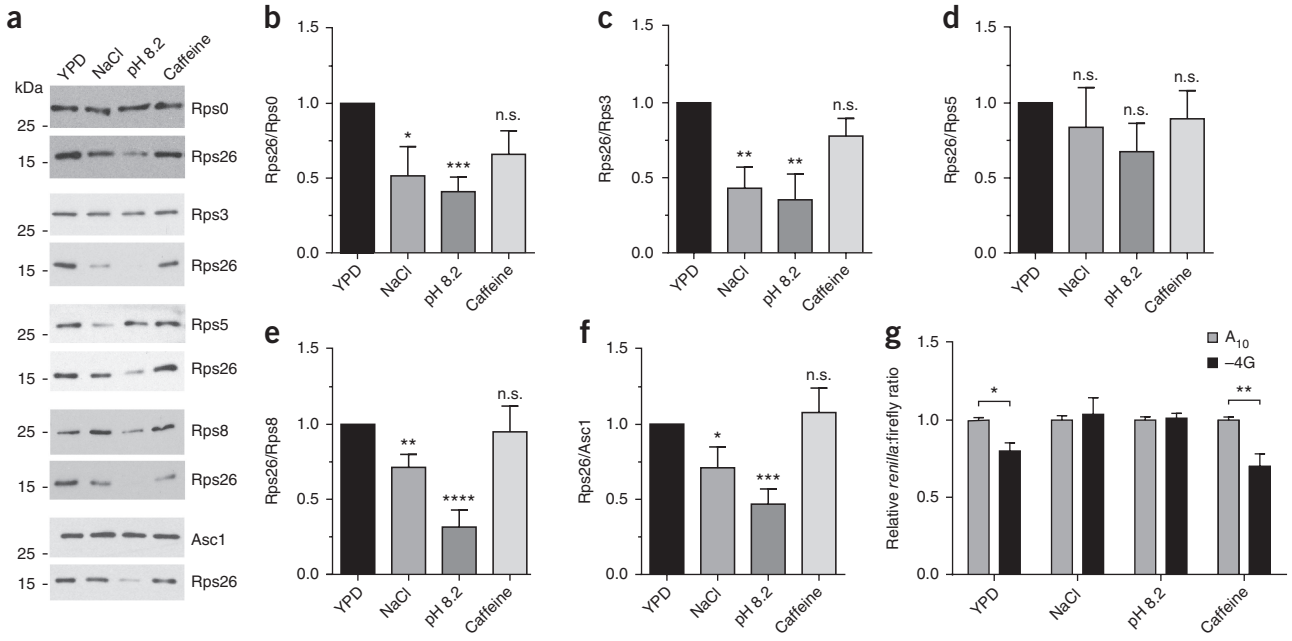

Figure 5 Cells generate ΔRps26 ribosomes in response to high-salt and high-pH conditions. (**a**) Western blots showing the levels of Rps0, Rps3, Rps5, Rps8, Rps26 and Asc1 in ribosomes purified from WT yeast grown in YPD alone, with 1 M NaCl, pH 8.2, or with 10 mM caffeine. Because all ribosomal proteins except Rps26 are nearly identical in size, separate gels were run to analyze each of these proteins relative to Rps26, and the Rps26 and RpsX bands from each of these gels are shown. (**b**–**f**) Quantification of data in **a**. Data are from ribosomes grown and purified in three independent replicates (four for caffeine) for each condition and each run on two different sets of gels. All comparisons are to the YPD-only condition. (**b**) NaCl, $*P = 0.0388$, $t = 2.469$, DF = 8; pH, $***P = 0.0003$, $t = 5.967$, DF = 8; caffeine, $P = 0.1018$, $t = 1.802$, DF = 10. (**c**) NaCl, $**P = 0.0025$, $t = 4.015$, DF= 10; pH, $**P = 0.0036$, $t = 3.773$, DF = 10; caffeine, $P = 0.1343$, $t = 1.606$, DF = 12. (**d**) NaCl, $P = 0.5534$, $t = 0.6132$, DF = 10; pH, $P = 0.1184$, $t = 1.708$, DF = 10; caffeine, $P = 0.5866$, $t = 0.5618$, DF = 10. (**e**) NaCl, $**P = 0.008$, $t = 3.299$, DF = 10; pH, $****P < 0.0001$, $t = 7.599$, DF = 8; caffeine, $P = 0.8089$, $t = 0.2472$, DF = 12. (**f**) NaCl, $*P = 0.048$, $t = 2.287$, DF = 9; pH, $***P = 0.0004$, $t = 5.132$, DF = 10; caffeine, $P = 0.6934$, $t = 0.4039$, DF = 12. Statistical values were determined by Student's *t* test. n.s., not significant. Data shown are mean and s.e.m. (**g**) Results of luciferase assays of WT cells after a 4-h exposure to different stress conditions. $n = 7$ (YPD) or 6 (all other conditions) independently grown cultures. YPD: $*P = 0.0219$, $t = 2.817$. NaCl: $P = 0.8563$ (n.s.), $t = 0.4983$. pH: $P = 0.8673$ (n.s.), $t = 0.1681$. Caffeine: $**P = 0.0014$, $t = 3.884$. DF = 42. Statistical values were determined by two-way ANOVA. Data shown are mean and s.e.m. Original blot images are in **Supplementary Data Set 3**. Source data for **d**–**g** are available online.

preferential translation of mRNAs that contain these adenosines. Thus, these data provide a molecular basis for the effects of mutations in the Kozak sequence uncovered 30 years ago[10–12,33] (**Fig. 6a**). Because of the critical role of Rps26 in recognition of the Kozak sequence, which enables efficient translation, Rps26 depletion decreases the translation of normally highly translated mRNAs, including those encoding ribosomal proteins, explaining the overall reduced ribosome numbers in Rps26-depleted cells (**Supplementary Fig. 3e**,**f**). Notably, ΔRps26 ribosomes were also found to selectively augment the translation of mRNAs containing −4G. This might arise simply from increased competitiveness of such mRNAs, as strong Kozak mRNAs are no longer favored. Additionally, or alternatively, Rps26 deficiency might ameliorate a steric hindrance that arises from the −4G in WT ribosomes, or −4G-containing mRNAs might adopt a different structure. Regardless, Rps26 enforces the translational program encoded by the Kozak sequence, whereas Rps26-deficient ribosomes decode a separate translational program.

Remarkably, we observed that mRNAs upregulated by Rps26 depletion are not randomly distributed throughout the yeast transcriptome but instead cluster in specific pathways, including the well-characterized and highly conserved Hog1 MAPK and the high-pH-response Rim101 pathways. Growth assays and immunoblotting showed that these pathways were constitutively activated in Rps26-deficient yeast, indicating that the decreased translation of otherwise well-translated mRNAs and the increased translation of mRNAs with mutations in the Kozak sequence lead to a perturbation in cellular protein homeostasis upon Rps26 depletion. Collectively, the data support a model whereby ΔRps26 ribosomes, such as those found in DBA, play pathogenic roles by disabling the translation of essential mRNAs, including those encoding ribosomal proteins, leading to reduced ribosome production, and by selectively enhancing the translation of otherwise poorly translated mRNAs characterized by long 5′ UTRs and weak Kozak sequences.

Interestingly, whereas the large majority of mRNAs encoding ribosome-assembly factors either showed no enrichment or were enriched in ΔRps26 ribosomes, the mRNA encoding Fap7 was enriched in +Rps26 ribosomes (**Supplementary Fig. 3i**). Consequently, Fap7 levels were decreased in Rps26-depleted cells (**Supplementary Fig. 3j**). This finding explains why 20S rRNA is moderately accumulated in Rps26-depleted cells[34], even though Rps26-deficient ribosomes contain 18S rRNA (**Fig. 1c** and **Supplementary Fig. 1f**,**g**). Furthermore, depletion of Fap7 in Rps26-depleted cells explains why overexpression of Fap7 rescues the growth of ΔTsr2 cells[41]; because Tsr2 stabilizes Rps26 (refs. 42,43), ΔTsr2 cells are expected to be Rps26 deficient, which would lead to Fap7 depletion.

How can ribosomal protein haploinsufficiency lead to cancer?

In mammalian cells, activation of pro-growth pathways by translational upregulation of specific mRNAs might account for the predisposition to cancer observed in humans with DBA, whereas the overall

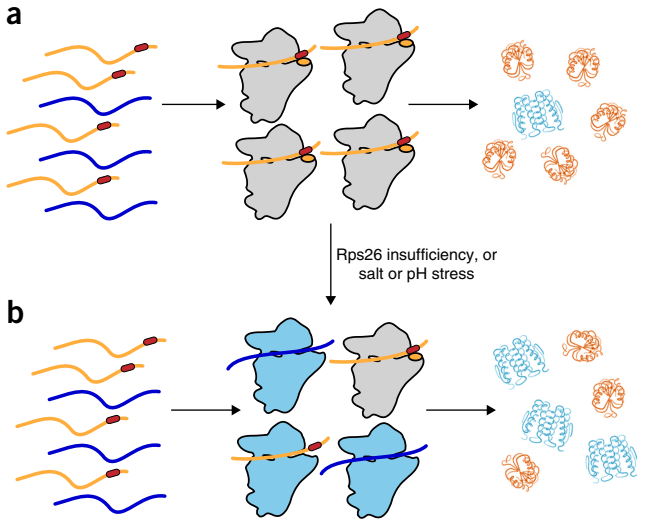

Figure 6 Disruption of protein homeostasis by Rps26 insufficiency. (**a**) Rps26 (orange ovals) recognizes positions –2 and –4 in the Kozak sequence (dark red ovals), thereby enhancing the translation of mRNAs containing canonical Kozak sequences. (**b**) When Rps26 is insufficient or in high-salt or high-pH conditions, ribosomes lacking Rps26 (light blue) accumulate, increasing the translation of mRNAs with deviations in the Kozak sequence, thereby leading to changes in protein homeostasis.

decrease in translation of ribosomal proteins (**Supplementary Data Sets 1** and **2**), and the resulting loss of ribosomes (**Supplementary Fig. 3e**,**f**), accounts for growth and developmental defects. Of note, other ribosomal proteins have also been linked to DBA[23], and several from the small subunit are also located near the mRNA exit channel. These proteins might be similarly involved in recognition of specific mRNA features or binding of translation-initiation factors, which can have mRNA-specific effects[44–46], with net effects that lead to alteration of the translational repertoire. Importantly, the method described herein for purification of ribosome subpopulations will allow assessment of mRNA-specific effects from depletion of other ribosomal proteins, or rRNA modifications, that have also been linked to cancer[47,48].

Physiological roles for ribosomes lacking ribosomal proteins

In their groundbreaking work, Barna and colleagues[49,50] demonstrated that deficiency of the large ribosomal subunit protein Rpl38 leads to specific developmental pathologies. Similar to Rps26, Rpl38 is required for translation of a specific subset of mRNAs from the Hox family. In contrast, Rps26 is required for ribosome recruitment of highly translated mRNAs, including those encoding ribosomal proteins.

Nevertheless, the data herein demonstrate not only that ΔRps26 ribosomes are deficient in translation of mRNAs containing a strong Kozak sequence, but also that ΔRps26 ribosomes selectively increase the translation of mRNAs with mutations at the –4 position in the Kozak sequence, including those encoding proteins from the Hog1 and Rim101 stress-response pathways. Most important, the data show that ΔRps26 ribosomes are produced by WT yeast upon exposure to high-salt and high-pH stress (**Fig. 6b**). Thus, ΔRps26 ribosomes also play physiological roles during stress, creating a feed-forward loop that coordinates translational and transcriptional programs that allow for cell survival in the face of changes in the extracellular milieu. Importantly, the physiological relevance of Rps26-deficient ribosomes may explain how these ribosomes can escape ribosomal quality-control mechanisms during assembly[51,52] and function of ribosomes[53,54].

How are Rps26-deficient ribosomes formed?

The observation that Rps26-deficient ribosomes are formed during high-salt and high-pH stress leads us to ask whether these ribosomes are formed from mature Rps26-containing ribosomes or whether, instead, newly made 40S subunits are produced that lack Rps26. To start addressing these questions, we used RT-qPCR to investigate whether levels of Rps26 mRNA decrease more under stress than those of other ribosomal proteins, which could allow for the production of Rps26-deficient ribosomes. The data shown in **Supplementary Figure 6a–c** demonstrate that Rps26 levels decreased similarly to those of other ribosomal proteins, perhaps less. Furthermore, forced expression of Rps26 via the galactose promoter did not have an effect on the stress sensitivity of yeast, even though the levels of Rps26 mRNA increased about 11-fold (**Supplementary Fig. 6d**,**e**). Together, these data indicate that it is not likely that downregulation of Rps26 levels leads to the *de novo* formation of ΔRps26 ribosomes, consistent with the observation that the stress conditions that lead to formation of ΔRps26 ribosomes do not support active ribosome assembly[55–57]. We thus speculate that ΔRps26 ribosomes are formed from pre-existing ribosomes. In that regard, it is interesting to note that a specific chaperone for Rps26, Tsr2, has been identified[42,43]. In addition to delivering Rps26 to nascent ribosomes, Tsr2 might store Rps26 and allow for the fully reversible loss and reincorporation of the protein.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank J. Cleveland, J. Joyce and members of the Karbstein lab for comments. This work was supported by the US National Institutes of Health (grants R01-GM086451 (to K.K.) and F31-GM116406 (to M.B.F.)), PGA National (Women's Cancer Awareness Day fellowship to H.G.), the Richard & Helen DeVos Foundation (graduate fellowship to M.B.F.), and the Howard Hughes Medical Institute (Faculty Scholar grant 55108536 to K.K.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the US National Institutes of Health.

AUTHOR CONTRIBUTIONS

Experiments were designed by M.B.F., H.G. and K.K. M.B.F., H.G. and E.L.P. performed and analyzed the experiments. E.A.W. wrote software scripts to analyze sequencing data. The manuscript was written and edited by M.B.F., H.G. and K.K., and all authors support the conclusions.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Holcik, M. & Sonenberg, N. Translational control in stress and apoptosis. *Nat. Rev. Mol. Cell Biol.* **6**, 318–327 (2005).
2. Silvera, D., Formenti, S.C. & Schneider, R.J. Translational control in cancer. *Nat. Rev. Cancer* **10**, 254–266 (2010).
3. Van Der Kelen, K., Beyaert, R., Inzé, D. & De Veylder, L. Translational control of eukaryotic gene expression. *Crit. Rev. Biochem. Mol. Biol.* **44**, 143–168 (2009).
4. Kong, J. & Lasko, P. Translational control in cellular and developmental processes. *Nat. Rev. Genet.* **13**, 383–394 (2012).
5. Ingolia, N.T., Ghaemmaghami, S., Newman, J.R.S. & Weissman, J.S. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* **324**, 218–223 (2009).
6. Ingolia, N.T., Lareau, L.F. & Weissman, J.S. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* **147**, 789–802 (2011).

7. Brar, G.A. *et al.* High-resolution view of the yeast meiotic program revealed by ribosome profiling. *Science* **335**, 552–557 (2012).
8. Zid, B.M. & O'Shea, E.K. Promoter sequences direct cytoplasmic localization and translation of mRNAs during starvation in yeast. *Nature* **514**, 117–121 (2014).
9. Stumpf, C.R., Moreno, M.V., Olshen, A.B., Taylor, B.S. & Ruggero, D. The translational landscape of the mammalian cell cycle. *Mol. Cell* **52**, 574–582 (2013).
10. Kozak, M. Point mutations close to the AUG initiator codon affect the efficiency of translation of rat preproinsulin *in vivo*. *Nature* **308**, 241–246 (1984).
11. Kozak, M. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44**, 283–292 (1986).
12. Kozak, M. At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. *J. Mol. Biol.* **196**, 947–950 (1987).
13. Dvir, S. *et al.* Deciphering the rules by which 5′-UTR sequences affect protein expression in yeast. *Proc. Natl. Acad. Sci. USA* **110**, E2792–E2801 (2013).
14. Kim, Y. *et al.* The immediate upstream region of the 5′-UTR from the AUG start codon has a pronounced effect on the translational efficiency in *Arabidopsisthaliana*. *Nucleic Acids Res.* **42**, 485–498 (2014).
15. Chen, S.-J., Lin, G., Chang, K.-J., Yeh, L.-S. & Wang, C.-C. Translational efficiency of a non-AUG initiation codon is significantly affected by its sequence context in yeast. *J. Biol. Chem.* **283**, 3173–3180 (2008).
16. Hussain, T. *et al.* Structural changes enable start codon recognition by the eukaryotic translation initiation complex. *Cell* **159**, 597–607 (2014).
17. Farrar, J.E. *et al.* Ribosomal protein gene deletions in Diamond-Blackfan anemia. *Blood* **118**, 6943–6951 (2011).
18. Bolze, A. *et al.* Ribosomal protein SA haploinsufficiency in humans with isolated congenital asplenia. *Science* **340**, 976–978 (2013).
19. Ebert, B.L. *et al.* Identification of RPS14 as a 5q– syndrome gene by RNA interference screen. *Nature* **451**, 335–339 (2008).
20. Burwick, N., Shimamura, A. & Liu, J.M. Non-Diamond Blackfan anemia disorders of ribosome function: Shwachman Diamond syndrome and 5q– syndrome. *Semin. Hematol.* **48**, 136–143 (2011).
21. Stumpf, C.R. & Ruggero, D. The cancerous translation apparatus. *Curr. Opin. Genet. Dev.* **21**, 474–483 (2011).
22. Armistead, J. & Triggs-Raine, B. Diverse diseases from a ubiquitous process: the ribosomopathy paradox. *FEBS Lett.* **588**, 1491–1500 (2014).
23. Boria, I. *et al.* The ribosomal basis of Diamond-Blackfan Anemia: mutation and database update. *Hum. Mutat.* **31**, 1269–1279 (2010).
24. Pisarev, A.V., Kolupaeva, V.G., Yusupov, M.M., Hellen, C.U.T. & Pestova, T.V. Ribosomal position and contacts of mRNA in eukaryotic translation initiation complexes. *EMBO J.* **27**, 1609–1621 (2008).
25. Koch, B. *et al.* Yar1 protects the ribosomal protein Rps3 from aggregation. *J. Biol. Chem.* **287**, 21806–21815 (2012).
26. Arava, Y. *et al.* Genome-wide analysis of mRNA translation profiles in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **100**, 3889–3894 (2003).
27. Hurowitz, E.H. & Brown, P.O. Genome-wide analysis of mRNA lengths in *Saccharomyces cerevisiae*. *Genome Biol.* **5**, R2 (2003).
28. Liu, H. *et al.* Characterization and evolution of 5′ and 3′ untranslated regions in eukaryotes. *Gene* **507**, 106–111 (2012).
29. Crooks, G.E., Hon, G., Chandonia, J.M. & Brenner, S.E. WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).
30. Kozak, M. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic Acids Res.* **12**, 857–872 (1984).
31. Hamilton, R., Watanabe, C.K. & de Boer, H.A. Compilation and comparison of the sequence context around the AUG startcodons in *Saccharomyces cerevisiae* mRNAs. *Nucleic Acids Res.* **15**, 3581–3593 (1987).
32. Nakagawa, S., Niimura, Y., Gojobori, T., Tanaka, H. & Miura, K. Diversity of preferred nucleotide sequences around the translation initiation codon in eukaryote genomes. *Nucleic Acids Res.* **36**, 861–871 (2008).
33. Kozak, M. Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. *Nucleic Acids Res.* **9**, 5233–5252 (1981).
34. Ferreira-Cerca, S., Pöll, G., Gleizes, P.-E., Tschochner, H. & Milkereit, P. Roles of eukaryotic ribosomal proteins in maturation and transport of pre-18S rRNA and ribosome function. *Mol. Cell* **20**, 263–275 (2005).
35. Ghalei, H. *et al.* Hrr25/CK1δ-directed release of Ltv1 from pre-40S ribosomes is necessary for ribosome assembly and cell growth. *J. Cell Biol.* **208**, 745–759 (2015).
36. Pertschy, B. *et al.* RNA helicase Prp43 and its co-factor Pfa1 promote 20 to 18 S rRNA processing catalyzed by the endonuclease Nob1. *J. Biol. Chem.* **284**, 35079–35091 (2009).
37. Carmona-Saez, P., Chagoyen, M., Tirado, F., Carazo, J.M. & Pascual-Montano, A. GENECODIS: a web-based tool for finding significant concurrent annotations in gene lists. *Genome Biol.* **8**, R3 (2007).
38. Hohmann, S. Control of high osmolarity signalling in the yeast *Saccharomyces cerevisiae*. *FEBS Lett.* **583**, 4025–4029 (2009).
39. Hayashi, M., Fukuzawa, T., Sorimachi, H. & Maeda, T. Constitutive activation of the pH-responsive Rim101 pathway in yeast mutants defective in late steps of the MVB/ESCRT pathway. *Mol. Cell. Biol.* **25**, 9478–9490 (2005).
40. Levin, D.E. Cell wall integrity signaling in *Saccharomyces cerevisiae*. *Microbiol. Mol. Biol. Rev.* **69**, 262–291 (2005).
41. Peña, C., Schütz, S., Fischer, U., Chang, Y. & Panse, V.G. Prefabrication of a ribosomal protein subcomplex essential for eukaryotic ribosome formation. *eLife* **5**, e21755 (2016).
42. Peng, W.T. *et al.* A panoramic view of yeast noncoding RNA processing. *Cell* **113**, 919–933 (2003).
43. Schütz, S. *et al.* A RanGTP-independent mechanism allows ribosomal protein nuclear import for ribosome assembly. *eLife* **3**, e03473 (2014).
44. Lee, A.S.Y., Kranzusch, P.J. & Cate, J.H.D. eIF3 targets cell-proliferation messenger RNAs for translational activation or repression. *Nature* **522**, 111–114 (2015).
45. Truitt, M.L. *et al.* Differential requirements for eIF4E dose in normal development and cancer. *Cell* **162**, 59–71 (2015).
46. Wolfe, A.L. *et al.* RNA G-quadruplexes cause eIF4A-dependent oncogene translation in cancer. *Nature* **513**, 65–70 (2014).
47. Dong, X.-Y. *et al.* SnoRNA *U50* is a candidate tumor-suppressor gene at 6q14.3 with a mutation associated with clinically significant prostate cancer. *Hum. Mol. Genet.* **17**, 1031–1042 (2008).
48. Dong, X.-Y. *et al.* Implication of snoRNA *U50* in human breast cancer. *J. Genet. Genomics* **36**, 447–454 (2009).
49. Kondrashov, N. *et al.* Ribosome-mediated specificity in Hox mRNA translation and vertebrate tissue patterning. *Cell* **145**, 383–397 (2011).
50. Xue, S. *et al.* RNA regulons in Hox 5′ UTRs confer ribosome specificity to gene regulation. *Nature* **517**, 33–38 (2015).
51. Strunk, B.S., Novak, M.N., Young, C.L. & Karbstein, K. A translation-like cycle is a quality control checkpoint for maturing 40S ribosome subunits. *Cell* **150**, 111–121 (2012).
52. Lebaron, S. *et al.* Proofreading of pre-40S ribosome maturation by a translation initiation factor and 60S subunits. *Nat. Struct. Mol. Biol.* **19**, 744–753 (2012).
53. LaRiviere, F.J., Cole, S.E., Ferullo, D.J. & Moore, M.J. A late-acting quality control process for mature eukaryotic rRNAs. *Mol. Cell* **24**, 619–626 (2006).
54. Cole, S.E., LaRiviere, F.J., Merrikh, C.N. & Moore, M.J. A convergence of rRNA and mRNA quality control pathways revealed by mechanistic analysis of nonfunctional rRNA decay. *Mol. Cell* **34**, 440–450 (2009).
55. Gasch, A.P. *et al.* Genomic expression programs in the response of yeast cells to environmental changes. *Mol. Biol. Cell* **11**, 4241–4257 (2000).
56. Warner, J.R. The economics of ribosome biosynthesis in yeast. *Trends Biochem. Sci.* **24**, 437–440 (1999).
57. Viladevall, L. *et al.* Characterization of the calcium-mediated response to alkaline stress in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **279**, 43614–43624 (2004).
58. Engel, S.R. *et al.* The reference genome sequence of *Saccharomyces cerevisiae*: then and now. *G3 (Bethesda)* **4**, 389–398 (2014).
59. Csárdi, G., Franks, A., Choi, D.S., Airoldi, E.M. & Drummond, D.A. Accounting for experimental noise reveals that mRNA levels, amplified by post-transcriptional processes, largely determine steady-state protein levels in yeast. *PLoS Genet.* **11**, e1005206 (2015).
60. Nagalakshmi, U. *et al.* The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* **320**, 1344–1349 (2008).
61. Wang, Y. *et al.* Precision and functional specificity in mRNA decay. *Proc. Natl. Acad. Sci. USA* **99**, 5860–5865 (2002).

ONLINE METHODS

Yeast strains and plasmids. *Saccharomyces cerevisiae* strains used in this study are listed in **Supplementary Table 1** and were obtained from the GE Dharmacon Yeast Knockout Collection or were created by standard recombination techniques[62]. The identity of generated strains was verified by PCR and western blotting. Plasmids used in this work are listed in **Supplementary Table 2**.

Isolation of mRNAs bound to ΔRps26 and wild-type ribosomes. We generated cells for this experiment by transforming strain yKK636 (GAL-Rp26; GAL-Rps3) with pKK3566 (TET–Rps3-TAP). An overnight culture in YPGal supplemented with 0.2 μg/ml dox was grown to mid-log phase. Cells were washed three times with prewarmed YPGal media and inoculated into YPD media at an OD_{600} of 0.15. After 4 h, cells were harvested after the addition of 100 μg/ml cycloheximide (Sigma Aldrich). TAP purification was performed as described[51,63], except that 100 μg/ml cycloheximide was included in all steps and only the IgG-binding and elution steps were performed. The IgG eluate, containing ΔRps26 ribosomes, and the flowthrough, containing +Rps26 ribosomes, were collected and fractionated on a 10–50% sucrose gradient as described[51]. Fractions containing the 80S and polysomes were pooled. Protein was precipitated from these fractions via the TCA–DOC method and analyzed by western blotting. RNA was isolated by phenol-chloroform-isoamyl alcohol extraction.

Illumina sequencing. Purified RNA was quantified in a Qubit 2.0 Fluorometer (Invitrogen) and run on an Agilent 2100 Bioanalyzer (Agilent Technologies) for quality assessment and then treated with DNase I (New England Biolabs). The DNase-treated total RNA was depleted of ribosomal RNA with the Ribo-Zero Gold Yeast Kit (Illumina) and then processed with the TruSeq Stranded Total RNA sample prep kit (Illumina). Briefly, chemically fragmented RNA was random-hexamer primed and reverse-transcribed to generate the first strand of cDNA. The second strand was synthesized incorporating dUTP in place of dTTP, preserving strand information. The double-stranded cDNA was then end-repaired, 3′ adenylated and ligated to PCR adaptors. The purified adaptor-ligated DNA was amplified via PCR to generate the final libraries. The final size of fragments was 200–600 bp with insert sizes in the range of 80–450 bp. The validated libraries were pooled at equimolar ratios and loaded onto the NextSeq 500 flow cell at a final concentration of 1.8 pM.

Bioinformatic processing. Demultiplexed and quality-filtered raw reads (fastq) generated from the NextSeq 500 were trimmed to remove adaptor sequences with Flexbar 2.4 and aligned to the *S. cerevisiae* genome (S288C, SacCer3) from the Saccharomyces Genome Database[58] with TopHat version 2.0.9 (ref. 64). HT seq-count version 0.6.1 was used to generate gene counts, and differential gene expression analysis on three biological replicates was done with DESeq2, using standard settings[65].

Identification and characterization of enriched transcripts. We used existing data sets of mRNA characteristics to analyze the transcripts considered enriched in each pool. Translational efficiency, 5′ and 3′ UTR lengths, mRNA abundance and half-life measurements were derived from existing data sets[59–61]. GO-term enrichment was determined using GENECODIS[37]. Enrichment of nucleotides upstream of the start codon in transcripts that were associated with Rps26-mediated selection was determined by Weblogo analysis[29]. ORF-length data were taken from the SGD_features file listed above. We obtained upstream bases from the *S. cerevisiae* reference genome[58] by using the custom-built program STARTSEQ written in Python.

Purification of ribosomes from TET-Rps26 cells. Cells were grown to mid-log phase and flash-frozen in ribosome buffer (20 mM Hepes/KOH, pH 7.4, 100 mM KOAc, 2.5 mM $Mg(OAc)_2$) supplemented with 1 mg/ml heparin, protease inhibitor cocktail (Sigma) and 2 mM DTT. 400 μL of each clarified yeast lysate was layered over 100 μL of sucrose cushion (Ribo Buffer, 500 mM KCl, 1 M sucrose, 2 mM DTT) and spun in a Beckman TLA 100.1 rotor at 100,000 r.p.m. for 2.5 h. The resulting pellets were resuspended in high-salt buffer (Ribo Buffer, 500 mM KCl, 1 mg/ml heparin, 2 mM DTT) and analyzed by SDS-PAGE followed by western blotting.

qPCR mRNA shift assay. We generated a yeast strain with dox-inducible Rps26 deficiency (TET-Rps26) by transforming yKK491 with pKK3792. We induced depletion of Rps26 by inoculating YPD media (with or without 0.2 μg/ml dox) with TET-Rps26 yeast at an OD_{600} of 0.05. Cultures were grown until they reached mid-log phase and were then prepared for sucrose-gradient fractionation as described above. 5,000 OD units of clarified lysate (~50 μl of lysate) were loaded onto a 25–45% sucrose gradient and spun in an SW-41Ti rotor (Beckman) at 40,000 r.p.m. for 165 min. Gradients were fractionated, and the following samples were collected: (i) unbound RNAs, (ii) free 40S and 60S ribosomal subunits, (iii) 80S monosomes, and one each (iv–xiii) containing two, three, four, etc. ribosomes. We were able to clearly resolve peaks for up to nine bound ribosomes. The final two samples contained 10–11 and 12–13 ribosomes, respectively. After the addition of 0.65 ng of *in vitro*–transcribed RNase P RNA from *Bacillus subtilis* to each sample, which was used for normalization, RNA was isolated by phenol-chloroform extraction from a fixed percentage of each sample's total volume, and reverse transcription was carried out with Protoscript II (New England Biolabs) per the manufacturer's instructions. qPCR was performed with Excella 2× SYBR master mix per the manufacturer's instructions on a BioRad IQ2, using the primers listed in **Supplementary Table 3**. We calculated the percentage of mRNA in each sample by comparing each sample's C_t value to the smallest C_t value obtained for that gradient. This ΔC_t was then transformed into arbitrary reference units of mRNA content, normalized to RNase P levels, and divided by the total mRNA content to represent the fraction of mRNA in that sample. We calculated translational units (TU) by multiplying the mRNA content in each gradient sample by the number of ribosomes bound in that sample and summing those values over all of the gradient fractions. We obtained the ΔTU value by subtracting the TU of each gene in Rps26-replete cells from the TU in Rps26-depleted cells. For each gradient fraction,

$$\text{mRNA content} = \frac{10 \times (\text{RNaseP normalization coefficient})}{2^{(\text{Ct value of gradient sample}) - (\text{Minimum Ct value of gradient})}}$$

$$\text{Fraction of mRNA in each gradient sample} = \frac{\text{mRNA content}}{\sum \text{mRNA content of entire gradient}}$$

$$\text{TU} = \sum (\text{Fraction of mRNA in each sample}) \times (\text{Number of ribosomes in fraction})$$

qPCR from total cellular RNA. Total RNA was isolated from cells growing in mid-log phase by hot phenol extraction. After ethanol precipitation, 1 μg of purified RNA was reverse-transcribed and analyzed by qPCR as described above.

Depletion of Rps3 and Rps15. We generated TET-Rps3 cells by transforming yKK493 with pKK4015, and TET-Rps17 cells by transforming yKK489 with pKK3968. Both strains were grown as described for TET-Rps26.

Luciferase assay. TET-Rps26 cells were transformed with plasmids encoding firefly luciferase preceded by a 10-adenosine upstream sequence and *Renilla* luciferase preceded by one of six sequences (listed in **Supplementary Table 1** and **Supplementary Fig. 4a**). Cells were then grown in selective media, depleted of Rps26 as described above and harvested in mid-log phase. Control experiments demonstrated that the copy numbers of these plasmids did not change upon dox addition (**Supplementary Fig. 4c**), ensuring that observed dox-dependent differences did not arise from differential translation, although we cannot exclude effects on mRNA transcription. For luciferase assays under stress, WT cells were grown for 4 h under stress conditions (or in minimal media as controls), as detailed below, and harvested in mid-log phase. Cells were lysed, and luciferase activity was measured with the Promega Dual-Luciferase Reporter Assay System on a PerkinElmer EnVision 2104 Multilabel Reader according to the manufacturer's protocol, with assay volumes scaled down to 15%.

Stress-response growth curves. For stress-tolerance tests, TET-Rps26 cells were grown and depleted of Rps26 as described above with either 0.2 or 0.1 μg/ml dox as indicated. Cells in mid-log phase were transferred to stress media (or control cultures) at OD 0.1 to test stress tolerance. The composition of stress media was as follows: YPD + 1 M NaCl, or YPD + 10 mM caffeine. For high-pH stress, cells were grown in YPD + 100 mM TAPS buffered to pH 7 (no stress) or pH 8.2 with NaOH. Cells were grown at 30 °C with rapid shaking, and doubling times were measured in the Bioscreen C Automated Microbiology Growth Curve Analysis System (Growth Curves USA) or with a Synergy 2 multi-mode microplate reader (BioTek).

Stress pathway activation. We tested HOG pathway activation by growing TET-Rps26 cells to mid-log phase at 30 °C in YPD with or without dox, as described above. Cells were then transferred to YPD containing 0 mM, 100 mM, 300 mM or 500 mM NaCl for 5 min, collected, and analyzed by western blotting. Rim101 pathway activation was tested as described[39] with TET-Rps26 cells transformed with pKK3678 (3HA-Rim101).

Ribosome purification from stressed cells. BY4741 yeast cells were grown to mid-log phase before being seeded into different media. Cells were inoculated into stress media at a starting OD of 0.7 and harvested after 4 h. Ribosomes were purified as previously described[66].

Antibodies. The phospho-p38 (D3F9) antibody from Cell Signaling Technologies was used to detect Hog1-p[67]. HA-tagged Rim101 was detected with anti-HA (HA.C5) from Abcam (ab18181). Anti-TAP (CAB1001) was from Thermo Fisher Scientific. Anti-Rps5 was from ProteinTech (16964-1-AP). Polyclonal antibodies were gifts from M. Seedorf, Universität Heidelberg (Rps3), G. Dieci, Università di Parma (Rps8)[68], L. Valášek, Czech Academy of Sciences (Rps0) and A. Link, Vanderbilt University (Asc1). The monoclonal antibody against Tub1, developed by J. Frankel, was obtained from the Developmental Studies Hybridoma Bank, created by the NICHD of the NIH and maintained at The University of Iowa. Antibodies to Fap7 and Rcl1 were raised against purified recombinant proteins by Josman, LLC. Antibodies to Rps10 and Rps26 were raised against peptides from each protein by New England Peptide. These four antibodies were tested against yeast lysates and either recombinant protein or purified 40S ribosomal subunits.

Statistical analysis. Various statistical tests were used as appropriate and as indicated in the respective figure legends. Unpaired, two-tailed, Student's *t* tests were used on small data sets. For larger data sets, or those in which there was not an assumption or normality, the nonparametric Kolmogorov–Smirnov test was used. For testing changes in growth-curve rates from a hypothetical value of 1 (no change), the nonparametric two-tailed Wilcoxon signed rank test was used. Finally, a two-way ANOVA with Holm–Sidak correction for multiple comparisons was used for data sets in which two factors (for example, cell type and reporter construct) were tested.

A **Life Sciences Reporting Summary** for this article is available.

Data availability statement. Sequencing data have been deposited at NCBI-GEO under accession code GSE86203. Source data for **Figures 1–5** and **Supplementary Figures 1–6** are available with the paper online. Other data and custom scripts will be made available upon request.

62. Longtine, M.S. *et al.* Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast* **14**, 953–961 (1998).
63. Strunk, B.S. *et al.* Ribosome assembly factors prevent premature translation initiation by 40S assembly intermediates. *Science* **333**, 1449–1453 (2011).
64. Trapnell, C., Pachter, L. & Salzberg, S.L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
65. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
66. Acker, M.G., Kolitz, S.E., Mitchell, S.F., Nanda, J.S. & Lorsch, J.R. Reconstitution of yeast translation initiation. *Methods Enzymol.* **430**, 111–145 (2007).
67. Tanigawa, M., Kihara, A., Terashima, M., Takahara, T. & Maeda, T. Sphingolipids regulate the yeast high-osmolarity glycerol response pathway. *Mol. Cell. Biol.* **32**, 2861–2870 (2012).
68. Dieci, G., Bottarelli, L. & Ottonello, S. A general procedure for the production of antibody reagents against eukaryotic ribosomal proteins. *Protein Pept. Lett.* **12**, 555–560 (2005).

nature research

Corresponding Author: Katrin Karbstein

Date: Jun 8, 2017

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work we publish. This form is published with all life science papers and is intended to promote consistency and transparency in reporting. All life sciences submissions use this form; while some list items might not apply to an individual manuscript, all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

   Describe how sample size was determined.

   For sequencing, ribosome purification and mRNA shift experiments, sample size was limited by the cost of assays. For growth curves and other experiments, the n was chosen to be large enough to identify any biologically relevant phenotypes.

2. Data exclusions

   Describe any data exclusions.

   For quantification of western blot images, bands were excluded if they were distorted or obscured by a transfer or blotting artifact. All other data was included.

3. Replication

   Describe whether the experimental findings were reliably reproduced.

   Experimental findings for pull downs, mRNA shift assays, phenotypic assays, etc were successfully replicated with minor variance between experiments (which is included in the data and accounted for via the indicated statistical tests).

4. Randomization

   Describe how samples/organisms/participants were allocated into experimental groups.

   Yeast strains used for this experiment were clonal. When different treatments were applied to a strain, cells always originated from a common culture.

5. Blinding

   Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

   No, data analysis was performed by the same researcher who performed the experiment so blinding was not possible. Instead, all data analysis was performed using a standard set of steps that were held consistent between samples.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

Nature Structural & Molecular Biology: doi:10.1038/nsmb.3442

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or the Methods section if additional space is needed).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒	A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly.
☐	☒	A statement indicating how many times each experiment was replicated
☐	☒	The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒	A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒	The test results (e.g. *p* values) given as exact values whenever possible and with confidence intervals noted
☐	☒	A summary of the descriptive statistics, including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒	Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.	Statistical tests were performed with GraphPad Prism 6, sequencing analysis is described in the methods section, upstream sequences of genes were retrieved using a custom script that is available upon request.

For all studies, we encourage code deposition in a community repository (e.g. GitHub). Authors must make computer code available to editors and reviewers upon request. The *Nature Methods* guidance for providing algorithms and software for publication may be useful for any submission.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.	All strains and plasmids generated for this study are freely available from our lab. Any commercial reagents are noted in the methods section.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).	All antibodies are described (with references or product numbers, as appropriate) in the methods section (page 22).

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	All yeast strains used are described in supplementary table 3
b. Describe the method of cell line authentication used.	Strains were validated by PCR and western blot (when possible)
c. Report whether the cell lines were tested for mycoplasma contamination.	N/A
d. If any of the cell lines used in the paper are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	N/A

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.	N/A

Nature Structural & Molecular Biology: doi:10.1038/nsmb.3442

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.	N/A

Nature Structural & Molecular Biology: doi:10.1038/nsmb.3442