


RESEARCH ARTICLE

Open Access



# Temperature-dependent regulation of upstream open reading frame translation in *S. cerevisiae*

Shardul D. Kulkarni<sup>1</sup>, Fujun Zhou<sup>1</sup>, Neelam Dabas Sen<sup>2,3</sup>, Hongen Zhang<sup>2</sup>, Alan G. Hinnebusch<sup>2\*</sup> and Jon R. Lorsch<sup>1\*</sup> 

## Abstract

**Background:** Translation of an mRNA in eukaryotes starts at an AUG codon in most cases, but near-cognate codons (NCCs) such as UUG, ACG, and AUU can also be used as start sites at low levels in *Saccharomyces cerevisiae*. Initiation from NCCs or AUGs in the 5'-untranslated regions (UTRs) of mRNAs can lead to translation of upstream open reading frames (uORFs) that might regulate expression of the main ORF (mORF). Although there is some circumstantial evidence that the translation of uORFs can be affected by environmental conditions, little is known about how it is affected by changes in growth temperature.

**Results:** Using reporter assays, we found that changes in growth temperature can affect translation from NCC start sites in yeast cells, suggesting the possibility that gene expression could be regulated by temperature by altering use of different uORF start codons. Using ribosome profiling, we provide evidence that growth temperature regulates the efficiency of translation of nearly 200 uORFs in *S. cerevisiae*. Of these uORFs, most that start with an AUG codon have increased translational efficiency at 37 °C relative to 30 °C and decreased efficiency at 20 °C. For translationally regulated uORFs starting with NCCs, we did not observe a general trend for the direction of regulation as a function of temperature, suggesting mRNA-specific features can determine the mode of temperature-dependent regulation. Consistent with this conclusion, the position of the uORFs in the 5'-leader relative to the 5'-cap and the start codon of the main ORF correlates with the direction of temperature-dependent regulation of uORF translation. We have identified several novel cases in which changes in uORF translation are inversely correlated with changes in the translational efficiency of the downstream main ORF. Our data suggest that translation of these mRNAs is subject to temperature-dependent, uORF-mediated regulation.

**Conclusions:** Our data suggest that alterations in the translation of specific uORFs by temperature can regulate gene expression in *S. cerevisiae*.

## Background

In eukaryotes, the recognition of the start codon in an mRNA during cap-dependent translation initiation is thought to occur by the scanning mechanism [1]. It begins with the formation of a ternary complex (TC) that consists of translation initiation factor 2 (eIF2) in its

GTP-bound form along with the methionyl initiator tRNA (Met-tRNA<sub>i</sub>). The TC binds to the small (40S) ribosomal subunit with the aid of eIFs 1, 1A, and 3 to form the 43S pre-initiation complex (PIC). The PIC binds to the mRNA near the 5'-cap with the aid of a group of initiation factors including the eIF4F complex, eIF3, and the poly(A) binding protein (PABP). The PIC then scans along the mRNA in a 5' to 3' direction in search of the start codon, which in most cases is an AUG. Upon recognition of the start codon, sequential events occur in the PIC that eventually leads to joining of the large (60S) ribosomal subunit and commencement of the elongation phase of protein synthesis. This series

\* Correspondence: [alanh@mail.nih.gov](mailto:alanh@mail.nih.gov); [jon.lorsch@nih.gov](mailto:jon.lorsch@nih.gov)

<sup>2</sup>Laboratory of Gene Regulation and Development, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD, USA

<sup>1</sup>Laboratory on the Mechanism and Regulation of Protein Synthesis, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD, USA

Full list of author information is available at the end of the article



of events starts with stable base pairing between the start codon and tRNA<sub>i</sub> anticodon, which triggers the ejection of eIF1 from its binding site on the 40S subunit. Ejection of eIF1 from the PIC leads to the conversion of eIF2 to its inactive, GDP-bound state as a result of gated phosphate (P<sub>i</sub>) release. This converts the open, scanning-competent PIC into a closed, scanning-arrested PIC. The release of the other initiation factors and joining of the 60S ribosomal subunit results in the formation of the 80S initiation complex which is now competent for translation elongation [2, 3].

The selection of the start codon in an mRNA by the translational machinery is one of the key steps of reading the genetic code. It defines the N-terminus of the translated protein as well as the reading frame for decoding. Although AUG is the start codon of most main ORFs, codons that differ from AUG by only one base (“near-cognate codons” (NCCs)) can also be utilized as start sites to varying degrees in *S. cerevisiae* [4–6]. The use of alternative start codons, such as AUGs in weak sequence context or NCCs, is a potential mechanism to regulate gene expression [7–10]. For example, in addition to their main open reading frames (mORFs), some mRNAs contain one or more upstream open reading frames (uORF) in their 5′-leaders which can begin with AUGs in a strong or weak sequence context or with NCCs [11]. The recognition and translation of uORFs can regulate the expression of the downstream mORF by various mechanisms, such as altering the level of the mRNA by triggering nonsense-mediated decay or by preventing PICs from reaching the mORF start codon [12, 13]. Several studies using ribosome-profiling have provided evidence that uORF translation is altered in response to a variety of stress conditions [8, 14]. It has also been reported that some mRNAs have multiple in-frame AUGs or NCCs that can be used as alternative start sites that can lead to the production of protein isoforms with N-terminal extensions. These alternative initiation codons can be conserved throughout eukaryotes, suggesting their functional importance [11]. The protein isoforms with and without the extension have been reported to localize differentially [5, 6, 15, 16], and a recent proteomic analysis of the yeast “N-terminome” indicates that ~10% of yeast mRNAs have alternative, in-frame start codons that are utilized some fraction of the time [17].

A number of components of the eukaryotic translation initiation machinery have been shown to be involved in start codon recognition, including initiation factors such as eIF1, eIF1A, eIF2, and eIF5, tRNA<sub>i</sub>, rRNA, and mRNA elements [18]. Mutations in these components can increase the efficiency with which NCCs are used as start sites, producing a phenotype referred to as suppressor of initiation codon mutation (Sui<sup>-</sup>). A mutation in eIF1A with the opposite effect on fidelity was shown to confer

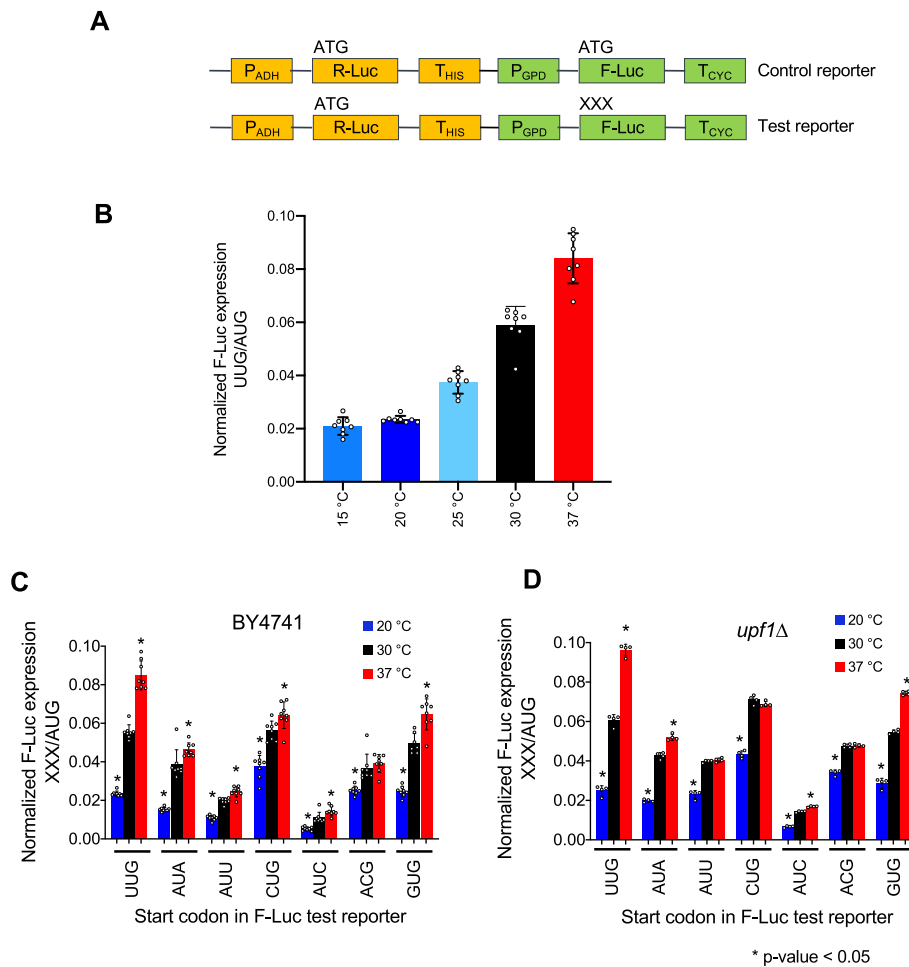
heightened discrimination genome-wide against AUGs in poor sequence context [19]. The use of NCC start sites in a reporter mRNA in yeast was also shown to be enhanced by two small molecules identified in a high-throughput screen [20], indicating that external agents can modulate the fidelity of start codon recognition.

We undertook this study starting with the hypothesis that the fidelity of the start codon recognition might be a point of post-transcriptional regulation of gene expression. Changes in the fidelity of the start codon recognition in response to external or internal stimuli could rapidly modify the proteome by changing the balance of translation of uORFs, N-terminal extensions, and main ORFs. In an attempt to test this hypothesis, we used the same dual-luciferase reporter assay used for the high-throughput chemical screen for compounds that alter the start codon recognition in *S. cerevisiae* [20] to search for other external stimuli that produce similar effects. We found that growth temperature modulates the use of near-cognate start codons in both the luciferase reporter system and an orthogonal *lacZ*-based system. However, when we used ribosome profiling to observe the effects of growth temperature on the translation of uORFs transcriptome-wide, we found a more complicated distribution of effects than was suggested by the reporter assays. Although the translation of most uORFs is not significantly affected by changes in growth temperature, a subset of uORFs are regulated by temperature shifts, with various combinations of increased or decreased translational efficiency at high or low temperature. Of the regulated uORFs, those starting with AUG are generally repressed at 20 °C and activated at 37 °C relative to their translation at 30 °C, whereas those starting with NCCs display a more distributed set of effects. The position of the uORF in the 5′-UTR and the length and degree of structure of the UTR appear to influence the effect of temperature on translation. We present a number of novel cases of temperature-dependent changes in uORF translation in which there are reciprocal changes in the main ORF translation, suggesting uORF-mediated regulation of main ORF expression.

## Results

### Growth temperature affects the efficiency of using non-AUG start codons in reporter mRNAs in yeast

We previously developed and validated a dual-luciferase assay to calculate the efficiency of utilization of near-cognate codons (NCCs) as translational start sites in yeast [20]. In this assay, Renilla luciferase (R-Luc) and Firefly luciferase (F-Luc) are expressed using separate promoters and transcription terminators from a single low-copy plasmid (Fig. 1a). R-Luc mRNA has an AUG as the start site and acts as an internal control for cell growth, lysis efficiency, and pipetting inconsistency. The



**Fig. 1** Changes in temperature result in changes in relative expression of firefly luciferase (F-Luc) reporters with AUG and near cognate codons (NCCs) as start sites. **a** Schematic of the reporters used in the study. P, promoter; T, terminator. Renilla luciferase (R-Luc) mRNA was produced from the ADH promoter and HIS terminator, and Firefly luciferase (F-Luc) mRNA was produced from the GPD promoter and CYC terminator. “XXX” in the F-Luc test reporter represents a start codon that varies from AUG by 1 bp (NCC). The relative expression from Firefly luciferase with respect to Renilla luciferase was calculated for both test (F-Luc<sup>XXX</sup>/R-Luc<sup>AUG</sup>) and control (F-Luc<sup>AUG</sup>/R-Luc<sup>AUG</sup>) reporters. The normalized expression was then calculated as the ratio of relative F-Luc<sup>XXX</sup> expression from the test reporter to relative F-Luc<sup>AUG</sup> expression from the control reporter (XXX/AUG ratio). **b** Normalized F-Luc<sup>UUG</sup> expression (UUG/AUG ratio) was measured in yeast strain BY4741 at different temperatures. **c** Normalized expression from F-Luc reporters (XXX/AUG ratio) with different NCCs as start sites in BY4741 cells cultured at multiple temperatures. **d** Same as **c**, except the experiments were done in  $\Delta upf1$  cells. For **b-d**, the results are represented as averages of at least four biological replicates with error bars representing standard deviation of mean. Each circle on the individual columns represents one biological replicate. For **b-d**, the asterisk denotes *p* values < 0.05 calculated by Student’s *t* test when compared with 30 °C

start codon of the F-Luc reporter is varied and can be AUG or any NCC (e.g., UUG, ACG). F-Luc expression is normalized to that of R-Luc to control for any global changes in gene expression or cell growth. F-Luc<sup>AUG</sup>/R-Luc<sup>AUG</sup> represents the normalized expression value for a F-Luc reporter starting with AUG. The relative expression from UUG (or any other near-cognate) is calculated by normalizing with respect to the normalized F-Luc<sup>AUG</sup> expression (F-Luc<sup>UUG</sup>/R-Luc<sup>AUG</sup>)/(F-Luc<sup>AUG</sup>/R-Luc<sup>AUG</sup>).

We used this assay to investigate the effects of changes in growth temperature on the use of NCCs as translational start sites. Yeast cells (BY4741) were transformed

with the dual-luciferase reporter plasmid with either an F-Luc<sup>AUG</sup> or F-Luc<sup>UUG</sup> gene, cultured at various temperatures, and the luciferase activity of the F-Luc<sup>UUG</sup> reporter relative to the F-Luc<sup>AUG</sup> reporter was measured. (F-Luc and R-Luc assays in cell extracts were performed at 24 °C regardless of the temperature at which the yeast cells were cultured). Elevating the growth temperature from 30 °C (which is the optimal yeast growth temperature) to 37 °C led to ~1.5-fold increase in the normalized expression of F-Luc<sup>UUG</sup>, while lowering the growth temperature from 30 to 25 °C or 20 °C led to ~1.6-fold and ~2.5-fold reduction, respectively, in the

normalized F-Luc<sup>UUG</sup> expression (Fig. 1b). These findings suggest that at lower growth temperatures, the efficiency of use of the F-Luc UUG start codon is decreased and that at higher temperatures, it is increased.

To try to confirm these findings and test the generalizability of the observed effects of temperature on start codon usage, we performed a similar experiment with two otherwise identical *HIS4-lacZ* fusion reporters with an AUG or UUG start codon [21] (Additional file 1: Figure S1A). The normalized expression of a *HIS4-lacZ* reporter with UUG as a translational start site was reduced ~5-fold at 20 °C with respect to 30 °C and was elevated ~2.8-fold at 37 °C (Additional file 1: Figure S1B). These results are consistent with the findings from the dual-luciferase reporter that the use of a UUG relative to an AUG start codon is reduced at 20 °C and increased at 37 °C. Frequently, the finding that two orthogonal reporter assays give similar results might be taken to indicate that the observed effect is generalizable to most mRNAs. However, as described below, this turns out not to be the case in this system.

We next tested the effects of changes in growth temperature on the normalized expression from F-Luc reporter mRNAs starting with other NCCs (GUG, CUG, ACG, AUA, AUC, and AUU), all of which have been shown to be utilized as start sites in yeast cells to varying extents [4, 20]. Like UUG, the normalized expression from all NCCs was significantly lowered (~2-fold) at 20 °C indicating that the effects of lowering the growth temperature were not specific to the UUG start site (Fig. 1c). On the other hand, elevation in growth temperature resulted in differential changes in normalized expression, ranging from no change for ACG to ~30% increase (for AUC and GUG) as compared to their expression at 30 °C. This suggested that the efficiency of the use of NCCs might be differentially affected at some temperatures.

The changes in the expression of the F-Luc reporters starting with different initiation codons could be due to the changes in mRNA stability induced by the nonsense-mediated decay (NMD) pathway triggered by altered translation of an upstream open reading frame (uORF) starting from an NCC. In the absence of an AUG start codon, it is possible that an upstream or out-of-frame NCC codon is used for initiation leading to premature termination, and potentially NMD [22]. Alternatively, it is also possible that the scanning 43S pre-initiation complexes bypass the NCC start site (leaky scanning), initiate at an out-of-frame AUG or NCC in the mORF, terminate in the mORF, and thereby trigger NMD [23]. To test these possibilities, we performed the luciferase assay in a strain in which the *UPF1* gene, which encodes a protein essential for NMD, had been deleted (*upf1Δ*) [24]. Deletion of *UPF1* did not affect the temperature-dependent

changes in expression observed in WT cells for the majority of the start codons tested at both 20 °C and 37 °C (Fig. 1d). This result suggests that the NMD pathway does not play a role in the observed changes in expression from these reporters at different temperatures.

The alterations in the use of UUG as a start site could be attributed to the changes in the levels of eIF1, which has been shown to be a “gatekeeper” in the start codon recognition process [25], helping to restrict start codon selection to AUGs and block the selection of NCCs. To test the possibility that temperature-dependent changes in eIF1 expression might be responsible for the observed effects of temperature on start codon usage, we assessed the levels of the factor using western blot analysis of whole-cell lysates from cells cultured at different temperatures. Levels of eIF1 protein were not significantly altered at 20 or 37 °C relative to 30 °C (Additional file 1: Figure S2A), suggesting changes in start codon use are not due to changes in eIF1 concentration.

We also tested the effect of over-expression of eIF1 on the observed temperature dependence of F-Luc start codon utilization. Over-expression of eIF1 from a high-copy (hc) plasmid has been shown to suppress the reduced stringency of start codon recognition (*Sui*<sup>-</sup>) phenotype caused by mutations in several initiation factors [26–29]. Consistent with its role as a central gatekeeper of start codon recognition, over-expression of eIF1 (hc-*SUI1*) suppressed the use of UUG as a start site at all three temperatures (Additional file 1: Figure S2B). The decrease in F-Luc<sup>UUG</sup> expression at 20 °C and the increase at 37 °C were still observed in the hc-*SUI1* strain, although the magnitude of the increase at 37 °C was reduced relative to WT cells in this experiment. Consistent with these results, reducing the concentration of eIF1 relative to WT cells by using a *SUI1/sui1Δ* heterozygous diploid strain resulted in increased expression of F-Luc<sup>UUG</sup> relative to F-Luc<sup>AUG</sup> at all three temperatures relative to expression in a WT diploid containing two wild-type chromosomal alleles of *SUI1* (Additional file 1: Figure S2C). No decrease in the magnitude of the temperature dependence of normalized F-Luc<sup>UUG</sup> expression was observed in the *SUI1/sui1Δ*-haploinsufficient diploid. In addition, haploinsufficiency of eIF1A (+/*tif11Δ*) or eIF5 (+/*tif5Δ*), also factors involved in the start codon recognition, did not significantly alter the effect of temperature on expression of F-Luc<sup>UUG</sup> relative to F-Luc<sup>AUG</sup> (Additional file 1: Figure S2C). It is noteworthy that, for reasons unknown, the increase at 37 °C is dampened in the WT *SUI1/SUI1* diploid versus the *SUI1* haploid strain analyzed in Fig. 1a and that the larger differences resurface in the *SUI1/sui1Δ* heterozygote. Although altering the dosage of the *SUI1* gene appears to modulate somewhat the effects of 37 °C on UUG initiation, overall, it appears that the effects of temperature

on NCC utilization are not dictated by altered cellular levels of eIF1.

Taken together, these results suggest that the effect of temperature on F-Luc<sup>UUG</sup> expression is not due to the changes in the concentrations of eIF1, eIF1A, or eIF5. Additionally, when we performed ribosome profiling in cells grown at different temperatures (see below), we did not observe any obvious changes in the ribosome protected fragments (RPF) counts and translational efficiencies for the mORF of eIF1, further confirming that the level of eIF1 does not change as a function of growth temperature in WT haploid cells (Additional file 1: Figure S2D). We cannot, however, rule out possible changes in subcellular localization or modification of eIF1 or other initiation factors.

#### Ribosome profiling elucidates temperature-dependent changes in start codon utilization transcriptome-wide

To investigate the effects of changes in growth temperature on the relative use of different codons as translational start sites throughout the transcriptome, we performed ribosome profiling in yeast cells cultured at multiple temperatures. WT yeast cells (BY4741) transformed with the F-Luc<sup>UUG</sup> reporter plasmid were cultured in SC-Ura for 16 h at 20, 30, and 37 °C, and ribosome profiling was performed (Fig. 2a) as previously described [8, 30], with some notable changes (see the “Materials and methods” section). In particular, we did not add cycloheximide to the intact cells to stop translation because of the known artifacts it creates [31–34], but instead, flash froze the cells in liquid nitrogen and added cycloheximide to the cell lysis buffer only.

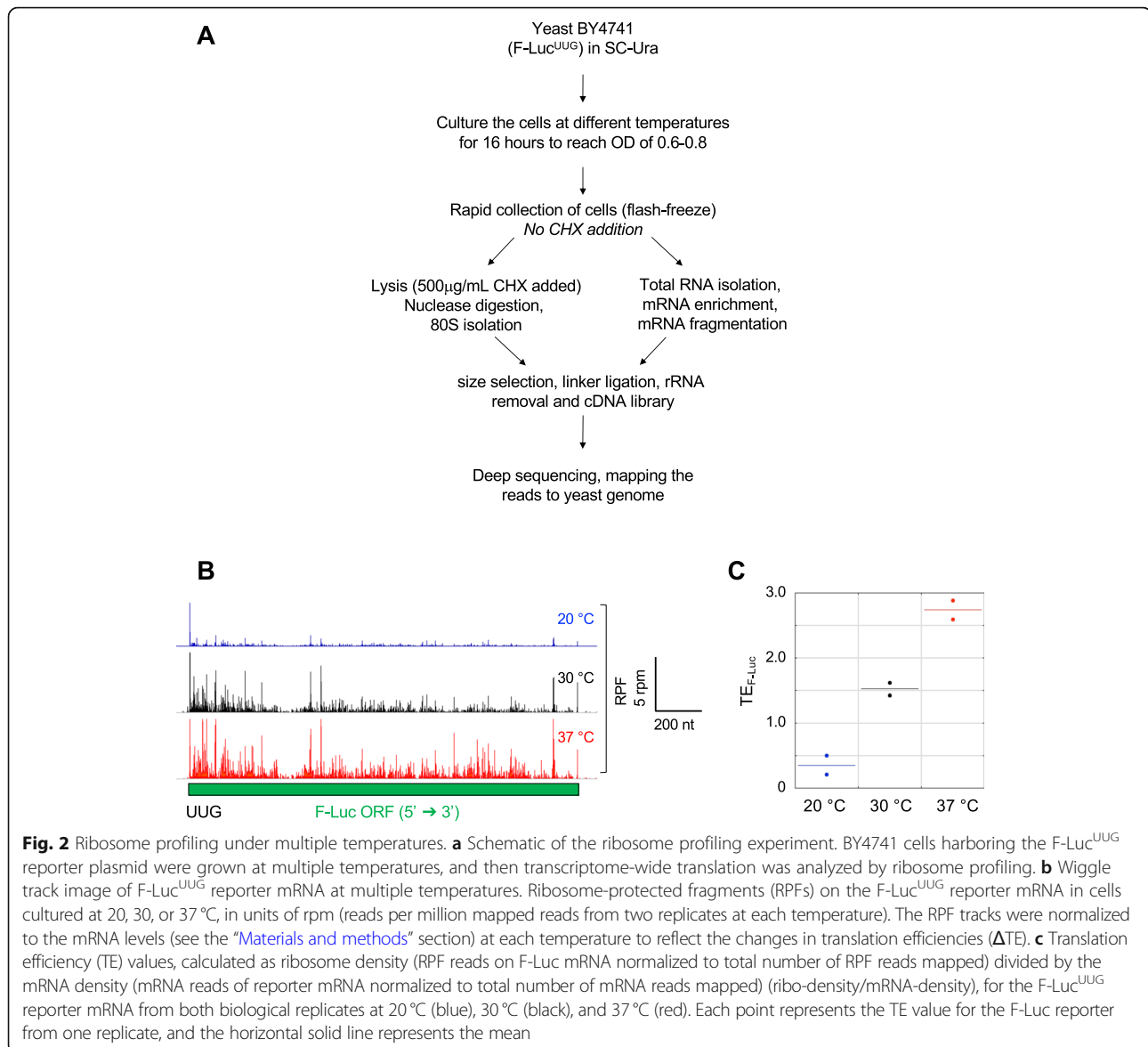
We calculated ribosomal read density as the number of 80S ribosomal footprint reads mapped to an mRNA sequence relative to the total number of reads in the footprint library (ribo-seq), and we calculated the mRNA read density by normalizing RNA-seq reads mapped to an mRNA sequence relative to the total number of reads in the RNA-seq library. The translation efficiency (TE) for each mRNA is calculated as ribosomal read density normalized to mRNA read density [8]. The ribosome footprint and RNA-seq data indicate that the two biological replicates were highly reproducible for all the temperatures (Pearson’s  $r > 0.98$ ; Additional file 1: Figure S3A).

We mapped the ribo-seq (ribosome protected fragments, RPFs) and RNA-seq reads to the F-Luc<sup>UUG</sup> mRNA reporter to confirm our findings with the reporter assay. The expression of F-Luc<sup>UUG</sup> mRNA reporter was significantly altered at multiple temperatures (Fig. 2b and Additional file 1: Figure S3B). The RPF read count for the F-Luc mRNA was decreased at 20 °C by ~25% and increased at 37 °C by ~46% as compared to 30 °C (Additional file 1: Figure S3B). In contrast, the mRNA read count for the F-Luc mRNA was increased

at 20 °C by ~300% and decreased at 37 °C by ~20%, as compared to 30 °C (Additional file 1: Figure S3B). To control for the changes in mRNA abundance, we calculated the TE of the F-Luc reporter (Fig. 2c). TE for F-Luc<sup>UUG</sup> mRNA was reduced at 20 °C by ~77% and elevated at 37 °C by ~80%, as compared to 30 °C, indicating that the translation of F-Luc mRNA is significantly altered at 20 °C and 37 °C, consistent with the findings from the luciferase assay where the normalized expression of this reporter (with respect to the F-Luc<sup>AUG</sup> control reporter, Fig. 1a) was altered in an analogous manner (Fig. 1b, c). We also examined the RPF and RNA-seq reads for the R-Luc<sup>AUG</sup> mRNA reporter (Additional file 1: Figure S3C, S3D), and the data suggested that R-Luc mRNA might undergo translational regulation at 20 °C. In addition, the R-Luc mRNA abundance increases ~5-fold at 20 °C relative to 30 or 37 °C. These data further indicate the caution that is warranted when interpreting results from reporter assays.

#### Identification of upstream open reading frames using ribosome profiling data

To better understand the effects of changes in temperature on start site selection, we identified a set of 1367 uORFs that show evidence of translation in our yeast strain at one or more temperatures. To this end, we employed a two-step strategy for the translated uORF discovery described previously [19] (Fig. 3a). The first step employs the Yassour-uORF algorithm [35] in which putative translated uORFs are identified from among the set of all possible uORFs initiating with an AUG or NCC on the basis of a strong peak of ribosome density at the start codon and the occurrence of >50% of downstream read counts in the zero frame of the start codon. After excluding uORFs shorter than 3 codons, we identified 6061 potential uORFs by applying this algorithm to several previously reported ribosome profiling datasets (see the “Materials and methods” section) as well as the dataset generated in the current study. In the second step, we examined which of these 6061 putative uORFs show evidence of translation in our ribosome profiling data using a different identification tool, RibORF [36], which is based on the criteria of 3-nt periodicity (a hallmark of mRNA fragments protected by actively translating ribosomes) and a uniform distribution of reads across uORF codons. This tool generates a *predicted translating probability* ranging from 0 to 1. Lower probability values indicate skewed distributions of reads and equally distributed fractions of reads at the zero, first and second reading frames. Higher values indicate a uniform distribution of reads and a majority of reads aligned in the zero reading frame. Applying a moderately stringent probability of prediction of >0.5, we found evidence for translation in our datasets for 1367 uORFs among the 6061 potential uORFs detected in the

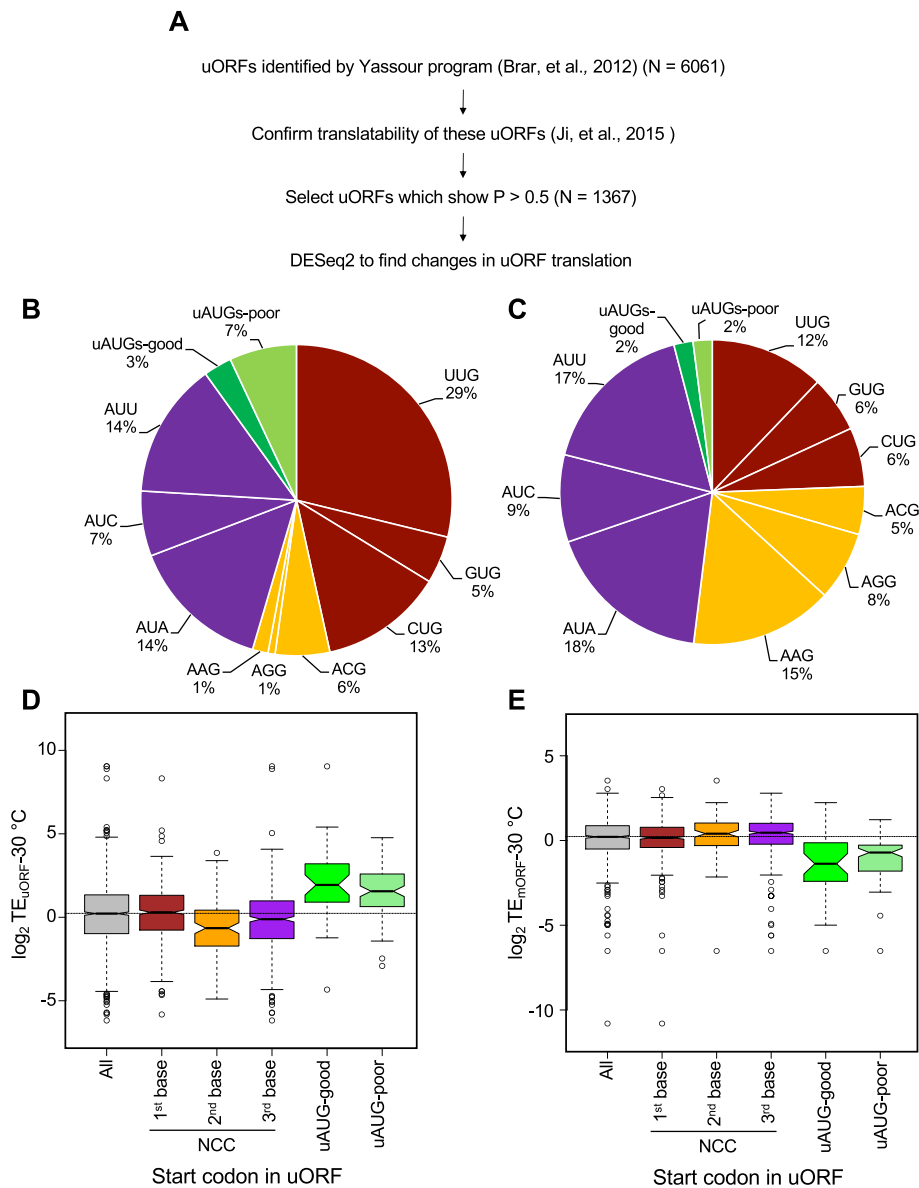


first step, located on 755 different mRNAs. Boxplot analysis of the fraction of reads from each of the three frames for all the identified uORFs confirmed the periodicity of RPF reads (Additional file 1: Figure S3E). These uORFs were further analyzed for changes in their expression at different temperatures using DESeq2 (see the “Materials and methods” section).

These 1367 translated uORFs start with either an AUG (~10%) or a NCC (~90%). We observed a range of NCCs as start sites (Fig. 3b), with UUG the most common (~30% of all uORFs) and AGG the least (~1% of all uORFs). The uORFs with near-cognate start codons with the second base changes from AUG (AAG, ACG, AGG) contributed only 8% of all the uORF start codons, with ACG at 6% and both AAG and AGG at 1%, which is consistent with previous findings indicating

that AAG and AGG are the least efficiently used near-cognate start codons in yeast cells [4]. Near-cognate codons with the first base changes (UUG, GUG, CUG) comprised ~50% of the total uORF start codons indicating that they are the most efficient near-cognate start sites, also consistent with previous studies [4] and our luciferase reporter analyses (Fig. 1c). Among the ~10% of all uORFs with an AUG start codon, ~33% have a preferred Kozak context at the -3 position (A/G), while the remaining ~66% have poor context (U/C at -3 position) [37, 38] (Fig. 3b), which is consistent with previous studies suggesting that good context uAUGs have been selected against evolutionarily [9].

We also looked at the overall abundance of AUGs and NCCs in the 5'-UTR transcriptome after removing any that initiate predicted uORFs less than three codons in



**Fig. 3** Identification of upstream open reading frames (uORFs) in *S. cerevisiae*. **a** Multiple datasets described previously in [19] were used to identify potential uORFs using a program described previously [35]. The potential uORFs  $\geq 3$  codons ( $N = 6061$ ) were run through a program previously described [36] to find 1367 uORFs showing evidence of translation in the combined dataset generated in this study (each replicate, all temperatures). The differential expression at 20 °C or 37 °C with respect to 30 °C was analyzed by DESeq2. **b** Distribution of start sites in translated uORFs. The AUG uORFs were sorted into good context (A/G) or poor (U/C) based on the nucleotide at the  $-3$  position with respect to AUG. Percentages were calculated with respect to the total number of translated uORFs. **c** Transcriptome-wide distribution of potential upstream start sites (USSs). The USSs that can lead to formation of uORFs  $< 3$  codons were excluded so that an appropriate comparison with the translated set of uORFs could be done. Percentages were calculated with respect to the total number of USSs ( $N = 58,353$ ). **d** Analysis of  $TE_{uORF}$  at 30 °C for both NCC and AUG uORFs. *All* indicates all the translated uORFs identified through our pipeline ( $N \sim 1350$ ); *first base change*: uORFs with UUG, CUG, GUG ( $N = 631$ ); *second base change*: uORFs with AAG, ACG, AGG ( $N = 109$ ); and *third base change*: uORFs with AUC, AUA, AUU ( $N = 483$ ). AUG uORFs were sorted into good context ( $N = 36$ ) or poor ( $N = 93$ ). The dotted horizontal line indicates the median TE value of *All*. **e** Analysis of TE of mORFs located downstream of the various uORFs at 30 °C. *All* indicates the set of mORFs located downstream of all translated uORFs analyzed in this study ( $N = 748$ )

length (in order to match the conditions used for identification of translated uORFs in our datasets, which also eliminated potential uORFs less than three codons long; Fig. 3c). Comparison of the start codon distribution of the set of translated uORFs we identified (Fig. 3b) with

the 5'-UTR transcriptome abundance of potential uAUG and NCC start codons (Fig. 3c) indicates that NCCs with first position changes are over-represented as start codons for the translated uORFs relative to their inherent abundance and those with second position changes are

under-represented. These results are consistent with the general efficiency of use of these classes of NCCs as start codons. Also as expected, AUG codons are used more frequently as start codons for translated uORFs than their representation among potential start codons in 5'-UTRs.

#### uORFs with different NCCs as start sites show differential translation efficiencies

To investigate the translatability of the uORFs, we calculated translational efficiency (TE) at 30 °C for all the uORFs starting with AUGs or NCCs. TE, as described above, is the ratio of ribosomal footprint read density to mRNA read density. As shown in Fig. 3d, uORFs starting with different initiation codons were translated with differing median efficiencies at 30 °C. As might be expected, the median TE for uORFs starting with AUG codons (AUG uORFs) was significantly higher than the medians for uORFs starting with any NCC (NCC uORFs). In this and all other box and whisker plots below, the lack of overlap in the notches of two adjacent plots indicates that their medians differ with > 95% confidence (Chambers et al., 1983 Graphical methods for data analysis. Wadsworth, Belmont). The median TE for AUG uORFs in good context was comparable to that for AUG uORFs in poor context (Fig. 3d), suggesting that other features of these mRNAs might modulate the effect of sequence context for this subset of uORFs.

The median TEs for NCC uORFs varied by a factor of about two to threefold, depending on which base varied from AUG. Consistent with our analysis of the start codon distribution of translated uORFs (Fig. 3b), uORFs starting with NCCs with the second base changes were the least efficiently translated, whereas uORFs starting with NCCs with the first base changes were translated the most efficiently (Fig. 3d). These data, together with the results described in Fig. 3b, suggest differential recognition and utilization of NCCs as start sites for uORFs in yeast in a manner consistent with previous analyses of the efficiency of different start codons for main ORF translation [5, 6, 20].

We next examined the relationship between the translation efficiency of the uORFs and that of the downstream main ORF (mORF). We calculated the TEs for the mORFs downstream of each subset of uORFs grouped according to the uORF start codon. We observed that the median TEs for the mORFs downstream of AUG uORFs were significantly lower than any other group (Fig. 3e), suggesting that uORFs with AUG start sites were typically inhibitory of downstream translation at 30 °C, as expected. The median TE of mORFs downstream of uORFs starting with near-cognate codons with the first base changes was essentially the same as for all mORFs with translated uORFs, whereas mORFs

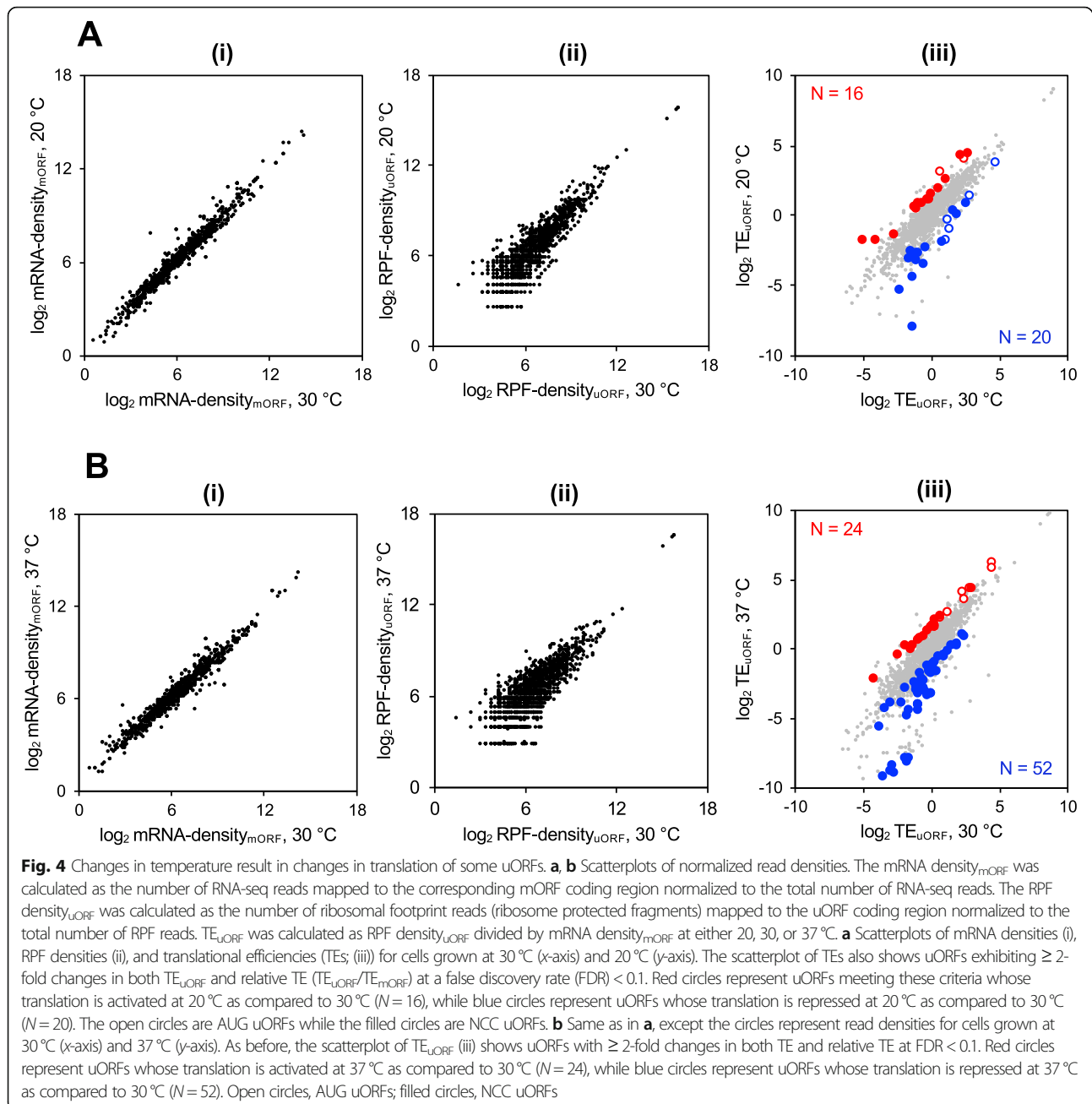
downstream of the more poorly translated uORFs starting with near-cognates with the second or third base changes had a slightly higher median TE compared to all mORFs (compare "All" in Fig. 3e with the first, second, third base changes and uAUGs). Overall, these data provide a transcriptome-wide view of uORF translation and the inverse correlation between uORF and mORF TE values. These data are consistent with previous reports that the presence of uAUGs in 5'-UTRs of yeast mRNAs is inversely correlated with the polysome density of the mRNAs [9].

#### Changes in growth temperature have varying effects on uORF translation for different mRNAs

To examine the effects of changes in growth temperature on the translation of AUG and NCC uORFs, we calculated the changes in TEs of uORFs at either 20 °C or 37 °C with respect to 30 °C. The TE of a uORF ( $TE_{uORF}$ ) at any given temperature was defined as the ribosome footprint (RPF) density of the uORF divided by the mRNA density of the downstream mORF. To calculate the changes in  $TE_{uORF}$  ( $\Delta TE_{uORF}$ ) at either 20 °C or 37 °C as compared to 30 °C, we performed DESeq2 analysis. As described in the "Materials and methods" section, DESeq2 is a statistical package using the framework of a generalized linear model (GLM) that can identify changes in RPF and mRNA densities, as well as TEs, for each ORF between two conditions; place confidence intervals on the magnitude of changes; and exclude genes with less than a minimum number of read counts or with aberrantly high variability. Applying this analysis to our data revealed greater changes in uORF RPF densities versus the corresponding mORF mRNA densities for cells grown at 20 °C versus 30 °C (Fig. 4a), and also for cells grown at 37 °C versus 30 °C (Fig. 4b), as indicated by the greater spread in the uORF RPF density versus mORF RNA density scatterplots [(ii) versus (i)]. These findings suggest that the changes in growth temperature led to more extensive changes in translation of uORFs than transcription and/or stability of the mRNAs containing the uORFs.

To identify uORFs showing changes in TE that appear to be activated or repressed by a change in growth temperature, we applied two criteria. First, we considered only those uORFs showing an increase or decrease in  $TE_{uORF}$  of  $\geq 2$ -fold at a given temperature with respect to 30 °C using a false discovery rate (FDR) of  $\leq 0.1$ . Thirty-nine uORFs showed significant TE changes at 20 °C versus 30 °C; whereas 84 uORFs displayed such TE changes at 37 °C versus 30 °C (Additional file 1: Figure S4A, S4B). We reasoned that changes in  $TE_{uORF}$  could result because of multiple mechanisms. For example, translation initiation on the mRNA as a whole could increase or decrease because of changes in the efficiency of PIC attachment or scanning processivity, leading to





corresponding increases or decreases in the TEs of both the uORF(s) and mORF. To exclude such changes in TE<sub>uORF</sub> occurring concurrently with similar changes in TE<sub>mORF</sub>, we devised a term called “relative-TE<sub>uORF</sub>” which is TE<sub>uORF</sub>/TE<sub>mORF</sub> at any given temperature. Calculating changes in relative-TE<sub>uORF</sub> ( $\Delta$ relative-TE<sub>uORF</sub>) helped to identify changes in uORF translation not occurring simultaneously with similar changes in the TE of the mORF. Thus, according to our second criterion, translation of a uORF was called *regulated* if there was  $\geq 2$ -fold change (increase or decrease) in relative-TE<sub>uORF</sub> at a given temperature with respect to 30 °C; that is,

TE<sub>uORF</sub> changed  $\geq 2$ -fold more than TE<sub>mORF</sub>, or their changes were in opposite directions. Applying these criteria, we identified uORFs whose translation is specifically regulated by changes in growth temperature. We classified uORF translation as *activated* if both  $\Delta$ TE<sub>uORF</sub> and  $\Delta$ relative-TE<sub>uORF</sub> are  $\geq 2$  and as *repressed* if both  $\Delta$ TE<sub>uORF</sub> and  $\Delta$ relative-TE<sub>uORF</sub> are  $\leq 0.5$ .

After applying these criteria, we found 36 uORFs showing temperature-dependent translational regulation at 20 °C. There were 16 uORFs whose translation was significantly activated (Fig. 4a, red circles in (iii)) of which 2 were AUG uORFs (open circles) and 14 were

NCC uORFs (solid circles). There were 20 uORFs whose translation was significantly repressed at 20 °C (Fig. 4a, blue circles in (iii)) as compared to 30 °C, of which 5 were AUG uORFs (open circles) and 15 were NCC uORFs (solid circles). We also found 76 uORFs showing temperature-dependent translational regulation at 37 °C. There were 24 uORFs whose translation was significantly activated (Fig. 4b, red circles in (iii)) as compared to 30 °C, of which 7 were AUG uORFs (open circles) and 17 were NCC uORFs (solid circles) and 52 uORFs whose translation was significantly repressed at 37 °C (Fig. 4b, blue circles in the right panel) all of which are NCC uORFs. The changes in translation efficiency of uORFs ( $\Delta TE_{uORF}$ ) were driven by the changes in ribosome density ( $\Delta RPF$ -density) and not by changes in mRNA levels ( $\Delta mRNA$ -density) for the uORFs regulated at either 20 °C or 37 °C (Additional file 1: Figure S4C), as well as for all 1359 translated uORFs identified in this study (Additional file 1: Figure S4D), as revealed by the high Spearman correlation coefficient values between  $\Delta TE_{uORF}$  and  $\Delta RPF$ -density and low coefficient values between  $\Delta TE_{uORF}$  and  $\Delta mRNA$ -density.

It has been reported that alternative transcription start sites can produce mRNA isoforms in yeast with different translational efficiencies [39]. Thus, we also calculated Spearman correlation coefficients between  $\Delta TE_{uORF}$  and the changes in reads of just the 5'-UTRs of these sets of mRNAs ( $\Delta mRNA_{5'-UTR}$ -density) and found that they are much smaller than those for  $\Delta RPF$ -density and are not statistically significant (Additional file 1: Figure S4C and S4D, orange bars). This result suggests that the changes in  $TE_{uORF}$  observed are not due to temperature-dependent alterations in transcriptional start sites that produce different levels of mRNA isoforms including or excluding the uORFs in question. Furthermore, high Spearman correlation coefficient values between  $\Delta TE_{uORF}$  and  $\Delta RPF$ -density and low coefficient values between  $\Delta TE_{uORF}$  and both  $\Delta mRNA$ -density and  $mRNA_{5'-UTR}$ -density were observed for all translated AUG and NCC uORFs (Additional file 1: Figure S4E and S4F), indicating that for both, these sets of uORFs the changes in translational efficiency were driven by changes in RPF density and not in mRNA or 5'-UTR density.

The F-Luc<sup>UUG</sup> and *HIS4*<sup>UUG</sup>-*LacZ* reporters showed decreased translation at 20 °C and increased translation at 37 °C (Figs. 1b, 2b, Additional file 1: Figure S1B) suggestive of altered efficiency of use of the NCCs. In contrast to these reporters, we found changes in growth temperature lead to a more diverse transcriptome-wide response of translation of uORFs starting with not only NCCs but also AUGs. As described above, we identified 112 uORFs (14 AUG uORFs and 98 NCC uORFs) (Additional file 2: Table S3) on 84 different mRNAs whose translation was activated or repressed in response to

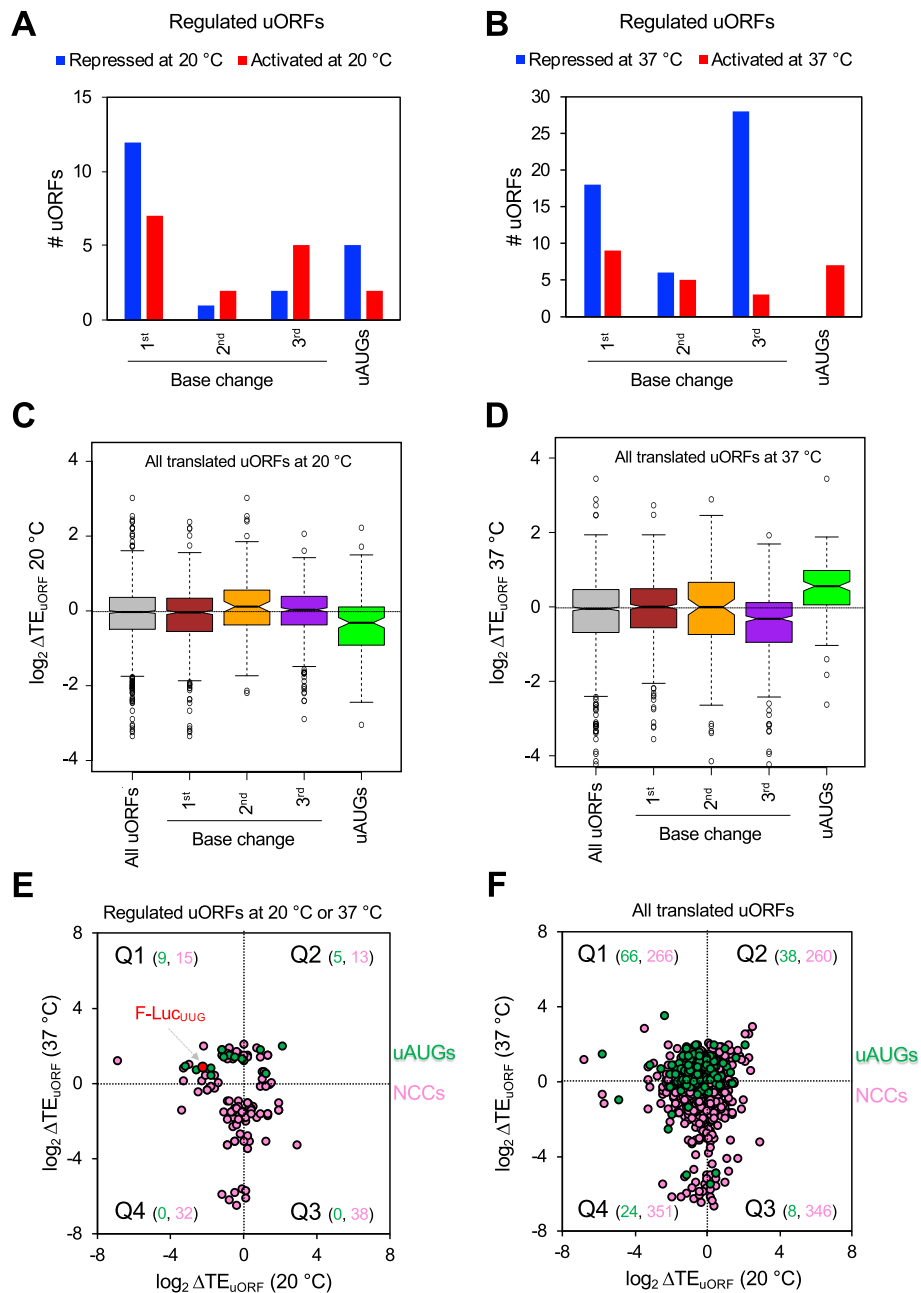
changes in growth temperature. It is noteworthy that less than 10% of the set of 1359 translated uORFs have significantly altered translation relative to changes in mORF translation at reduced or elevated growth temperatures (20 or 37 °C), indicating that temperature changes in this range do not produce global effects on the initiation efficiency of uORFs, but rather have specific, mRNA-dependent effects.

#### Influence of uORF start codon sequence on temperature-dependent changes in translation

The distribution of start codons of the 112 regulated uORFs described above is shown in Fig. 5a and b. We separated the uORFs based on their start sites into four classes: uAUGs; NCCs with the first base changes with respect to AUG (UUG, CUG, GUG), second base changes (AAG, ACG, AGG), and third base changes (AUC, AUA, AUU). The number of uORFs in each bin is too small to make inferences about the statistical significance of the differences. Thus, we next analyzed the changes in TE ( $\Delta TE_{uORF}$ ) of all 1359 translated uORFs by binning them into groups based on their start codon triplets without applying the criteria used to identify significant changes (Fig. 5c, d). The black horizontal dotted line indicates the median  $\Delta TE_{uORF}$  for all uORFs analyzed in this study at each temperature, which is close to unity. The NCC uORFs did not show a significant difference in median  $\Delta TE_{uORF}$  at 20 °C versus 30 °C when compared to all uORFs (Fig. 5c). In contrast, the AUG uORFs showed a significantly lower median  $\Delta TE_{uORF}$  at 20 °C when compared to all uORFs, suggesting that the use of uAUGs tends to be decreased at 20 °C.

We performed a similar analysis with uORF translation at 37 °C (Fig. 5d). The TE of uORFs starting with NCCs with the third base changes had a significant tendency to be downregulated at 37 °C, whereas those with the first and second base changes showed no clear trend. AUG uORFs displayed a strong overall increase in  $TE_{uORF}$  (positive  $\Delta TE_{uORF}$ ) at 37 °C. Thus, the use of uAUGs is significantly altered at both temperatures: reduced at 20 °C and elevated at 37 °C, similar to the behavior of the F-Luc and *HIS4*-*LacZ* reporters with UUG start codons. In contrast to AUG uORFs, the translation of NCC uORFs displays no clear trends in response to changes in growth temperature, save for reduced utilization of NCCs with the third base changes at 37 °C.

To better understand the global translational response of uORFs to the changes in growth temperature, we plotted  $\Delta TE_{uORF}$  at 20 °C versus  $\Delta TE_{uORF}$  at 37 °C (both with respect to 30 °C). First, we examined the 112 uORFs showing temperature-dependent translational regulation at either 20 °C or 37 °C (Fig. 5e). The distribution of these uORFs on the scatterplot reflects their translational behavior at each growth temperature



**Fig. 5** Varied response to temperature of uORF translation. **a, b** Responses of uORF translation to growth temperature grouped by start codons triplets. The temperature-regulated uORFs [see Fig. 4a, b; highlighted in red or blue in (iii)] were binned according to their start site. **a** Analysis done for changes observed at 20 °C relative to 30 °C. **b** Analysis done for changes observed at 37 °C relative to 30 °C. **c, d** Notched boxplot analysis for changes in  $TE_{uORF}$  ( $\Delta TE_{uORF}$ ) on all the uORFs starting with each near-cognate start codon triplet. *All uORFs* indicates all translated uORFs identified and analyzed in this study ( $N = 1359$ ); *uAUGs* ( $N = 136$ ); *first base change* ( $N = 631$ ); *second base change* ( $N = 109$ ); and *third base change* ( $N = 483$ ). The dotted horizontal line shows the median for *All uORFs*. **c** Analysis done for  $\Delta TE_{uORF}$  at 20 °C. **d** Analysis done for  $\Delta TE_{uORF}$  at 37 °C. **e, f** Scatterplots of  $\Delta TE_{uORF}$  at 20 °C (x-axes) versus  $\Delta TE_{uORF}$  at 37 °C (y-axes). **e** The  $\Delta TE_{uORF}$  values of temperature-regulated uORFs [Fig. 4a, b, (iii)]. Green circles represent AUG uORFs (uAUGs,  $N = 14$ ), and pink circles represent NCC uORFs (NCCs,  $N = 98$ ). The plot is broken into four quadrants, Q1–4, depending on the direction of  $\Delta TE_{uORF}$  at each temperature. The numbers in parenthesis represent the number of uORFs present in that specific quadrant. The red circle shows the main ORF of F-Luc<sup>uUG</sup> reporter mRNA. **f** Same as in **e**, except all translated uORFs identified in this study whose changes in TE could be determined are shown. AUG uORFs ( $N = 136$ , green circles) and NCC uORFs ( $N = 1223$ , pink circles)

relative to 30 °C: repression at 20 °C and activation at 37 °C (quadrant 1), activation at both temperatures (quadrant 2), repression at 37 °C and activation at 20 °C (quadrant 3), and repression at both temperatures (quadrant 4). The position of the F-Luc<sup>UUG</sup> reporter in quadrant 1 (red circle) indicates translational regulation at altered temperatures, with decreased expression at 20 °C and increased expression at 37 °C. Similarly, the majority (9/14) of the AUG uORFs (Fig. 5e, green circles) are positioned in quadrant 1, indicating that translation of regulated uORFs starting with AUGs tends to be decreased at 20 °C ( $\log_2 TE_{uORF} < 0$ ) and increased at 37 °C ( $\log_2 TE_{uORF} > 0$ ). All of the AUG uORFs that met our criterion for regulated changes in translation are in quadrant 1 or 2, indicating that translation of AUG uORFs is generally increased at 37 °C relative to 30 °C. The regulated NCC uORFs (Fig. 5e, pink circles), on the other hand, showed a scattered distribution across all the quadrants indicating that changes in translation of these NCC uORFs are more variable.

We performed a similar analysis with all translated uORFs ( $N = 1359$ ) (Fig. 5f). The overall distribution on this scatter plot is similar to that for the regulated uORFs in Fig. 5e, except that many uORFs whose translation is unaffected by temperature are present near the middle of the plot. The significant numbers of points representing NCC uORFs along the vertical axis between quadrants 3 and 4 in both plots (Fig. 5e, f, pink circles) indicates that many of these uORFs are repressed at 37 °C relative to 30 °C, but are relatively unaffected at 20 °C, consistent with the box plot analyses of these same uORFs (Fig. 5c, d). As with the AUG uORFs that met the criteria for regulated changes (Fig. 5e), 66/136 of all translated AUG-uORFs are repressed at 20 °C and activated at 37 °C (Fig. 5f, green circles; quadrant 1) and most (104/136) are activated at 37 °C (Fig. 5f, green circles; quadrants 1 and 2), indicating that this is a general phenomenon. In contrast, NCC uORFs have less coherent behavior upon change in growth temperature and can be unaffected, activated, or repressed (Fig. 5f, pink circles) in an apparently mRNA-specific manner.

Recently, 982 uORFs were identified from *S. cerevisiae* in 791 mRNAs using a comparative genomics approach to identify translated uORFs that are conserved in length or sequence among yeast species [10]. Approximately 44% of these are AUG uORFs and ~31% are UUG uORFs. When we interrogated this conserved uORF set, we found that, similar to our observations with the translated uORFs described above, translation of the conserved AUG uORFs is significantly repressed at 20 °C and activated at 37 °C (Additional file 1: Figure S5A-C, green boxplots and circles). Intriguingly, the TE of the conserved NCC uORFs is on average slightly elevated at 20 °C and slightly repressed at 37 °C (Additional file 1:

Figure S5A-C, pink boxplots and circles). Thus, a set of conserved AUG uORFs identified in a different manner than was our set of translated AUG uORFs displays a similar overall response to temperature.

#### Translated AUG uORFs tend to be in shorter, less structured 5'-UTRs than do NCC uORFs

To look for possible mechanisms influencing translation of NCC and AUG uORFs, we investigated whether intrinsic properties of these uORFs or their mRNAs display any significant correlations with uORF translational efficiencies. We first used a dataset of ~2700 yeast mRNA 5'-UTR lengths and propensities of forming secondary structures [40] to look for trends in the translated uORFs. We found that the 5'-UTRs of mRNAs with translated uORFs are on average longer and have a higher propensity to form secondary structure than the genomic average (Additional file 1: Figure S6A-C; see Additional file 4 for details). This result might be expected because shorter 5'-UTRs have less space in which to have a uORF and fewer possibilities for base pairing. Translated AUG uORFs tend to be significantly closer to the 5'-cap (Additional file 1: Figure S6D) and shorter (Additional file 1: Figure S6F) than are all translated uORFs or NCC uORFs. AUG uORFs also tend to be on shorter 5'-UTRs (Additional file 1: Figure S6G) and have less overall structure in their 5'-UTRs and in their mORFs near the start codon (Additional file 1: Figures S6H, I, K, L, M) compared to the translated NCC uORFs (green versus pink boxplots) or all translated uORFs (green versus gray boxplots). No statistical difference is seen between AUG and NCC uORFs or all uORFs when distance from the mAUG is compared (Additional file 1: Figure S6E).

Next, we calculated the “context adaptation scores” for the uORFs, as described previously [19, 41], quantifying the similarity between the start codon context of each uORF to that of the mORF AUGs of the 2% of yeast mRNAs with the highest ribosomal loads [42]. The start codons of the AUG uORFs have a significantly lower context score when compared to NCC uORFs (Additional file 1: Figure S6N, green versus pink boxplots) or all translated uORFs (green versus gray boxplots), consistent with the notion that strong AUG codons in the 5'-UTR have likely been selected against evolutionarily [9]. Overall, these data indicate that AUG uORFs, on average, occur on shorter, less structured 5'-UTRs closer to the cap, and exhibit poorer context, compared to NCC uORFs.

#### Position of a uORF within a 5'-UTR influences its translational response to altered growth temperatures

We next investigated whether any of these intrinsic properties of the translated uORFs and their mRNAs

influence the temperature-dependent regulation of uORF translation by calculating the Spearman's correlation coefficients between  $\Delta TE_{uORF}$  and the features analyzed in Additional file 1: Figures S6D-N. We found a modest but significant positive correlation between the distance from the 5'-cap and  $\Delta TE_{uORF}$  at 20 °C versus 30 °C for all uORFs, AUG uORFs, and NCC uORFs (Fig. 6, top panel). We also found a significant positive correlation between the length of the 5'-UTR and  $\Delta TE_{uORF}$  at 20 °C for all NCC uORFs and all uORFs. Similar correlation was observed for AUG uORFs, but they did not meet statistical significance because of the smaller number of these uORFs. At 37 °C, the distance between the uORF start codon and the mAUG, the length of the uORF, and the length of the 5'-UTR all had a significant positive correlation with  $\Delta TE_{uORF}$  for all uORFs and NCC uORFs (Fig. 6, bottom panel). Again, similar correlations were observed for these same parameters for AUG uORFs, but they did not meet statistical significance because of the smaller number of these uORFs. Taken together, these data suggest that the position of a uORF in the 5'-UTR and the length of the 5'-UTR can influence how translation of the uORF responds to changes in growth temperature.

In an effort to confirm these relationships between TE changes and uORF position relative to the cap or mORF, we examined the groups of NCC uORFs exhibiting the greatest TE changes at either high or low temperature. Sorting all ~1200 translated NCC uORFs according to  $\Delta TE_{uORF}$  values at 20 °C or 37 °C revealed that TE changes at both temperatures vary over an ~1000-fold range, from 8-fold to ~0.01-fold. We then selected the 100 uORFs with the largest increases in TE (TE<sub>up</sub>) or the largest decreases in TE (TE<sub>down</sub>) for subsequent analysis (Fig. 7a, braces and dotted boxes). Boxplots of the TE values of TE<sub>up</sub> and TE<sub>down</sub> uORFs confirm that the median TE values of these groups of uORFs differ significantly between 30 and 20 °C (Fig. 7b) or between 30 and 37 °C (Fig. 7c).

As shown in Fig. 7d, the group of TE<sub>down</sub> uORFs at 20 °C (col. 2) are located significantly closer to the 5'-cap relative to all NCC uORFs. In contrast, the TE<sub>up</sub> uORFs at 20 °C (col. 3) are located significantly farther from the 5'-cap. The median distances between the 5'-cap and uORF start site for TE<sub>down</sub> (20 °C), TE<sub>up</sub> (20 °C), and all NCC uORFs are 38, 79, and 63 nt, respectively. These data are consistent with the correlation analysis performed on all translated uORFs (Fig. 6, col. 1, upper plot, pink) in suggesting that NCC uORFs located farther from the cap tend to exhibit increased translation at 20 °C, whereas those closer to the cap tend to exhibit the opposite trend. In contrast, there is no significant difference in median distance from the cap for the two groups of NCC uORFs classified as TE<sub>down</sub>

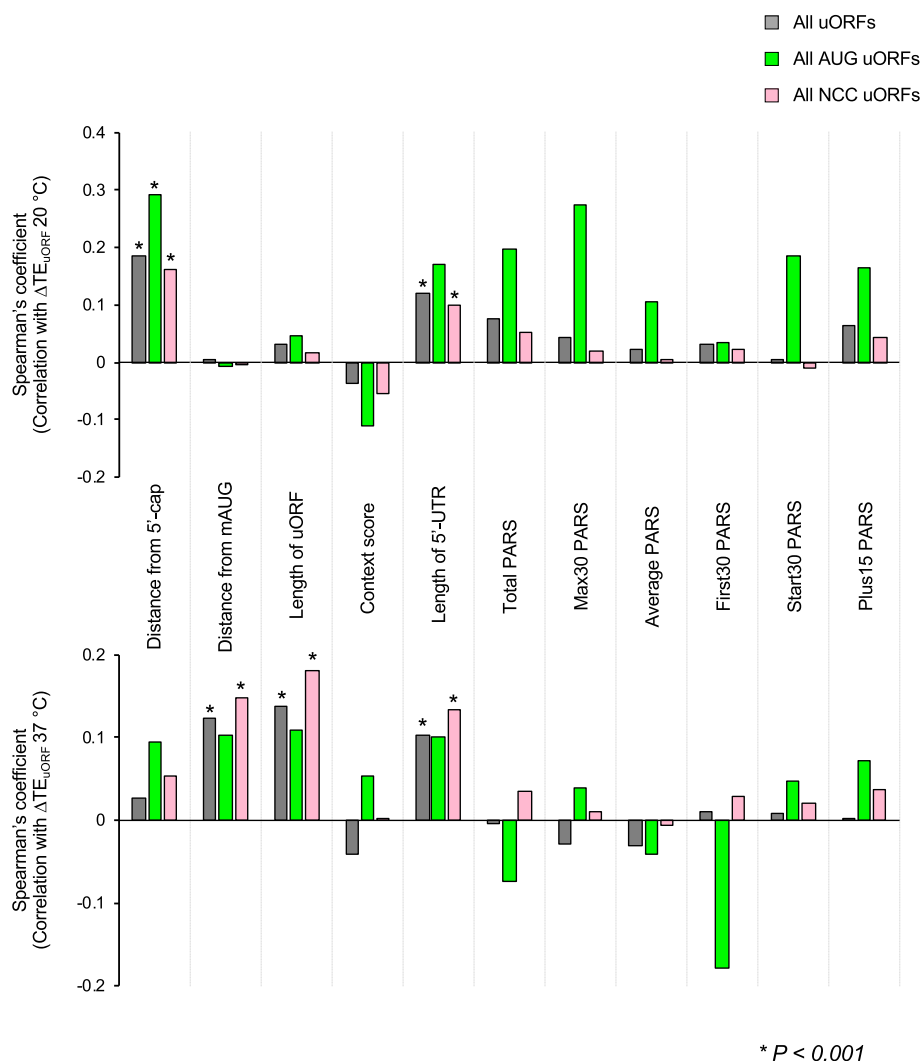
and TE<sub>up</sub> at 37 °C (Fig. 7d, cols. 4–5). These data also are consistent with the correlation analysis performed for all translated NCC uORFs at 37 °C (Fig. 6, col. 1, lower plot, pink).

The TE<sub>down</sub> (37 °C) NCC uORFs are located closer to the mAUG (median 51 nt) compared to all NCC uORFs (median 68 nt) and the TE<sub>up</sub> (37 °C) group of NCC uORFs (median 71 nt) (Fig. 7e, columns 4 versus 1 and 5). Although the TE<sub>down</sub> (20 °C) NCC uORFs had a significantly shorter median distance from the mAUG compared to all NCC uORFs (median 55 and 68 nt, respectively), the two groups of NCC uORFs with TE<sub>up</sub> or TE<sub>down</sub> at 20 °C do not differ significantly from one another in this parameter (median 55 and 56 nt, respectively; Fig. 7e, compare columns 1–3). These results are consistent with the correlation analysis for all translated NCC uORFs (Fig. 6, col. 3, upper and lower plots, pink) in suggesting that proximity to the mORF is associated with reduced translation of NCC uORFs at 37 °C but not at 20 °C.

We did not observe any significant influences of start codon context, the uORF length, or propensity to form secondary structure, on the temperature-dependent changes in translational efficiency for the TE<sub>up</sub> or TE<sub>down</sub> NCC uORFs at either temperature (Additional file 1: Figures S7A, S7C, and S8A-F). The mRNAs with TE<sub>down</sub> uORFs tend to have shorter 5'-UTRs than do the mRNAs with TE<sub>up</sub> uORFs (Additional file 1: Figure S7B).

#### Disparate effects of changes in uORF translation on mORF translation

We next investigated the effects of the temperature-dependent changes of uORF translation on the translational efficiencies of their downstream mORFs. Scatterplots of mORF TEs at either 20 versus 30 °C or 37 versus 30 °C revealed the absence of widespread TE changes as a function of growth temperature (Additional file 1: Figures S9A and S9B); only 25 mRNAs exhibit  $\geq 2$ -fold changes (activation or repression) in mORF TE (FDR < 0.1) at either 20 or 37 °C. To visualize the effects of changes in uORF translation on mORF expression, we plotted  $\Delta TE_{uORF}$  versus  $\Delta TE_{mORF}$  at 20 or 37 °C for our set of uORFs (described in Additional file 1: Figures S4A and S4B) showing significant changes in translation at either 20 or 37 °C (Fig. 8a, b). Changes in uORF translation had varying effects on the translation of the mORF, which can be categorized according to the quadrant in which the uORF/mORF pair falls in the scatterplots in Fig. 8. Cases in which a decrease in  $TE_{uORF}$  is accompanied by an increase in  $TE_{mORF}$  (quadrant 1) or an increase in  $TE_{uORF}$  is associated with a decrease in  $TE_{mORF}$  (quadrant 3) suggest that the uORF plays a canonical inhibitory role in terms of its effect on translation of the mORF. Quadrants 2 and 4 represent cases in which uORF and



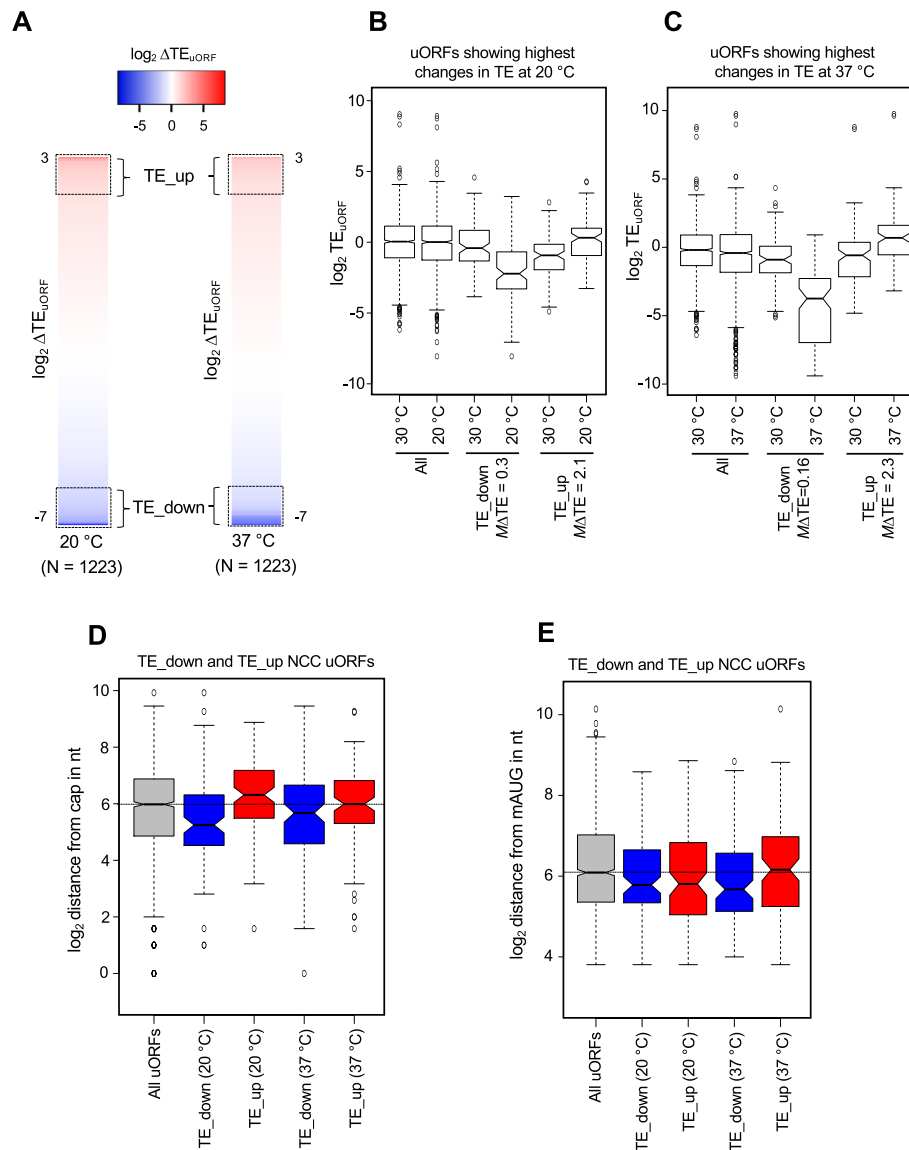
**Fig. 6** Position of a uORF within a 5'-UTR influences its translational response to altered growth temperatures. Spearman's correlation coefficients between  $\Delta TE_{uORF}$  at 20 °C (top panel) or 37 °C (bottom panel) and the parameters described in Additional file 1: Figure S6. The uORFs whose  $\Delta TE$  could be calculated were analyzed here. Coefficients with significance levels of  $p < 0.001$  are indicated with asterisks

mORF translation both increase or decrease, respectively. The simplest explanation for this behavior is that overall initiation on these mRNAs (e.g., PIC loading onto the 5'-UTR) increases or decreases, leading by mass action to increases or decreases in translation of both the uORF and the mORF. Alternatively, it is possible that some cases in these two quadrants represent mRNAs on which re-initiation after translation of the uORF is very efficient and thus an increase or decrease in uORF translation has a corresponding effect on mORF translation.

We performed a similar analysis using the  $TE_{up}$  and  $TE_{down}$  NCC uORF sets described above and in Fig. 7a–c (Fig. 8c, d). As with the set of regulated NCC uORFs, this set was also distributed into all four quadrants, with a preponderance in quadrants 1 and 2.

We also performed this analysis for all translated AUG uORFs (Fig. 8e, f). Again, the plot shows a distribution of mRNAs into all four quadrants. However, consistent with the behavior of AUG uORFs described above, a majority (66%) are in quadrants 1 and 4 (i.e., negative  $\Delta TE_{uORF}$ ) at 20 °C, whereas at 37 °C a majority (76%) are in quadrants 2 and 3 (i.e., positive  $\Delta TE_{uORF}$ ).

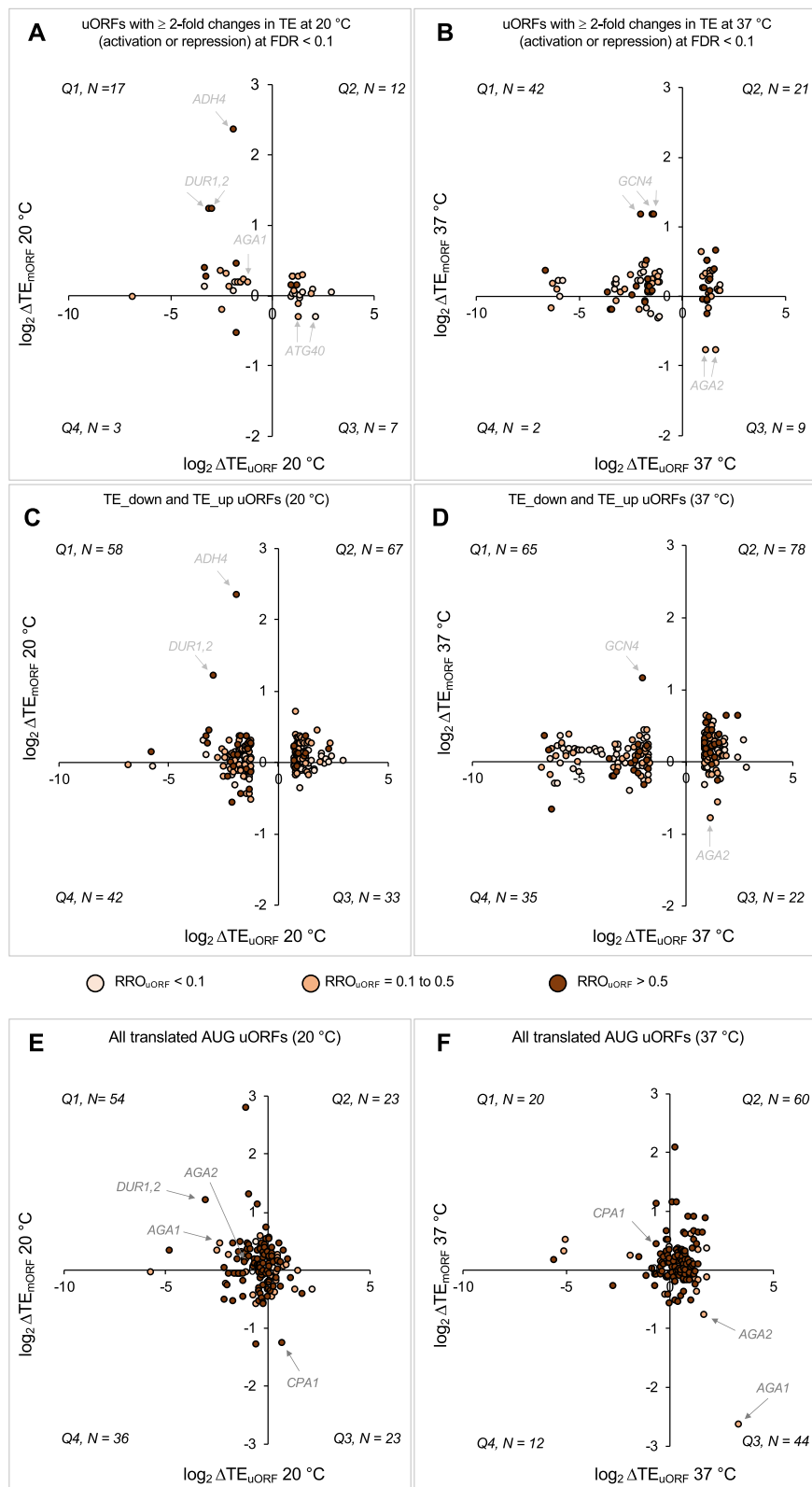
Why  $TE_{mORF}$  increases for the mRNAs described in Fig. 8a–f more often than it decreases regardless of the direction of  $\Delta TE_{uORF}$  is unclear, although it is possible that mRNAs whose main ORF translation decreases (negative  $\Delta TE_{mORF}$ ) are more likely to be degraded due to the coupling between active translation and mRNA stability [43, 44] and thus less likely to appear in the ribo- or RNA-seq data.



**Fig. 7** Identification and analysis of NCC uORFs with the greatest changes in TE at 20°C and 37°C. **a** Heatmaps of  $\Delta TE_{uORF}$  values for all translated NCC uORFs ( $N = 1223$ ), ranked according to  $\Delta TE$ . The scale on the top indicates the range of  $\Delta TE$  values. The dotted boxes indicate the 100 uORFs with the greatest decreases in  $TE_{uORF}$  (TE\_down) or the greatest increases in  $TE_{uORF}$  (TE\_up) at either 20°C or 37°C with respect to 30°C. **b** Boxplot analysis of the 100 uORFs with the greatest decrease (TE\_down) or increase (TE\_up) in  $TE_{uORF}$  at 20°C, identified as described in **a**. The median  $\Delta TE_{uORF}$  values for each group are shown below the plot (m $\Delta TE$ ). All indicates all translated NCC uORFs ( $N = 1223$ ). **c** Same as in **b** but for  $\Delta TE_{uORF}$  at 37°C. **d** Boxplot analysis of the distance between the start site of the uORF and the 5'-end of the mRNA. The dotted horizontal line indicates the median distance of All translated NCC uORF start codons from the 5'-ends of the mRNAs. **e** Boxplot analysis of the distance between uORF start sites and the downstream mORF start (mAUG) codons. The dotted horizontal line indicates the median distance of All translated NCC uORF start codons from the mORF start codons

In order to assess possible trends in the relative translation of the uORFs and mORFs, we colored the circles corresponding to each uORF in the plots in Fig. 8 according to their relative ribosome occupancies ( $RRO_{uORF}$ ) values at 30°C: tan for an  $RRO < 0.1$ , light brown for an  $RRO$  between 0.1 and 0.5, and dark brown for an  $RRO > 0.5$ . In general, no obvious

trends in RROs emerge for mRNAs with temperature-regulated uORFs. However, the plots for AUG uORFs (Fig. 8e, f) make clear that most AUG uORFs are well-translated relative to their downstream uORFs (dark brown circles,  $RRO > 0.5$ ), in contrast to the situation with NCC uORFs (Fig. 8a–d), where a wide range of RROs is observed and a majority are  $\leq 0.5$ .



**Fig. 8** (See legend on next page.)



(See figure on previous page.)

**Fig. 8** Correlations between changes in uORF and mORF translation suggest possible novel cases of uORF-mediated translational regulation. **a–f** Correlation analysis between  $\Delta TE_{uORF}$  and  $\Delta TE_{mORF}$  at multiple growth temperatures. Each circle in the plots represents a uORF. The circles are color-coded according to the relative ribosome occupancy of the uORF ( $RRO_{uORF}$ ; color key is shown at the bottom) which is the ribosome occupancy on the uORF ( $RO_{uORF}$ ) normalized to the ribosome occupancy on the mORF ( $RO_{mORF}$ ).  $RO_{uORF}$  is the RPF density on the uORF at 30 °C normalized to its length, and  $RO_{mORF}$  is the RPF density on the mORF at 30 °C normalized to its length. The plot is divided into four quadrants (Q1–Q4). The number of uORFs in each quadrant is shown on every plot. Multiple sets of uORFs are used in this analysis. **a, b** The uORFs with  $\geq 2$ -fold changes in TE (activation or repression) at FDR < 0.1 were analyzed (highlighted circles in red and blue in Additional file 1: Figures S4A and S4B). **a** Analysis of changes observed at 20 °C. **b** Analysis of changes observed at 37 °C. **c, d** NCC uORFs with the greatest changes in TE ( $TE_{down}$  and  $TE_{up}$ ) at 20 °C (**c**) and 37 °C (**d**) as described in Fig. 7a were analyzed. **e, f** All translated AUG uORFs ( $N = 136$ ) were analyzed for changes in translation at 20 °C (**e**) and 37 °C (**f**). For **a–f**, cases representing the uORFs with a possible canonical inhibitory function for the uORFs are indicated with an arrow

### uORF-dependent regulation of translation by changes in growth temperature

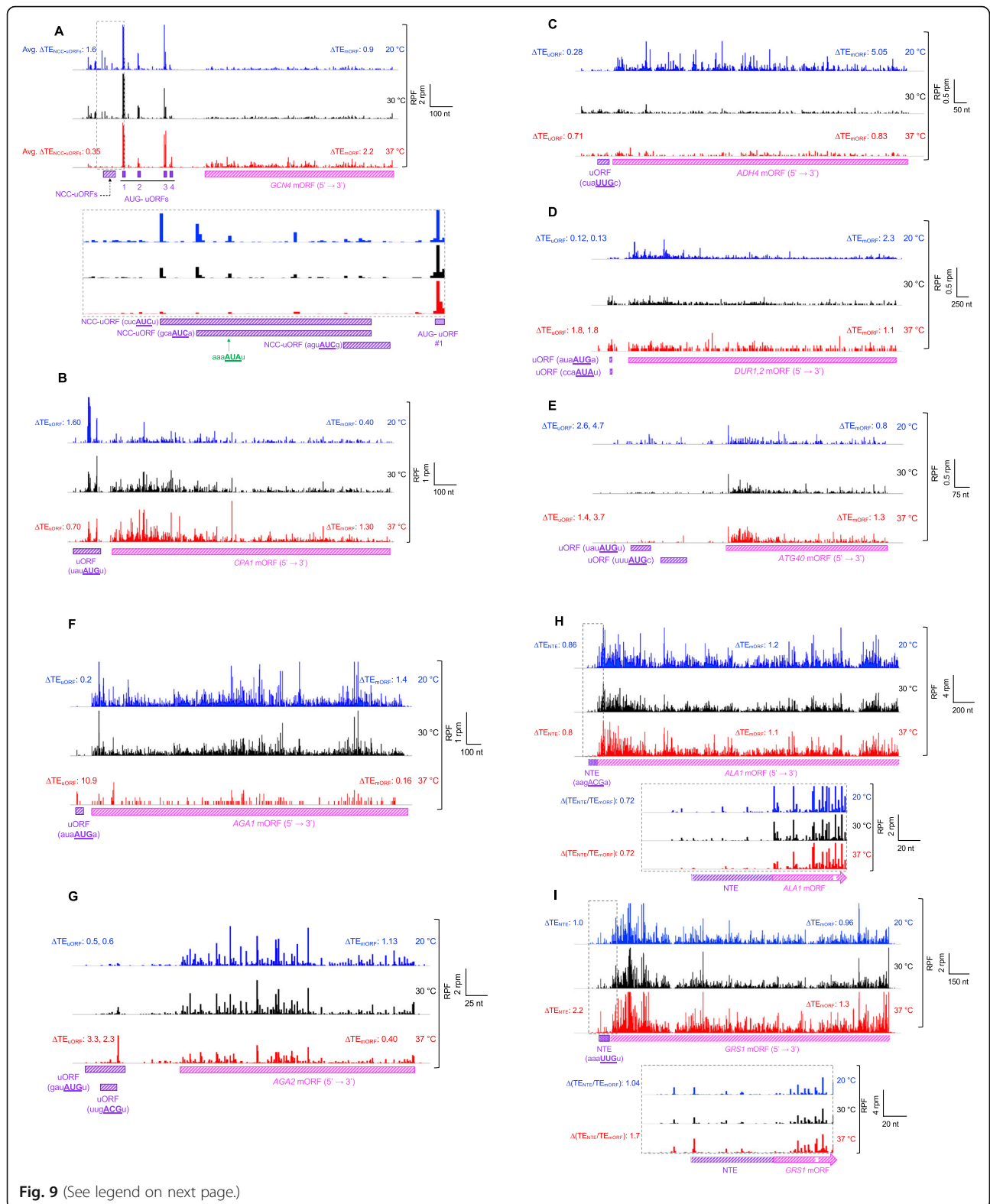
Figure 8 highlights a number of clear cases in which uORF TE decreases and the TE of the downstream mORF correspondingly increases (labeled circles in quadrant 1) or the uORF TE increases and the TE of the downstream mORF decreases (labeled circles in quadrant 3). One very well-established example of uORF-dependent translational regulation is the *GCN4* mRNA in which four AUG uORFs are involved in controlling nutrient-dependent modulation of the translation of the mORF [45]. Recently, ribosome profiling studies indicated that there was also a translated uORF starting with an NCC (AUA) upstream of the canonical AUG uORFs [46] and that its TE was increased upon amino acid starvation in yeast [8], and under sustained histidine limitation [46]. However, evidence was presented that increased or decreased translation of this NCC uORF is not associated with a significant change in *GCN4* expression under non-starvation, starvation, or stress conditions at 30 °C [46].

Intriguingly, the *GCN4* NCC uORFs have the strongest reciprocal increases in mORF translation when their translation is repressed at 37 °C of any NCC uORFs (Fig. 8b, d, quadrant 1). We therefore examined the ribo and mRNA read density plots (wiggle tracks) from the 20, 30, and 37 °C data (Fig. 9a, Additional file 1: Figure S10A). At all three temperatures, very strong ribosome footprint density was observed for AUG uORFs 1 and 3 with clear but lower density at AUG uORFs 2 and 4. We also observed significant ribosome occupancy at several NCC uORFs upstream of the canonical AUG uORFs, including the AUA uORF reported previously [46] (Fig. 9a, left and bottom panel, Additional file 1: Figure S10A). At 20 °C, the TEs of these NCC uORFs were increased on average 1.6-fold relative to 30 °C, which had no detectable effect on the TE of the mORF (Fig. 9a, compare black and blue traces, effective increase in  $RRO_{uORF} \sim 2$ -fold). In contrast, at 37 °C, the TEs of the NCC uORFs decreased  $\sim 3$ -fold, which was accompanied by an  $\sim 2$ -fold increase in mORF translation (effective decrease in  $RRO_{uORF} \sim 6$ -fold), suggesting that the NCC uORFs may

be involved in regulating translation of the *GCN4* mORF at elevated temperatures. Further, we calculated the TEs of the AUG uORFs manually as our pipeline does not allow the validation of uORFs shorter than three codons. We found that these uORFs also show temperature-dependent changes in TEs. At 20 °C, the TE of uORF1 is decreased  $\sim 1.7$ -fold relative to 30 °C, while the TE of uORF3 is increased by  $\sim 1.5$ -fold. At 37 °C, the TE of uORF2 is decreased  $\sim 3$ -fold relative to 30 °C, while the TEs of uORFs 3 and 4 are increased  $\sim 1.7$ - and  $2.7$ -fold, respectively. The differing behavior of each *GCN4* uORF with respect to temperature seems to underscore the conclusion that no single variable is solely responsible for the observed effects of growth temperature on uORF translation.

Another well-studied example of uORF-mediated translational regulation is the *CPA1* mRNA, which encodes an arginine biosynthetic enzyme. At high arginine levels, translation of the AUG uORF represses translation of the downstream mORF [47–49]. Strikingly, in our data, the *CPA1* mRNA has the highest degree of negative regulation of mORF translation of any mRNA with an AUG uORF when the translation of the uORF increases at 20 °C relative to 30 °C (Fig. 8e, quadrant 3). At 20 °C,  $TE_{uORF}$  was elevated  $\sim 1.6$ -fold while  $TE_{mORF}$  was repressed  $\sim 2.5$ -fold (effective increase in  $RRO_{uORF} \sim 4$ -fold) (Fig. 9b, Additional file 1: Figure S10B, compare black and blue traces). Conversely, at 37 °C  $TE_{uORF}$  was repressed  $\sim 1.4$ -fold and  $TE_{mORF}$  was elevated by 1.3-fold (effective decrease in  $RRO_{uORF} \sim 2$ -fold) (Fig. 9b, compare black and red traces). These data suggest that *CPA1* translation may be regulated by temperature-dependent changes in uORF translation, either through a direct effect of temperature on uORF TE or indirectly, for example, due to temperature-dependent changes in arginine levels in the cell.

A number of other mRNAs that were not previously known to be subject to uORF-mediated translational control appear in our data as having reciprocal, temperature-dependent changes in uORF and mORF TE values suggestive of a regulatory relationship (Fig. 8). For example, *ADH4* mRNA, which encodes an alcohol



(See figure on previous page.)

**Fig. 9** mRNAs that show reciprocal changes in the translation of uORFs and mORFs at multiple temperatures. **a** Wiggle track images showing ribosome-protected fragments (RPF) on the *GCN4* mRNA in cells cultured at either 20, 30, or 37 °C, in units of rpm (reads per million mapped reads from two replicates at each temperature). The RPF tracks were normalized to the mRNA levels at each temperature to reflect the changes in translation efficiencies ( $\Delta TE$ ) of uORF and mORF as described in the legend to Fig. 2 and the “Materials and methods” section. The schematic shows the position of the uORFs (purple rectangles) and mORF (striped pink rectangle). NCC uORFs are shown with striped purple rectangle. AUG uORFs are in purple rectangles. Average change in the TEs of the three NCC uORFs showing significant changes in translation at 20 or 37 °C (Avg.  $\Delta TE_{NCC \text{ uORFs}}$ ) is shown. The enlargement of the boxed area is also shown below with start sites of NCC uORFs (bold, underlined) and the -3 to -1 and +4 context nucleotides. The green arrow shows the NCC uORF start site (AUA) that has been previously shown to be used as an upstream start site [46]. **b–g** Same as in **a** but for **b** the *CPA1* mRNA, **c** *ADH4* mRNA, **d** *Dur1,2* mRNA, **e** *ATG40* mRNA, **f** *AGA1* mRNA, **g** *AGA2* mRNA. **h, i** Wiggle track images of the *ALA1* (**h**) and *GRS1* (**i**) mRNAs as described in **a**. The N-terminal extension (striped purple rectangle) and mORF (striped pink rectangle) are shown. Relative  $TE_{NTE}$  ( $TE_{NTE}/TE_{mORF}$  ratio) reflects the ratio of translation efficiency of initiation at the start site of the NTE (ACG in the case of *ALA1* and UUG in the case of *GRS1*) to that of the combined initiation events at NTE and mAUG

dehydrogenase enzyme, undergoes a 5-fold increase in mORF TE at 20 °C while the TE of its NCC uORF decreases 3.6-fold (effective decrease in  $RRO_{uORF} \sim 18$ -fold) (Fig. 8a, c; Fig. 9c; Additional file 1: Figure S10C, compare black and blue traces). Likewise, the *DUR1,2* mRNA, which encodes a urea amidolyase enzyme, has one NCC uORF and one AUG uORF. The TEs of these uORFs decrease nearly 10-fold at 20 °C with a corresponding increase of 2.3-fold in the TE of the mORF (effective increase in  $RRO_{uORF} \sim 20$ -fold) (Fig. 8a, c; Fig. 9d; Additional file 1: Figure S10D, compare black to blue).

The *AGA1* and *AGA2* mRNAs, which encode the subunits of the a-agglutinin receptor, are particularly striking examples of possible uORF-dependent translational regulation (Fig. 8a, b, d–f, quadrants 1 and 3). Translation of the *AGA1* AUG uORF increases over 10-fold at 37 °C relative to 30 °C, which is accompanied by a 6-fold decrease in TE for the mORF (effective increase in  $RRO_{uORF} \sim 60$ -fold) (Fig. 9f and Additional file 1: Figure S10F, compare black and red traces). TE of the uORF is also decreased 5-fold at 20 °C, with a corresponding increase in mORF TE of 1.4-fold (effective decrease in  $RRO_{uORF} \sim 7$ -fold). The fact that the increase in  $TE_{mORF}$  is smaller than the decrease in  $TE_{uORF}$  can be explained by the fact that at 30 °C the uORF is translated at only a small percentage of the level of the main ORF, with an  $RRO$  value of 0.18. *AGA2* mRNA, which has two translated uORFs, displays similar behavior (Fig. 9g and Additional file 1: Figure S10G). Intriguingly, the constitutive agglutinability of yeast cells, the process mediated by a-agglutinin, has been reported to be diminished at growth above 30 °C [50], consistent with the putative uORF-mediated regulation of *AGA1* and *AGA2* expression suggested by our ribosome profiling data.

In addition to altering uORF translation, changes in the efficiency of initiation site usage can alter the levels of inclusion or exclusion of N-terminal extensions (NTEs) that are in-frame with the main ORF of an mRNA, changing the balance between different protein isoforms. We found 130 cases where initiation upstream of the annotated mAUG appears to lead to formation of

an NTE (Additional file 2: Table S4). Interestingly, > 95% of these upstream start sites are NCCs. Two well-characterized examples of N-terminal extensions that encode mitochondrial localization signals are the yeast alanyl- and glutamyl-tRNA synthetases encoded by the *ALA1* and *GRS1* mRNAs, respectively [5, 6]. Both of these extensions initiate with NCCs, ACG for *ALA1*, and UUG for *GRS1*. We see evidence of translation of both NTEs in our ribosome profiling data (Fig. 9h, i and Additional file 1: Figure S10H, I), with relative ribosome occupancies for the NTE versus mORF of 0.1 and 0.23, respectively, at 30 °C. Although the TE of the NTE of *ALA1* is only modestly affected by growth temperature (Fig. 9h and Additional file 1: Figure S10H) with an effective decrease in  $RRO_{uORF} \sim 1.4$ -fold at both 20 °C and 37 °C, the TE of the *GRS1* NTE increases over 2-fold at 37 °C relative to 30 °C (Fig. 9i and Additional file 1: Figure S10I, compare black and red traces, effective increase in  $RRO_{uORF} \sim 1.7$ -fold). This change could potentially lead to an increase in the mitochondrial concentration of the glutamyl-tRNA synthetase at elevated growth temperatures.

#### uORF-dependent regulation of expression of *AGA1* mRNA by changes in growth temperature

In order to help validate the roles of the newly identified uORFs in temperature-dependent regulation of mORF expression, we further investigated the regulation of *AGA1* mRNA translation. We cloned the *AGA1* coding region along with its native promoter with a C-terminal 3XHA tag on a high-copy plasmid [Additional file 1: Figure S11A, *AGA1-HA (WT)*] and measured the protein levels when cells were grown at 20, 30, and 37 °C. Western blotting of the resulting cell lysates using anti-HA tag antibodies showed that, relative to 30 °C, the levels of Aga1-HA protein increased  $\sim 2$ -fold at 20 °C and decreased  $\sim 100$ -fold at 37 °C (Additional file 1: Figure S11B; S11D, E, left panels for WT). These data are consistent with our ribosome profiling results ( $RPF_{mORF}$  at 20 °C is increased  $\sim 3$ -fold and  $RPF_{mORF}$  at 37 °C is decreased  $\sim 17$ -fold). To test the role of the uORF in

regulating mORF expression, we mutated the uORF start site (AUG) to a non-cognate (AAA) and assessed the levels of Aga1-HA protein. With this mutant reporter, Aga1-HA protein levels were identical to the WT construct at 30 °C and still increased ~2–3-fold at 20 °C, indicating that the uORF does not play a significant role in the change in expression at lower growth temperature (Additional file 1: Figure S11C, S11D; right panel for mutant). These data are consistent with the fact that  $TE_{\text{mORF}}$  only changes 1.4-fold between 30 and 20 °C (Fig. 9f), likely due to the small RRO between uORF and mORF at 30 °C, as noted above. In this case, the increased expression levels are likely driven by changes in transcription or mRNA stability, which can be seen in the increased *AGAI* mRNA abundance at 20 °C (Additional file 1: Figure S10F, compare green mRNA traces at 20 and 30 °C).

In contrast, changing the *AGAI* uORF start site to AAA [Additional file 1: Figure S11A, *AGAI-HA (Mutant)*] led to a significant loss of regulation at 37 °C, from ~100-fold in the case of WT to ~10-fold in the case of the mutant (Additional file 1: Figure S11C; S11E, compare black and red columns). These data are consistent with the interpretation that an increase in translation of the uORF at 37 °C results in a reciprocal decrease in the translation of the mORF. The remaining 10-fold temperature dependence in the expression of the mutant could be due to transcriptional regulation or changes in mRNA stability, consistent with the ~2.5-fold decrease in mRNA density at 37 °C relative to 30 °C observed in the ribosome profiling data (Additional file 1: Figure S10F, compare green mRNA traces at 30 and 37 °C). In addition, consistent with the proposed inhibitory role of uORF, loss of the uORF start codon led to an ~10-fold increase in Aga1-HA protein expression at 37 °C relative to WT (Additional file 1: Figure S11E, compare WT and mutant columns in red). Together, these results support the conclusion that the AUG-initiated uORF in the 5'-UTR of *AGAI* mRNA regulates the translation of the downstream mORF at 37 °C.

## Discussion

In this report, we provide evidence for growth temperature-induced changes in the efficiency of translation of a subset of uORFs in the 5'-UTRs of mRNAs in *S. cerevisiae*. Using a multi-filter pipeline, we identified 1367 uORFs with strong evidence of translation at 1 or more growth temperatures in our ribosome profiling datasets. Most (90%) of the translated uORFs begin with near-cognate codons (NCCs), over half of which are first position changes from AUG (UUG, GUG, CUG). The 10% that begin with AUG codons have higher TEs at the optimum growth temperature of 30 °C than those beginning with NCCs, as expected, and their downstream main ORFs have correspondingly lower TEs than the

mORFs downstream of NCC uORFs. Moreover, the ribosome occupancies of the uORFs relative to the mORFs (relative ribosome occupancy; RRO) were higher for the AUG uORFs than for the NCC uORFs (median value of 1.8 and 0.17, respectively). These findings suggest that translation of AUG uORFs (at 30 °C) frequently reduces the fraction of scanning ribosomes that reach the mORF start codon, consistent with the usual inhibitory effect of uORF translation (for review see [13]), whereas the majority of NCC uORFs are translated at levels too low to exert this regulatory function.

Of the 1367 translated uORFs in our dataset, we found ~10% exhibited changes (activation or repression) in their translation at 20 and/or 37 °C relative to 30 °C that met our dual criteria for temperature-regulated uORF translation (Fig. 4a, b (iii)). A majority of the regulated uORFs that begin with AUG codons have reduced translational efficiencies at 20 °C and increased efficiencies at 37 °C (Fig. 5a, b, e). These trends were also evident for the entire set of translated AUG uORFs (Fig. 5c, d, f). In contrast, regulated NCC uORFs do not display a consistent trend as their translation can be activated, repressed, or unaffected at either temperature. One possible explanation for this difference might be that the rate-limiting step for initiation on most AUG uORFs, which are generally well translated relative to the NCC uORFs (Fig. 3d) and reside on shorter, less structured 5'-UTRs (Additional file 1: Figure S6), has a temperature dependence such that its rate increases with growth temperature. NCC uORFs, which tend to be on longer, more structured 5'-UTRs (Additional file 1: Figure S6), might have different rate-limiting steps for initiation depending on a variety of factors (e.g., position in the 5'-UTR, structural features of the 5'-UTR, length of uORF), and these differing rate-limiting steps could result in differing temperature dependencies. It is noteworthy that most of the regulated NCC uORFs (70/98) are in Q3–Q4 (Fig. 5e), indicating repression of TE at 37 °C despite no clear trend at 20 °C. One possible explanation is that the imperfect codon-anticodon helices formed by NCCs are destabilized at high temperatures, whereas the perfect pairing with AUGs is stable enough to resist this effect.

The position of the uORF in the 5'-UTR exerts a significant influence on the direction and magnitude of the temperature dependence of translation of uORFs. Translation of uORFs that are closer to the 5'-cap than the average distance for all uORFs tends to be inhibited at 20 °C relative to 30 °C, whereas the translation of uORFs that are farther from the cap than the average tends to be activated at 20 °C (Figs. 6 and 7d). In addition, for NCC uORFs, the distance from the mORF AUG codon

also correlates with the temperature dependence of translation such that those farther from the mORF AUG are more likely to be activated at 37 °C (Figs. 6 and 7e).

One simple explanation for some of the observed effects could be that low temperature stabilizes the structure in 5'-UTRs, which is generally inhibitory towards uORF initiation, whereas higher temperature tends to destabilize the overall 5'-UTR structure. Such an effect could explain uORFs whose translation decreases at 20 °C and increases at 37 °C, including the majority of AUG uORFs. On the other hand, in particular cases, stabilization of mRNA structures at low temperature and destabilization at higher temperatures could have the opposite effects if the structures are located downstream from the uORF start codon and cause PICs to pause near the sub-optimal start codons, increasing the probability of initiation on them [51, 52]. This mechanism could account for the activation of uORF translation at 20 °C and its repression at 37 °C. If a structural element is already unstable at 30 °C, it could be the case that increasing the temperature to 37 °C has no significant effect, and if a structural element is already stable enough to produce a maximal effect at 30 °C, decreasing the temperature to 20 °C might produce no additional observable effect. Thus, it is possible to rationalize most of the classes of effects we see at decreased or elevated growth temperatures simply by invoking the influence of temperature on RNA structure.

Nonetheless, it is likely that the temperature-dependent effects on uORF translation we observe are not all the result of a single mechanism. Other possible influences might include the rate of scanning, as slower scanning could lead to increased initiation on suboptimal start codons, changes in the thermodynamics of codon:anticodon pairings, alterations of the levels or activities of specific mRNA binding proteins or other factors, or the temperature dependence of required structural rearrangements within the PIC or of enzymatic reactions such as GTP hydrolysis by eIF2 or eIF5B. Other factors such as the position of the uORF in the 5'-UTR, and hence the distance the PIC must scan to reach it, could further influence most of these possible mechanisms of regulation, leading to some of the observed correlations. Other possible explanations for the observed effects might include temperature dependence of the efficiency of tRNA modification or of the competitiveness of mRNAs for interactions with components of the translational machinery.

Although our data indicate that translation of a significant number of uORFs is regulated by growth temperature and suggest some cases in which these effects influence expression of the main ORF in the mRNA, it is also striking that this is not a general effect and that translation of most uORFs is relatively

insensitive to changes in temperature, at least between 20 and 37 °C. Only 8% of the translated uORFs we identified (112/1359) had changes in TE that met our criteria for temperature-regulated translation, showing altered uORF translation without a similar increase or decrease in the translation of the downstream mORF. Translation of the large majority of uORFs was refractory to temperature, which is remarkable in that temperature exerts effects on most reactions and interactions and has been shown to influence transcription, metabolism, and overall cellular physiology in yeast [53, 54]. One possible explanation for this seeming conundrum could be that yeast cells evolved homeostatic mechanisms to damp down the influence of growth temperature on general translation in order to maintain appropriate levels of protein products. Because read numbers and TEs in each ribosome profiling experiment reflect values relative to the average observed value for the parameter in that experiment, changes reported are also relative to the population averages. Thus, it is possible that the absolute rates of translation are changing as a function of temperature for most uORFs, but in a linear fashion such that we do not observe changes relative to the population averages in each ribosome profiling experiment. Those uORFs for which we observe significant changes in TE are ones in which translation is increasing or decreasing more than the average change in the experiment. Nonetheless, our data indicate that translation of most uORFs behaves the same with respect to temperature, which implies a general mechanism to prevent relative translation rates of different ORFs from diverging when the growth temperature shifts and thereby changing global proteomic ratios in suboptimal ways.

More studies will be required to understand the mechanistic basis of the temperature-dependent regulation of uORF translation described here and of the physiological consequences of these phenomena. The observed effects of temperature on the translation of N-terminal extensions also warrant additional analysis because protein isoforms with different N-termini can have altered cellular localization patterns, functions, or activities [5, 6, 55]. The set of temperature-dependent uORFs and N-terminal extensions we have identified using transcriptome-wide approaches should serve as a useful starting point for in-depth studies to elucidate the roles and underlying mechanisms of these intriguing systems.

## Materials and methods

### Yeast strains and plasmids

The yeast cells were transformed as described previously [56], and the transformants were selected on the appropriate media lacking the nutrients corresponding to auxotrophic marker/s. The detailed list of the strains and plasmids used in the study is provided in

Additional file 2: Tables S1 and S2, respectively. The *upf1Δ* strain has previously been described [24]. Plasmid sr. no. 1–8 are described in [4], plasmid sr. no. 9–10 are described in [21], and plasmid sr. no. 11 is described in [26].

### Biochemical assays

The dual-luciferase assay was carried out as previously described [4, 20] with some minor modifications. The wild-type/mutant yeast cells having either control reporter ( $R^{AUG}FF^{AUG}$ ) or test reporter ( $R^{AUG}FF^{XXX}$ ) (Fig. 1a) were grown overnight to saturation. The cells were then diluted to reach a desired  $OD_{600}$  of 0.6–0.8 in 16 h at the tested temperature (e.g., 20 °C, 30 °C). To calculate the luciferase activity, 2  $\mu$ l of culture was added to 50  $\mu$ l of 1 $\times$  Passive Lysis Buffer (Promega #E1941) which was aliquoted in a reader plate (Corning #CLS3912), followed by lysis at room temperature for 50 min. The luciferase activity was measured using a Turner Modulus Microplate Reader at 24 °C. Briefly, 50  $\mu$ l of F-Luc reagent (15 mM Tris pH 8.0, 25 mM glycylglycine, 4 mM EGTA, 15 mM  $MgSO_4$ , 1 mM DTT, 2 mM ATP, 0.1 mM CoA, and 75  $\mu$ M luciferin) was added to each well. The activity was measured with a delay time (the duration between the injection of the reagent and taking a measurement) of 2 s and an integration time (the duration of measurement per well) of 1 s. The R-Luc activity in the same well was immediately measured by adding 50  $\mu$ l of R-Luc reagent (0.22 M citric acid-sodium citrate pH 5, 1.1 M NaCl, 2.2 mM  $Na_2EDTA$ , 1.3 mM  $NaN_3$ , 0.44 mg/ml BSA, 1.43  $\mu$ M coelenterazine) with the same settings of measurement as for F-Luc. The relative activity of firefly luciferase (F-Luc) was calculated by normalizing with the activity of the Renilla luciferase (R-Luc) to yield either  $FF^{AUG}/R^{AUG}$  (AUG) or  $FF^{UUG}/R^{AUG}$  (UUG) values. To calculate the normalized activity of firefly luciferase starting with UUG as an initiation codon, the  $FF^{UUG}/R^{AUG}$  value was normalized to  $FF^{AUG}/R^{AUG}$  for the UUG/AUG ratio. The normalized expression of other near cognate codons (e.g., ACG, AUU) was calculated in a similar manner.

$\beta$ -Galactosidase activity assays in whole-cell extracts (WCEs) were performed as described previously [57]. To assay the expression of eIF1 (encoded by *SLI1*) by western blot analysis, the WCEs were made under denaturing conditions as described previously [58] from four biological replicates (independent transformants). The immunoblot analysis was performed [59] using antibodies against eIF1 [60] and Ded1 (a kind gift from Tien-Hsien Chang). Two technical replicates were performed using the same extracts, and two-fold different amounts of each extract were loaded in two successive lanes. Enhanced chemiluminescence (Amersham #RPN2106) was used to detect the immune complexes using a ProteinSimple imager (FluorChem systems #FM0261), and

the signal intensities were quantified by densitometry using Adobe Photoshop after inversion of the image as described [61].

### Ribosome profiling

#### Cell culture and lysis

BY4741 cells harboring  $pR^{AUG}FF^{UUG}$  plasmid (dual-luciferase reporter with R-Luc (AUG)-F-Luc (UUG) in a *URA3* vector (see Additional file 2: Table S2) were used for the ribosome profiling, which was performed as described [8, 19, 30, 62] with some modifications. Seven hundred fifty milliliters of yeast cells in log phase ( $OD_{600}$  0.6–0.8) grown in synthetic complete medium lacking uracil (SC-Ura) at either 20 °C, 30 °C, or 37 °C for 16 h were harvested rapidly ( $\leq 1$  min) by vacuum filtration at room temperature and snap-frozen in liquid nitrogen. No cycloheximide was added to the media before harvesting to avoid cycloheximide-induced artifacts [31–34]. This was followed by addition of ice-cold lysis buffer (20 mM Tris [pH 8.0], 140 mM KCl, 1.5 mM  $MgCl_2$ , 1% Triton, 500  $\mu$ g/ml cycloheximide). The concentration of cycloheximide in this lysis buffer was five times the concentration originally used in ribosome profiling experiments [8] to minimize continued translation elongation during lysis and processing. The frozen pellets in lysis buffer were lysed in a freezer mill (Freezer/Mill® Dual Chamber Cryogenic Grinder, #6870, with the settings of 15 cycles, at 15 Hz, 5 min precool, 1 min run followed by 2 min cooling). The frozen lysate was transferred to a 50-ml conical tube and thawed on ice with frequent agitation. The cell lysate was centrifuged at 3000 rpm (Eppendorf #5810R), at 4 °C for 5 min, and the supernatant was collected and further centrifuged at 13,000 rpm ( $\sim 18,000$  rcf) in a table top centrifuge (Eppendorf #5417R) at 4 °C for 10 min. The final supernatant was collected, and the  $OD_{260}$  was measured. Aliquots of lysate containing 30  $OD_{260}$  units were stored at  $-80$  °C until used.

#### Isolation of 80S ribosomal footprints for sequencing library preparation

An aliquot of lysate containing 30  $OD_{260}$  units was thawed on ice and treated with 500 U of RNase I (Ambion™ #AM2294) for 60 min at 26 °C in a thermomixer at 700 rpm. Five microliters of SUPERase-In RNase inhibitor (Ambion™ #AM2694) was added to the reaction, which was then used for isolation of 80S monosomes as previously described [30]. Briefly, to isolate 80S monosome, the reaction was loaded on a 10–50% (w/v) sucrose gradient made in lysis buffer, followed by centrifugation for 3 h at 40,000 rpm in a SW 41 Ti rotor. The fractions were separated on density gradient fractionation system (Brandel) using 60% sucrose solution (made in lysis buffer) to pump the gradients. The

fractions corresponding to 80S were collected, and the ribosome-protected mRNA fragments (RPFs) were purified with SDS/hot acid phenol and chloroform.

#### **Isolation of total mRNA for RNA-seq library preparation**

Total RNA was isolated from the lysate using miRNeasy Mini Kit (Qiagen #217004) following the manufacturer's protocol. The random fragmentation was carried out by the addition of fragmentation reagent (Ambion #AM8740) and incubation at 70 °C for 8 min, followed by the addition of stop solution (from the same kit).

#### **Sequencing library construction**

The ribosome protected fragments (RPFs) and fragmented total mRNA were each resolved by electrophoresis on a 15% TBE-Urea gel (Novex #EC68852BOX). Following the size selection, the RNA was gel extracted and was dephosphorylated using polynucleotide kinase (NEB #M0201S). A universal miRNA cloning linker (NEB #S1315S) was ligated to the 3'-ends of the RNA using T4 RNA Ligase 2 (NEB #M0242L) in the presence of PEG 8000 (2.5% w/v), DMSO and SUPERase-In at 37 °C for 2.5 h. The ligated products were resolved by electrophoresis on 15% TBE-Urea gel and the appropriate size fragments eluted from the gel. The linker-ligated RNA recovered from RPFs was directly subjected to reverse transcription, while that extracted from fragmented total RNA samples was used as input for the RiboZero reaction (Illumina Ribo-Zero Gold rRNA Removal Kit -Yeast) to remove rRNAs, and subsequently subjected to reverse transcription. For reverse transcription, 10 µl RNA (dissolved in 10 mM Tris pH 8.0) from the previous reaction was mixed with 2 µl of 1.25 µM reverse transcription primer [5'-(Phos) AGATCGGAAGAGCG TCGTGTAGGGAAAGAGTGTAGATCTCGG TGGTCGC (SpC18) CACTA (SpC18) TTCAGACGTG TGCTCTCCGATCTATTGATGGTGCCTACAG], denatured at 80 °C for 2 min and incubated on ice. This was followed by the addition of SuperScript™ III Reverse Transcriptase (#AM2694), dNTPs, DTT, and Superase-In and incubation at 48 °C for 30 min. The RNA template was removed by the addition of 2.2 µl of 1 N NaOH followed by incubation at 98 °C for 20 min. The reaction was resolved on a 15% TBE-Urea gel, and the cDNA was extracted from the gel. Circularization using CircLigase (Epicentre #CL4111K) was conducted by mixing 15 µl of cDNA (dissolved in 10 mM Tris pH 8.0) with 2 µl of 10× CircLigase buffer, 1 µl of 1 mM ATP, 1 µl of 50 mM MnCl<sub>2</sub>, and 1 µl of CircLigase. The reaction was incubated at 60 °C for 1 h, followed by heat inactivation at 80 °C for 10 min.

The circularized products were used as templates for PCR amplification of the total RNA sample, or first subjected to subtractive hybridization to remove rRNA-

derived sequences in case of RPF libraries. For the latter, the circularization reaction was mixed with a subtraction pool of biotinylated oligonucleotides [62] in the presence of 2× SSC (#AM9763), denatured at 100 °C for 90 s, followed by annealing at 37 °C. The reaction was then mixed with Dynabeads and incubated at 37 °C in a thermomixer at 1000 rpm. The eluate was recovered and treated as the rRNA-depleted sample. The latter, as well as the circularized product derived from total RNA, was used as templates for PCR amplification to produce sequencing libraries using previously published primers [62] with slight changes in the barcodes. The resulting libraries were sequenced using the Illumina HiSeq system at the DNA Sequencing and Genomics Core, NHLBI, NIH. The reads were trimmed to remove the linker sequences (using fastx\_toolkit/fastx\_trimmer [[http://hannonlab.cshl.edu/fastx\\_toolkit/index.html](http://hannonlab.cshl.edu/fastx_toolkit/index.html)] in the relevant command line code) and then aligned to the *S. cerevisiae* rRNA database using Bowtie [63]. The reads that did not align with the reference database (non-rRNA reads) were aligned to the *S. cerevisiae* genome or to the FASTA file made from the reporter sequence (pR<sup>AUG</sup>F-F<sup>UUG</sup>) to generate alignments to the F-Luc reporter mRNA using TopHat [64].

#### **Data analysis, statistics, and web tools**

For ribosomal profiling, each experiment was performed with two independent cultures as biological replicates. The statistical analysis between the replicates was performed using DESeq2 [65]. The DESeq2 statistical package addresses the typical challenge in next-generation sequencing experiments of having only two biological replicates for each condition by pooling information about the variances of read counts across the thousands of genes being analyzed in order to model count variances for genes of similar expression levels. The modeled variances are used in the framework of a generalized linear model (GLM), an extension of linear regression allowing for non-normal error distributions, to identify expression changes and place confidence intervals on the magnitude of changes, and also to exclude genes showing aberrantly high variability. Transcriptional and translational changes can be analyzed together in a GLM by including library type (mRNA-Seq or Ribo-Seq) as one of the factors, in addition to experimental variables like genotype or drug treatment, in a multi-factor design. The translational efficiency (TE) emerges as the effect of the Ribo-seq library type against the mRNA-Seq baseline, and significant interactions of TE with the experimental variables indicate translational control [66]. While the analysis is under-powered with two versus three biological replicates, only the chance of false negatives, not false positives, is increased. Using DESeq2, we calculated the changes in the ribosome density (ribosome protected fragments per million mapped reads),

mRNA density (RNA-seq reads per million mapped reads), and the translation efficiency (TE, ribosome density/mRNA density) at different temperatures using a cutoff of  $\geq 10$  average mRNA reads in four samples. To calculate TE of a uORF ( $TE_{uORF}$ ), mRNA read counts for the mORF were employed rather than the mRNA read counts in the uORF alone because the mORFs have lower noise. Spearman's correlation was calculated using an online tool at [https://www.wessa.net/rwasp\\_spearman.wasp](https://www.wessa.net/rwasp_spearman.wasp). Notched box and whisker plot analysis was conducted using a web tool <http://shiny.chemgrid.org/boxplot/>. The notches indicate  $\pm 1.58 \times \text{interquartile range (IQR)} / \sqrt{n}$ , where IQR is the difference between the 75th and 25th percentiles and  $n$  represents the number of data points in that bin. Non-overlapping notches give roughly 95% confidence that the two medians differ. Heatmaps were generated using a web tool "Heatmapper" (<http://www.heatmapper.ca/>) [67].

### Wiggle tracks

A combined alignment file (Bam file) was generated using two alignment files (one for each of the two biological replicates). The combined files were generated for both RPF samples and total RNA samples (for 20 °C, 30 °C, and 37 °C). Wiggle files were generated from this combined alignment file. Files were generated for each gene on the Watson or Crick strand. The tracks were visualized using Integrative Genomics Viewer (IGV 2.4.14). The tracks were normalized according to the total number of mapped reads in the combined file. To normalize the effects of the changes in mRNA levels, the total read-normalized peaks were scaled with respect to the changes in mRNA levels to reflect the changes in translation efficiency (see the "Results" section). Wiggle tracks, both with and without the scaling by changes in mRNA levels, are shown.

### Finding upstream ORFs

We took a similar approach to identify possible uORFs and confirm their translation in our experiments as described previously [19]. First, putative translated uORFs were identified essentially as described [35]. Briefly, for all open reading frames in annotated 5'-UTRs that initiate from either AUGs or near cognate codons, the ratio of RPF counts at the +1 position (start codon of uORF) to -1 position (upstream of the start codon) was calculated. Those uORFs with ratios  $>4$ , with  $>14$  RPF counts at the +1 and -1 positions combined, and with at least 50% of the count reads in the 0 frame with respect to the start site (i.e., the relevant line of code is -c15-r4-z0.5), were selected for further analysis. The multiple ribosome profiling datasets we used to identify potential uORFs were described previously [19] and have been submitted to the NCBI Gene Expression Omnibus, and the accession numbers are listed at <https://elife-sciences.org/articles/31250/figures#suppl> in

Additional file 2: Table S2. The RNA-seq data have been submitted to the NCBI Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo>) with GEO accession number GSE137021. The details of the RNA-seq and Ribo-seq datasets generated in this study are provided in Additional file 2: Table S5.

To determine which of these putative uORFs were translated in our experiments, we employed an ORF identification tool (RibORF) described previously [36], which uses 3-nucleotide periodicity and a uniform distribution of RPFs counts across the uORF as scoring criteria. We applied a moderately stringent cutoff of the probability of prediction of 0.5 and used a combined alignment file generated from footprint libraries of all 6 biological replicates grown at 3 different temperatures to identify uORFs with evidence of translation at 1 or more growth temperatures. After excluding uORFs shorter than 3 codons, we identified 1367 uORFs starting with AUG ( $N=142$ ) or a NCC ( $N=1225$ ). Quantifying the total mRNA and RPF counts in the 5'-UTR, uORF or mORF was done as previously described [19].

To identify potential uORFs in all yeast mRNA 5'-UTR transcriptome, 5'-UTR sequences for all mRNAs were extracted, and each AUG and NCC nucleotide triplets were searched throughout the 5'-UTR. We identified the first in frame stop codon for each start codon as the end of the uORF. For uORFs without an in-frame stop codon in 5'-UTR, we defined the end of the 5'-UTR as the end of uORF. uORFs with less than 3 codons in length were excluded in downstream analysis.

### Generation of C-terminally HA-tagged clones of *AGA1* gene

Gibson assembly master mix (NEB# E2611) was used to generate C-terminally HA-tagged *AGA1* clones [*AGA1*-HA (WT), see plasmid #12, Additional file 2: Table S2]. The protocol for four to six fragment assemblies was employed. The four fragments assembled are as follows: double-stranded HA tag (~108 nt), high-copy vector (pRS426 digested with BamHI and SacI, followed by gel extraction), a PCR product (P1, ~3000 bp) obtained by using a forward primer binding 650 nt upstream of the start codon and reverse primer binding immediately downstream of stop codon (on the Crick strand) of the *AGA1* coding region, and a PCR product (P2, 300 bp) obtained by using a forward primer binding immediately downstream of the stop codon and reverse primer binding 300 nt upstream of stop codon (on the Crick strand) of the *AGA1* coding region. The primers were designed using NEBuilder assembly tool (v2.2.5) with minor changes. The genomic DNA from BY4741 cells was used in all PCR reactions. All four fragments were gel extracted and then ligated using 2 $\times$  Gibson master mix according to the manufacturer's instructions. One



microliter of the assembly was transformed in NEB® 5-alpha Competent *E. coli*-high efficiency cells (#C2987) followed by selection on LB-carbenicillin plates. The colonies were screened by colony PCR using M13 forward (M13F) and M13 reverse (M13R) primers, followed by restriction mapping and sequencing using M13F and M13R primers. The uORF-start mutant [AGA1-HA (Mutant), see plasmid #13, Additional file 2: Table S2] was generated from AGA1-HA (WT) plasmid by a quick change PCR using PfuTurbo DNA Polymerase (Agilent Technologies #600252). The mutation was confirmed by sequencing with M13F primer. Yeast cells harboring WT or mutant plasmids were selected on the appropriate media lacking the nutrient(s) corresponding to the auxotrophic marker(s). The yeast cell cultures, lysate preparation, and western blot analysis were performed as described above. The C-terminally HA-tagged Aga1 protein was detected using an HA tag monoclonal antibody (2-2.2.14) from Invitrogen (#26183).

## Supplementary information

**Supplementary information** accompanies this paper at <https://doi.org/10.1186/s12915-019-0718-5>.

**Additional file 1: Figure S1.**  $\beta$ -galactosidase activity assay to calculate usage of NCC start site at multiple growth temperature. **Figure S2.** Temperature dependent alterations in usage of NCC start site are not evoked due to changes in the levels of initiation factors. **Figure S3.** Ribosome profiling at multiple growth temperatures. **Figure S4.** Changes in growth temperature lead to changes in translation of uORFs. **Figure S5.** Analysis of translation efficiency changes of conserved uORFs upon changes in growth temperature. **Figure S6.** Analysis of the properties of various sets of uORFs and the mRNAs harboring them. **Figure S7.** Analysis of uORFs showing highest changes in translation at 20 °C and 37 °C. **Figure S8.** Box-plot analyses of 5'-UTR features for the most highly regulated NCC uORFs. **Figure S9.** Analysis of TE-changes of mORFs. **Figure S10.** Wiggle track images of examples of mRNAs that show temperature-dependent changes in translation of one or more uORF and the mORF. **Figure S11.** uORF-mediated temperature-dependent regulation of AGA1 mRNA.

**Additional file 2: Table S1.** Yeast strains used in this study. **Table S2.** Plasmids used in this study. **Table S3.** List of uORFs showing temperature-dependent translational regulation. **Table S4.** List of putative N-terminal extensions identified in this study. **Table S5.** List of datasets generated in this study and their GEO accession numbers.

**Additional file 3.** uORF and mORF TE information: The file includes following datasheets: **Sheet 1.** Includes information regarding all the uORFs identified and studied here. **Sheet 2.** DESeq2 output for calculation of  $\Delta TE_{uORF}$  when cultured at 20 °C with respect to 30 °C. **Sheet 3.** DESeq2 output for calculation of  $\Delta TE_{uORF}$  when cultured at 37 °C with respect to 30 °C. **Sheet 4.** DESeq2 output for calculation of  $\Delta TE_{mORF}$  when cultured at 20 °C with respect to 30 °C. **Sheet 5.** DESeq2 output for calculation of  $\Delta TE_{mORF}$  when cultured at 37 °C with respect to 30 °C.

**Additional file 4.** Supplementary Methods [68].

## Acknowledgements

We thank Thomas Dever and Nicholas Guydosh for the thoughtful comments and suggestions. We are grateful to Nick Ingolia for the advice in the early phases of this project. We would also like to thank all the members of Lorsch, Hinnebusch, Dever, and Guydosh labs for their useful comments and our reviewers for the helpful suggestions for improving the manuscript.

## Authors' contributions

All authors read and approved the final manuscript. SDK contributed to the conceptualization, resources, data curation, formal analysis, investigation, visualization, methodology, writing—original draft—review, and editing. FZ and NDS contributed to the resources, data curation, formal analysis, and methodology. HZ contributed to the methodology. AGH and JRL contributed to the conceptualization, formal analysis, supervision, funding acquisition, visualization, project administration, writing, review, and editing.

## Funding

This work was supported by the Intramural Research Program of the National Institutes of Health (AGH and JRL). The funders had no role in the study design, data collection, interpretation, or the decision to submit the work for publication.

## Availability of data and materials

The RNA-seq data generated in this study have been submitted to the NCBI Gene Expression Omnibus (GEO accession number GSE137021) [69], and the details are provided in Additional file 2: Table S5, sr. no. 1–6. The ribo-seq data analyzed were published in [19], and accession numbers are listed in Additional file 2: Table S5, sr. no. 7–12. The additional datasets generated in this study (uORF information, TE-changes, etc.) are provided in Additional file 3.

## Ethics approval and consent to participate

Not applicable

## Consent for publication

Not applicable

## Competing interests

The authors declare that they have no competing interests.

## Author details

<sup>1</sup>Laboratory on the Mechanism and Regulation of Protein Synthesis, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD, USA. <sup>2</sup>Laboratory of Gene Regulation and Development, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD, USA. <sup>3</sup>Present Address: School of Life Sciences, Jawaharlal Nehru University, New Delhi 110067, India.

Received: 2 July 2019 Accepted: 1 November 2019

Published online: 06 December 2019

## References

- Hinnebusch AG. The scanning mechanism of eukaryotic translation initiation. *Annu Rev Biochem.* 2014;83:779–812.
- Aitken CE, Lorsch JR. A mechanistic overview of translation initiation in eukaryotes. *Nat Struct Mol Biol.* 2012;19(6):568–76.
- Hinnebusch AG, Lorsch JR. The mechanism of eukaryotic translation initiation: new insights and challenges. *Cold Spring Harb Perspect Biol.* 2012;4(10):a011544.
- Kolitz SE, Takacs JE, Lorsch JR. Kinetic and thermodynamic analysis of the role of start codon/anticodon base pairing during eukaryotic translation initiation. *RNA.* 2009;15(1):138–52.
- Chang KJ, Wang CC. Translation initiation from a naturally occurring non-AUG codon in *Saccharomyces cerevisiae*. *J Biol Chem.* 2004;279(14):13778–85.
- Tang HL, Yeh LS, Chen NK, Ripmaster T, Schimmel P, Wang CC. Translation of a yeast mitochondrial tRNA synthetase initiated at redundant non-AUG codons. *J Biol Chem.* 2004;279(48):49656–63.
- Ingolia NT, Brar GA, Stern-Ginossar N, Harris MS, Talhouarne GJ, Jackson SE, Wills MR, Weissman JS. Ribosome profiling reveals pervasive translation outside of annotated protein-coding genes. *Cell Rep.* 2014;8(5):1365–79.
- Ingolia NT, Ghaemmaghami S, Newman JR, Weissman JS. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science.* 2009;324(5924):218–23.
- Arriberre JA, Gilbert WW. Roles for transcript leaders in translation and mRNA decay revealed by transcript leader sequencing. *Genome Res.* 2013;23(6):977–87.

10. Spealman P, Naik AW, May GE, Kuersten S, Freeberg L, Murphy RF, McManus J. Conserved non-AUG uORFs revealed by a novel regression analysis of ribosome profiling data. *Genome Res.* 2018; 28(2):214–22.
11. Bazykin GA, Kochetov AV. Alternative translation start sites are conserved in eukaryotic genomes. *Nucleic Acids Res.* 2011;39(2):567–77.
12. Barbosa C, Peixeiro I, Romao L. Gene expression regulation by upstream open reading frames and human disease. *PLoS Genet.* 2013; 9(8):e1003529.
13. Morris DR, Geballe AP. Upstream open reading frames as regulators of mRNA translation. *Mol Cell Biol.* 2000;20(23):8635–42.
14. Gerashchenko MV, Lobanov AV, Gladyshev VN. Genome-wide ribosome profiling reveals complex translational regulation in response to oxidative stress. *Proc Natl Acad Sci U S A.* 2012;109(43):17394–9.
15. Outten CE, Culotta VC. Alternative start sites in the *Saccharomyces cerevisiae* GLR1 gene are responsible for mitochondrial and cytosolic isoforms of glutathione reductase. *J Biol Chem.* 2004;279(9):7785–91.
16. Slusher LB, Gillman EC, Martin NC, Hopper AK. mRNA leader length and initiation codon context determine alternative AUG selection for the yeast gene MOD5. *Proc Natl Acad Sci U S A.* 1991;88(21):9789–93.
17. Helsens K, Van Damme P, Degroeve S, Martens L, Arnesen T, Vandekerckhove J, Gevaert K. Bioinformatics analysis of a *Saccharomyces cerevisiae* N-terminal proteome provides evidence of alternative translation initiation and post-translational N-terminal acetylation. *J Proteome Res.* 2011;10(8):3578–89.
18. Hinnebusch AG. Structural insights into the mechanism of scanning and start codon recognition in eukaryotic translation initiation. *Trends Biochem Sci.* 2017;42(8):589–611.
19. Martin-Marcos P, Zhou F, Karunasiri C, Zhang F, Dong J, Nanda J, Kulkarni SD, Sen ND, Tamame M, Zeschnigk M, et al. eIF1A residues implicated in cancer stabilize translation preinitiation complexes and favor suboptimal initiation sites in yeast. *Elife.* 2017;6:e31250.
20. Takacs JE, Neary TB, Ingolia NT, Saini AK, Martin-Marcos P, Pelletier J, Hinnebusch AG, Lorsch JR. Identification of compounds that decrease the fidelity of start codon recognition by the eukaryotic translational machinery. *RNA.* 2011;17(3):439–52.
21. Donahue TF, Cigan AM. Genetic selection for mutations that reduce or abolish ribosomal recognition of the HIS4 translational initiator region. *Mol Cell Biol.* 1988;8(7):2955–63.
22. Amrani N, Sachs MS, Jacobson A. Early nonsense: mRNA decay solves a translational problem. *Nat Rev Mol Cell Biol.* 2006;7(6):415–25.
23. Welch EM, Jacobson A. An internal open reading frame triggers nonsense-mediated decay of the yeast SPT10 mRNA. *EMBO J.* 1999;18(21):6134–45.
24. He F, Peltz SW, Donahue JL, Rosbash M, Jacobson A. Stabilization and ribosome association of spliced pre-mRNAs in a yeast upf1<sup>-</sup> mutant. *Proc Natl Acad Sci U S A.* 1993;90(15):7034–8.
25. Maag D, Fekete CA, Gryczynski Z, Lorsch JR. A conformational change in the eukaryotic translation preinitiation complex and release of eIF1 signal recognition of the start codon. *Mol Cell.* 2005;17(2):265–75.
26. Cheung YN, Maag D, Mitchell SF, Fekete CA, Algire MA, Takacs JE, Shirokikh N, Pestova T, Lorsch JR, Hinnebusch AG. Dissociation of eIF1 from the 40S ribosomal subunit is a key step in start codon selection in vivo. *Genes Dev.* 2007;21(10):1217–30.
27. He H, von der Haar T, Singh CR, Li M, Li B, Hinnebusch AG, McCarthy JE, Asano K. The yeast eukaryotic initiation factor 4G (eIF4G) HEAT domain interacts with eIF1 and eIF5 and is involved in stringent AUG selection. *Mol Cell Biol.* 2003;23(15):5431–45.
28. Saini AK, Nanda JS, Lorsch JR, Hinnebusch AG. Regulatory elements in eIF1A control the fidelity of start codon selection by modulating tRNA(i)(Met) binding to the ribosome. *Genes Dev.* 2010;24(1):97–110.
29. Valasek L, Nielsen KH, Zhang F, Fekete CA, Hinnebusch AG. Interactions of eukaryotic translation initiation factor 3 (eIF3) subunit NIP1/c with eIF1 and eIF5 promote preinitiation complex assembly and regulate start codon selection. *Mol Cell Biol.* 2004;24(21):9437–55.
30. Guydosh NR, Green R. Dom34 rescues ribosomes in 3' untranslated regions. *Cell.* 2014;156(5):950–62.
31. Hussmann JA, Patchett S, Johnson A, Sawyer S, Press WH. Understanding biases in ribosome profiling experiments reveals signatures of translation dynamics in yeast. *PLoS Genet.* 2015;11(12):e1005732.
32. Gerashchenko MV, Gladyshev VN. Translation inhibitors cause abnormalities in ribosome profiling experiments. *Nucleic Acids Res.* 2014;42(17):e134.
33. Schuller AP, Wu CC, Dever TE, Buskirk AR, Green R. eIF5A functions globally in translation elongation and termination. *Mol Cell.* 2017;66(2):194–205 e195.
34. Urban J, Soulard A, Huber A, Lippman S, Mukhopadhyay D, Deloche O, Wanke V, Anrather D, Ammerer G, Riezman H, et al. Sch9 is a major target of TORC1 in *Saccharomyces cerevisiae*. *Mol Cell.* 2007;26(5):663–74.
35. Brar GA, Yassour M, Friedman N, Regev A, Ingolia NT, Weissman JS. High-resolution view of the yeast meiotic program revealed by ribosome profiling. *Science.* 2012;335(6068):552–7.
36. Ji Z, Song R, Regev A, Struhl K. Many lncRNAs, 5'UTRs, and pseudogenes are translated and some are likely to express functional proteins. *Elife.* 2015;4: e08890.
37. Chen SJ, Lin G, Chang KJ, Yeh LS, Wang CC. Translational efficiency of a non-AUG initiation codon is significantly affected by its sequence context in yeast. *J Biol Chem.* 2008;283(6):3173–80.
38. Shabalina SA, Ogurtsov AY, Rogozin IB, Koonin EV, Lipman DJ. Comparative analysis of orthologous eukaryotic mRNAs: potential hidden functional signals. *Nucleic Acids Res.* 2004;32(5):1774–82.
39. Rojas-Duran MF, Gilbert WV. Alternative transcription start site selection leads to large differences in translation activity in yeast. *RNA.* 2012;18(12): 2299–305.
40. Kertesz M, Wan Y, Mazor E, Rinn JL, Nutter RC, Chang HY, Segal E. Genome-wide measurement of RNA secondary structure in yeast. *Nature.* 2010; 467(7311):103–7.
41. Miyasaka H. The positive relationship between codon usage bias and translation initiation AUG context in *Saccharomyces cerevisiae*. *Yeast.* 1999; 15(8):633–7.
42. Zur H, Tuller T. New universal rules of eukaryotic translation initiation fidelity. *PLoS Comput Biol.* 2013;9(7):e1003136.
43. Parker R. RNA degradation in *Saccharomyces cerevisiae*. *Genetics.* 2012; 191(3):671–702.
44. Presnyak V, Alhusaini N, Chen YH, Martin S, Morris N, Kline N, Olson S, Weinberg D, Baker KE, Graveley BR, et al. Codon optimality is a major determinant of mRNA stability. *Cell.* 2015;160(6):1111–24.
45. Hinnebusch AG. Translational regulation of GCN4 and the general amino acid control of yeast. *Annu Rev Microbiol.* 2005;59:407–50.
46. Zhang F, Hinnebusch AG. An upstream ORF with non-AUG start codon is translated in vivo but dispensable for translational control of GCN4 mRNA. *Nucleic Acids Res.* 2011;39(8):3128–40.
47. Gaba A, Jacobson A, Sachs MS. Ribosome occupancy of the yeast CPA1 upstream open reading frame termination codon modulates nonsense-mediated mRNA decay. *Mol Cell.* 2005;20(3):449–60.
48. Werner M, Feller A, Messenguy F, Pierard A. The leader peptide of yeast gene CPA1 is essential for the translational repression of its expression. *Cell.* 1987;49(6):805–13.
49. Delbecq P, Werner M, Feller A, Filipkowski RK, Messenguy F, Pierard A. A segment of mRNA encoding the leader peptide of the CPA1 gene confers repression by arginine on a heterologous yeast gene transcript. *Mol Cell Biol.* 1994;14(4):2378–90.
50. Lipke PN, Kurjan J. Sexual agglutination in budding yeasts: structure, function, and regulation of adhesion glycoproteins. *Microbiol Rev.* 1992; 56(1):180–94.
51. Guenther UP, Weinberg DE, Zubradt MM, Tedeschi FA, Stawicki BN, Zagore LL, Brar GA, Licatalosi DD, Bartel DP, Weissman JS, et al. The helicase Ded1p controls use of near-cognate translation initiation codons in 5' UTRs. *Nature.* 2018;559(7712):130–4.
52. Kozak M. Downstream secondary structure facilitates recognition of initiator codons by eukaryotic ribosomes. *Proc Natl Acad Sci U S A.* 1990;87(21): 8301–5.
53. Pizarro FJ, Jewett MC, Nielsen J, Agosin E. Growth temperature exerts differential physiological and transcriptional responses in laboratory and wine strains of *Saccharomyces cerevisiae*. *Appl Environ Microbiol.* 2008; 74(20):6358–68.
54. Strassburg K, Walther D, Takahashi H, Kanaya S, Kopka J. Dynamic transcriptional and metabolic responses in yeast adapting to temperature stress. *OMICS.* 2010;14(3):249–59.
55. Touriol C, Bornes S, Bonnal S, Audigier S, Prats H, Prats AC, Vagner S. Generation of protein isoform diversity by alternative initiation of translation at non-AUG codons. *Biol Cell.* 2003;95(3–4):169–78.
56. Gietz RD, Woods RA. Yeast transformation by the LiAc/SS carrier DNA/PEG method. *Methods Mol Biol.* 2006;313:107–20.

57. Moehle CM, Hinnebusch AG. Association of RAP1 binding sites with stringent control of ribosomal protein gene transcription in *Saccharomyces cerevisiae*. *Mol Cell Biol*. 1991;11(5):2723–35.
58. Reid GA, Schatz G. Import of proteins into mitochondria. Extramitochondrial pools and post-translational import of mitochondrial protein precursors in vivo. *J Biol Chem*. 1982;257(21):13062–7.
59. Nanda JS, Cheung YN, Takacs JE, Martin-Marcos P, Saini AK, Hinnebusch AG, Lorsch JR. eIF1 controls multiple steps in start codon recognition during eukaryotic translation initiation. *J Mol Biol*. 2009;394(2):268–85.
60. Olsen DS, Savner EM, Mathew A, Zhang F, Krishnamoorthy T, Phan L, Hinnebusch AG. Domains of eIF1A that mediate binding to eIF2, eIF3 and eIF5B and promote ternary complex recruitment in vivo. *EMBO J*. 2003;22(2):193–204.
61. Kulkarni SD, Muralidharan B, Panda AC, Bakthavachalu B, Vindu A, Seshadri V. Glucose-stimulated translation regulation of insulin by the 5' UTR-binding proteins. *J Biol Chem*. 2011;286(16):14146–56.
62. Ingolia NT, Brar GA, Rouskin S, McGeachy AM, Weissman JS. The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nat Protoc*. 2012;7(8):1534–50.
63. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 2009;10(3):R25.
64. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*. 2009;25(9):1105–11.
65. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
66. Ingolia NT. Ribosome footprint profiling of translation throughout the genome. *Cell*. 2016;165(1):22–33.
67. Babicki S, Arndt D, Marcu A, Liang Y, Grant JR, Maciejewski A, Wishart DS. Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Res*. 2016;44(W1):W147–53.
68. Sen ND, Zhou F, Ingolia NT, Hinnebusch AG. Genome-wide analysis of translational efficiency reveals distinct but overlapping functions of yeast DEAD-box RNA helicases Ded1 and eIF4A. *Genome Res*. 2015;25(8):1196–205.
69. Kulkarni SD, Zhou F, Sen ND, Zhang H, Hinnebusch AG, Lorsch JR. Temperature-dependent regulation of upstream open reading frame translation in *S. cerevisiae*. *Suppl Datasets*. 2019; NCBI accession: GSE137021.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

