

RESEARCH

Open Access

A comparative genomics study on the effect of individual amino acids on ribosome stalling

Renana Sabi¹, Tamir Tuller^{1,2*}

From 13th Annual Research in Computational Molecular Biology (RECOMB) Satellite Workshop on Comparative Genomics

Frankfurt, Germany. 4-7 October 2015

Abstract

Background: During protein synthesis, the nascent peptide chain emerges from the ribosome through the ribosomal exit tunnel. Biochemical interactions between the nascent peptide and the tunnel may stall the ribosome movement and thus affect the expression level of the protein being synthesized. Earlier studies focused on one model organism (*S. cerevisiae*), have suggested that certain amino acid sequences may be responsible for ribosome stalling; however, the stalling effect at the individual amino acid level across many organisms has not yet been quantified.

Results: By analyzing multiple ribosome profiling datasets from different organisms (including prokaryotes and eukaryotes), we report for the first time the organism-specific amino acids that significantly lead to ribosome stalling. We show that the identity of the stalling amino acids vary across the tree of life. In agreement with previous studies, we observed a remarkable stalling signal of proline and arginine in *S. cerevisiae*. In addition, our analysis supports the conjecture that the stalling effect of positively charged amino acids is not universal and that in certain conditions, negative charge may also induce ribosome stalling. Finally, we show that the beginning part of the tunnel tends to undergo more interactions with the translated amino acids than other positions along the tunnel.

Conclusions: The reported results support the conjecture that the ribosomal exit tunnel interacts with various amino acids and that the nature of these interactions varies among different organisms. Our findings should contribute towards better understanding of transcript and proteomic evolution and translation elongation regulation.

Background

mRNAs translation is a fundamental intracellular process which occurs in all living organisms. Translation elongation is an iterative stage of translation in which the ribosome scans the mRNA sequence and decodes it into a specific protein by adding one amino acid at the time to the growing peptide chain. It has been suggested that the speed by which ribosomes progress along the mRNA is affected by different local features of the coding sequence. One determinant of the translation elongation speed is the identity of the codon at the P-site; it has been suggested that the codon decoding rate is influenced by several factors related to the P-site, including: the cellular concentration of the

paired tRNA [1-6]; the efficiency of the codon-anticodon pairing which occurs non-optimally for wobble base pairing [7-9] and the efficiency of incorporation of the decoded amino acid into the polypeptide which is mainly poor in the case of proline [10-12]. Other coding sequence features thought to slow down ribosomes include: the folding energy of the mRNA sequence downstream from the ribosomal P-site [13-16]; the identity of the tRNA at the A-site [17]; and the charge of the amino acids in the exit tunnel [16,18,19].

The Ribosomal Exit Tunnel (*RET*) is the site through which nascent peptides leave the ribosome during translation. The non-uniform biochemical characteristics of the tunnel allow it to play an important role in affecting translation rates and protein folding rather than being a passive conduit for the nascent polypeptide. First, the overall electrostatic potential of the *RET* is negative and

* Correspondence: tamirtul@post.tau.ac.il

¹Department of Biomedical Engineering, Tel Aviv University (TAU), Tel Aviv, Israel

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

varies in magnitude along the tunnel [19-23]; thus, it was suggested that a nascent peptide that contains charged amino acids may undergo electrostatic interaction with the exit tunnel [19]. Second, the diameter of the tunnel varies between 10\AA^0 and 20\AA^0 [24-27]; thus, the interaction between that nascent peptide and the exit tunnel may also be dictated by geometrical constraints. Although the expansion in diameter enables the partial folding of the translated peptide [28], the cramped dimensions of the tunnel prohibit a folding of whole protein domains and only tertiary/secondary structures of small segments are allowed [29].

Evidences of ribosome pausing mediated by nascent peptide have been manifested in several studies [30-36]. These studies, however, either conducted a small scale experiment or focused on one organism only.

The development of the ribosome profiling technique has significantly broadened the comprehension of *in vivo* translation by enabling the detection of the momentary positions of ribosomes along the transcripts at nucleotide resolution [37]. During the past few years, the high throughput quantitative data obtained by ribosome profiling experiments has been widely used to study gene translation [10,16,18,38-51].

Specifically, ribosome profiling data was used to show that ribosome stalling is induced in response to the presence of certain amino acid [10,16,18]. Specifically, it has been suggested that positively charged amino acids are implicated in transient ribosomal pauses by interacting with the negatively charged exit tunnel [16,18,19]. A more recent study of Artieri and Fraiser [10], on the other hand, emphasized the possibility that the incorporation of proline into the nascent peptide has the major effect on ribosome stalling.

In order to investigate the organism-specific influence of each individual amino acid on substantial ribosome stalling, we performed a large scale analysis based on multiple ribosome profiling datasets of 9 organisms including eukaryotes (*H.sapiens*, *C.elegans*, *S.cerevisiae*, *S.pombe*, *A.thaliana*, *P.falciparum*, *D.melanogaster*, *M.musculus*) and bacteria (*C.crescentus*).

Results

Ribosome profiling experiments include the following major stages (Figure 1A): cells are treated with cycloheximide (for example) to arrest translating ribosomes; then, RNA fragments protected by ribosomes from RNases are isolated and processed for high-throughput sequencing, resulting in reads of ribosomes protected footprints. As slowly decoded codons are covered by ribosomes for a larger amount of time, they tend to create higher amount of protected fragments, in comparison to faster decoded codons on the same transcript. Finally, using a computational method, the obtained sequenced footprints are

mapped to the genome of the analyzed organism creating for each gene a ribosomal footprints read count profile. This profile will be referred here as a *RD* profile for Ribosomal Density.

In this work, we aim at understanding whether extreme ribosomal stalling occurs at a specific codon is affected by an interaction between the *RET* and the amino acids encoded by the codons upstream from the pause (Figure 1B). To this end, we use ribo-seq and mRNA-seq data to generate normalized profiles of *RD*/mRNA and extract peak positions in each normalized profile (Figure 1C). These positions presumably represent the positions along the mRNA where ribosomes have been significantly stalled (see details in the Methods). In the next step, we define for each peak the corresponding Upstream Stalling Region (*USR*) which is the sequence of amino acids encoded by the codons upstream from the peak. These amino acids occupy the *RET* while the codon at the peak position is being translated. Specifically, since the length of peptide required to fill the tunnel is approximately 31 amino acids [52], we have focused on the 31 amino acids before each peak (Figure 1D).

The folding of the nascent peptide inside the exit tunnel [28] and additional factors may alter the distance of a specific amino acid in the tunnel from the P-site during the translation process [21,53,54]; thus, we have decided to use measures that are based on the enrichment of different amino acids in the *USR* instead of constraining the amino acids to appear at a specific position relative to the P-site.

The organism-specific stalling effect of each amino acid

At the first step, we determined the enrichment of each amino acid in the *USRs* based on the following test: we calculated the probability to observe the amino acid in the real *USRs*; then, we calculated the probability to observe the amino acid in randomized ribosome profiling with similar properties as the original data (see details in the Methods section: *Quantifying the enrichment of each amino acid in the USRs*). Finally, based on the real and randomized ribosomal profiling data, we calculated a p-value which determines the extent to which each amino acid tends to occupy the *RET* while a codon at a highly stalled position is being translated (Figure 2A). For the bacteria, we performed one additional test (Figure 2B) to show that the reported results cannot be explained by the fact that hybridization between the prokaryotic ribosomal RNA and sequences that resemble the Shine-Dalgarno (SD) sequence can also cause pauses [50,55]; this phenomenon was controlled by filtering peaks that appear downstream from such sequences (see details in the Methods section: *Controlling for translational pausing driven by Shine-Dalgarno-like sequences*).

For each dataset, we classified each amino acid in one of three possible classifications based on the output of

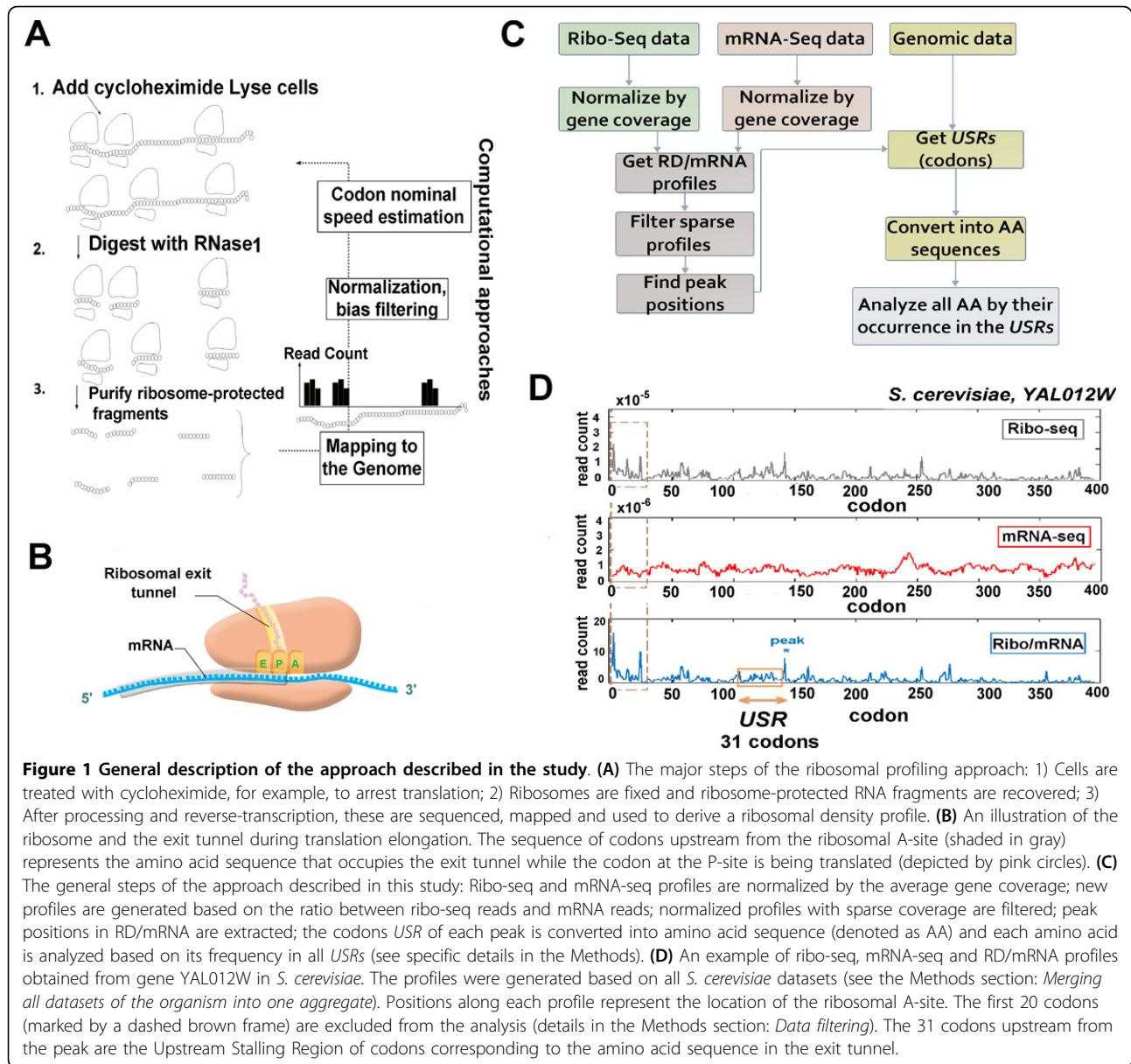


Figure 1 General description of the approach described in the study. **(A)** The major steps of the ribosomal profiling approach: 1) Cells are treated with cycloheximide, for example, to arrest translation; 2) Ribosomes are fixed and ribosome-protected RNA fragments are recovered; 3) After processing and reverse-transcription, these are sequenced, mapped and used to derive a ribosomal density profile. **(B)** An illustration of the ribosome and the exit tunnel during translation elongation. The sequence of codons upstream from the ribosomal A-site (shaded in gray) represents the amino acid sequence that occupies the exit tunnel while the codon at the P-site is being translated (depicted by pink circles). **(C)** The general steps of the approach described in this study: Ribo-seq and mRNA-seq profiles are normalized by the average gene coverage; new profiles are generated based on the ratio between ribo-seq reads and mRNA reads; normalized profiles with sparse coverage are filtered; peak positions in RD/mRNA are extracted; the codons *USR* of each peak is converted into amino acid sequence (denoted as AA) and each amino acid is analyzed based on its frequency in all *USRs* (see specific details in the Methods). **(D)** An example of ribo-seq, mRNA-seq and RD/mRNA profiles obtained from gene YAL012W in *S. cerevisiae*. The profiles were generated based on all *S. cerevisiae* datasets (see the Methods section: *Merging all datasets of the organism into one aggregate*). Positions along each profile represent the location of the ribosomal A-site. The first 20 codons (marked by a dashed brown frame) are excluded from the analysis (details in the Methods section: *Data filtering*). The 31 codons upstream from the peak are the Upstream Stalling Region of codons corresponding to the amino acid sequence in the exit tunnel.

the randomized *USR* test: If the test turned out to be significant, the amino acid was classified as ‘overrepresented’, meaning that the analysis supports the hypothesis that this amino acid tends to appear upstream of peaks more than expected by the null model (this may suggest that the amino acid contributes to the ribosomal stalling via its interaction with the tunnel). If the test turned out to be significant in the opposite direction (*i.e.*, the probability of observing the amino acid in *USRs* was significantly *smaller* than in random regions), the amino acid was classified as ‘underrepresented’. In case the test turned out to be insignificant at the 0.05 level, the amino acid was classified as ‘not significant’. The

classification of the 20 amino acids for each of the analyzed datasets appears in Figure 3.

As can be seen in Figure 3, our analysis suggests that the amino acids which significantly tend to occupy the exit tunnel when ribosomes stall, are organism (or condition)-specific.

Two remarkable stalling signals were produced by Proline (P) and Arginine (R) in *S. cerevisiae*, a finding that is well supported by the study of Artieri and Fraser [10]. In addition, our results suggest that Proline has also a stalling effect in more organisms including *S. pombe*, *D. melanogaster*, *H. sapiens* (G1 phase), *P. falciparum* (Late trophozoite) and *C. crescentus* (PYE). Another new

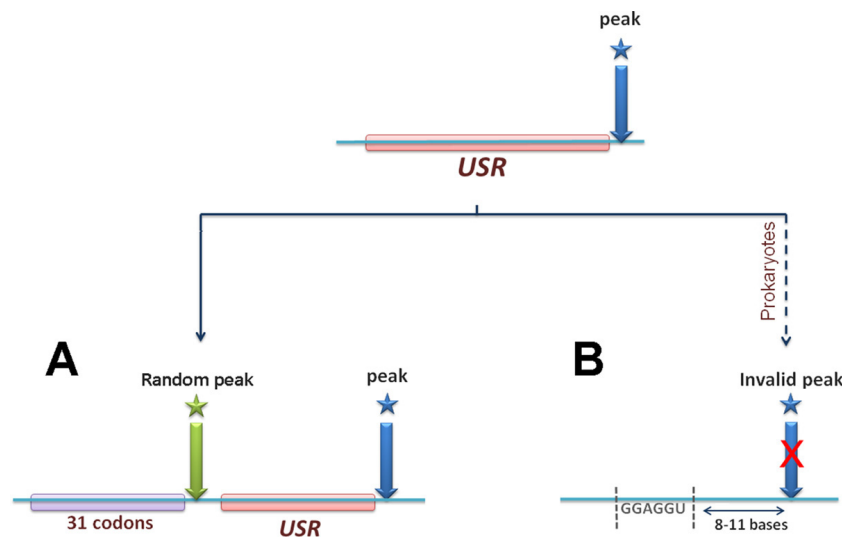


Figure 2 Tests for identifying stalling amino acids. The position of peak in the RD/mRNA profile is marked by a blue arrow. The USR corresponding to the peak is the upstream sequence of 31 codons which codes for the 31 amino acid that occupy the tunnel when ribosomes stall at the peak position. **(A)** In the first test, the frequency of each amino acid upstream of real peak positions is compared with its frequency in the 31 codons upstream of random positions. The number of randomly drawn positions per profile is equal to the number of real peaks in the original profile (see details in the Methods). **(B)** The additional test is performed for bacteria. In case a sequence that resembles the SD sequence was observed 8-11 bases upstream from the peak position, the peak was excluded from the analysis.

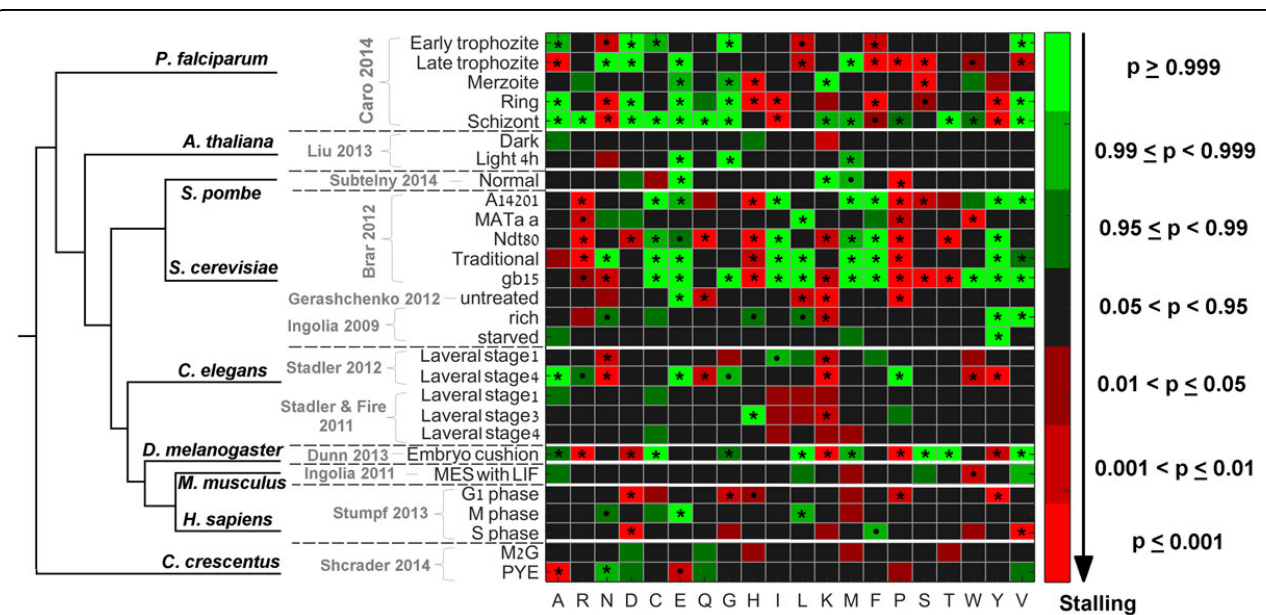


Figure 3 Dataset-specific classification of the 20 amino acids. Each amino acid was classified as significantly stalling (red), significantly non-stalling (green) or insignificant (black) according to the frequency of its codons in the USRs. All analyzed datasets are listed to the left. A color bar with the different significance levels is provided to the right. Stalling amino acids that passed FDR at the 0.05 level are marked with asterisk and those that passed FDR at the 0.1 level are marked by black dots. Thick horizontal white lines are plotted to separate the different organisms which are ordered in accordance with their evolutionary tree based on iTOL [76,77].

prominent stalling effect was observed for Lysine (K) in all datasets of *C.elegans*.

The negatively charged glutamic acid (E) was not found to be significantly stalling in any eukaryotic dataset. Moreover, it exhibited a significant signal of non-stalling in 6 of the 8 eukaryotes. Aspartic acid (D), the second negatively charged amino acid exhibits a stalling signal in specific datasets from 3 organisms (*S. cerevisiae*, *D. melanogaster* and *H. sapiens*).

Identifying the regions in the exit tunnel that tend to interact with the growing peptide

Since the biochemical, geometrical and electrostatic properties of the tunnel varies along its length, specific regions in the tunnel may have higher potential to induce interactions with certain amino acids. In order to identify such regions, we calculated the probability for each amino acid to occupy a specific position along the length of the exit tunnel when the ribosome stalls (*i.e.* upstream from peak positions). The resultant position-specific probabilities for each organism are presented in Figure 4.

As can be seen, for most of the amino acids, the positions with the most extreme probabilities tend to be concentrated in the part of the tunnel that is close to the P-site (~5 amino acids in length). This might suggest that in most cases this part of the tunnel tends to undergo more interactions with the translated amino acids than other positions in the tunnel. It can also be seen that proline, tends to specifically appear in the ribosomal P-site in 6 of the 9 organism, in line with previous studies [10-12].

Charged amino acids do not contribute to ribosome stalling in all organisms

Earlier studies have suggested that charged amino acids tend to interact with the exit tunnel and thus, contribute to ribosomal stalling [16,18,19,56]. Our analysis demonstrates that indeed in a few cases, the *USRs* tend to be enriched with charged amino acids (Figure 5). In order to understand whether the factor for the stalling is specifically the charge or other property of the amino acid, we tested the composite effect of charged amino acids on ribosome stalling. Similarly to the tests described in Figure 2, the frequency of occurrence of the charged amino acids was compared between real and randomized *USRs* (See details in the Methods section : *Quantifying the enrichment of charged amino acids in USRs*).

As can be seen in Figure 5, enrichment of positively charged amino acids among *USRs* was observed in 11 datasets from 3 eukaryotic organisms. Enrichment of *negatively* charged amino acids among *USRs* was less common and was observed only in 4 datasets from 3 organisms. On the contrary, in 5 organisms the probability to observe

negatively charged amino acid before peaks was significantly higher in random (Figure 5C); this may suggest that negatively charged amino acids can prevent the halting of the ribosome. We found no cases of greater enrichment in random for the positively charged amino acids.

Discussion

Our analysis identified nascent single amino acids that with high confidence contribute to ribosome stalling. The approach taken here to detect these amino acids is based on strict definitions and includes important controls on the analyzed genes such as control for amino acids bias and for possible experimental/protocol biases. In addition, we performed for the first time multi-organismal study of this topic which includes the analysis of both prokaryotes and eukaryotes.

The statistical tests performed here are based on the enrichment of amino acids *upstream* from the ribosomal P-site, thus, features such as mRNA folding strength which tends to slow ribosomes down when it occurs downstream from the P-site cannot trivially explain our results. In addition, previous studies (*e.g.* [18,39]) have suggested that the effect of rare codons on ribosome stalling tends to be less extreme than the effect of the interaction between the ribosomal exit tunnel and the nascent chain; thus, we also believe that the reported results cannot be trivially explained by the use of rare/non-efficient codons (see Additional file 1 for analysis supporting this point).

It is important to mention that currently the biases arise from the ribosome profiling approach and the effect of different protocols are not completely understood [57,58]. Much effort was spent here to consider these possible biases by 1) excluding from the analysis the first 20 codons which are known to be biased [37,44,57]; 2) filter low-coverage profiles; 3) normalizing each profile by its mean coverage to account for coverage differences [10] 4) normalizing ribo-seq data by mRNA-seq data to account for shared biases between the two fractions [10] 5) analyzing many datasets corresponding to a few different experimental conditions; 6) analyzing and comparing nine organisms (including eukaryotes and prokaryotes); 7) excluding pauses which might have been caused by SD sequences that hybridize with the prokaryotic ribosomal RNA [50,55]. Taken together, the reported results are based on a very conservative approach.

One of the major conclusions is related to the relation between positively charged amino acids and ribosome stalling. Previous studies have suggested that in *S. cerevisiae* positively charged amino acids play a role in ribosome stalling. Our analysis supports this conjecture in *S. cerevisiae* and also in specific datasets from *D. melanogaster* and *P. falciparum*. Therefore, our

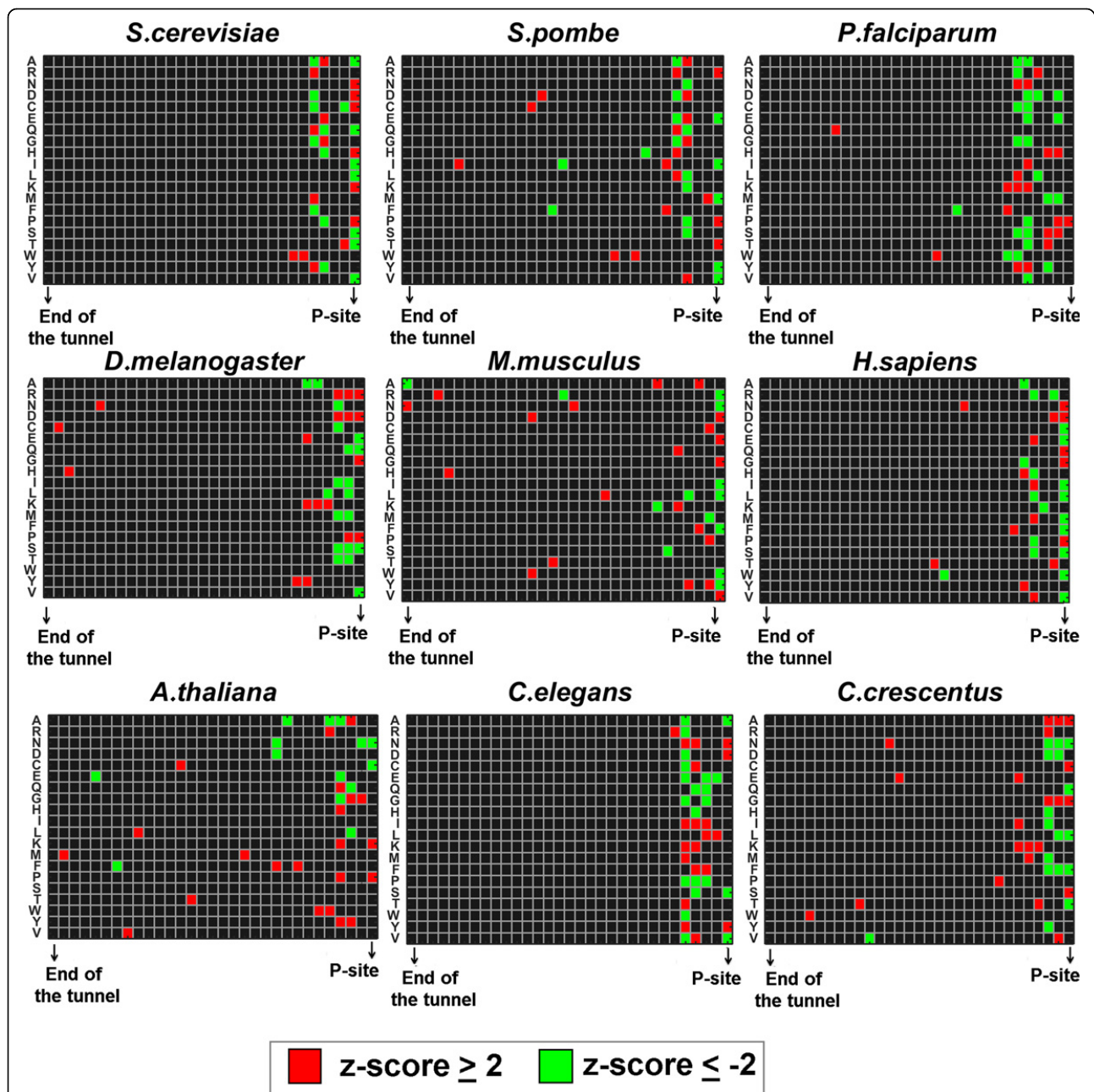


Figure 4 The distribution of amino acids along the tunnel when ribosomes stall. The position-specific probabilities were calculated for the 31 positions in the tunnel (based on the *USRs*). Results are presented per organism based on an aggregate that merges all analyzed datasets of the organism (see details in the Methods section: *Merging all datasets of the organism into one aggregate*). The probabilities were standardized to have a mean of zero per amino acid. We defined a square to be red/green if the probability to observe the amino acid in the corresponding position is significantly higher/lower than other positions in the tunnel.

study suggests that the relation between amino acids charge and ribosomal halting is not universal.

In addition, our analysis suggests that not only positively charged amino acids interact with the *RET*. Specifically, we show that negatively charged amino acids tend to halt the ribosome via interactions with the exit tunnel in *S. cerevisiae* (Ingolia *et al.* 2009, starved condition growth [37]),

D. melanogaster (Dunn *et al.* 2013, Embryos cushion [59]) and *H. sapiens* (Stumpf *et al.* 2013, G1 and S phase of HeLa cells [60]). Since the *RET* is negatively charged [19-21] it makes sense that it may undergo interactions with both positively and negatively charged amino acids. Furthermore, interestingly our analysis suggests that in some cases the negatively charged amino acids may

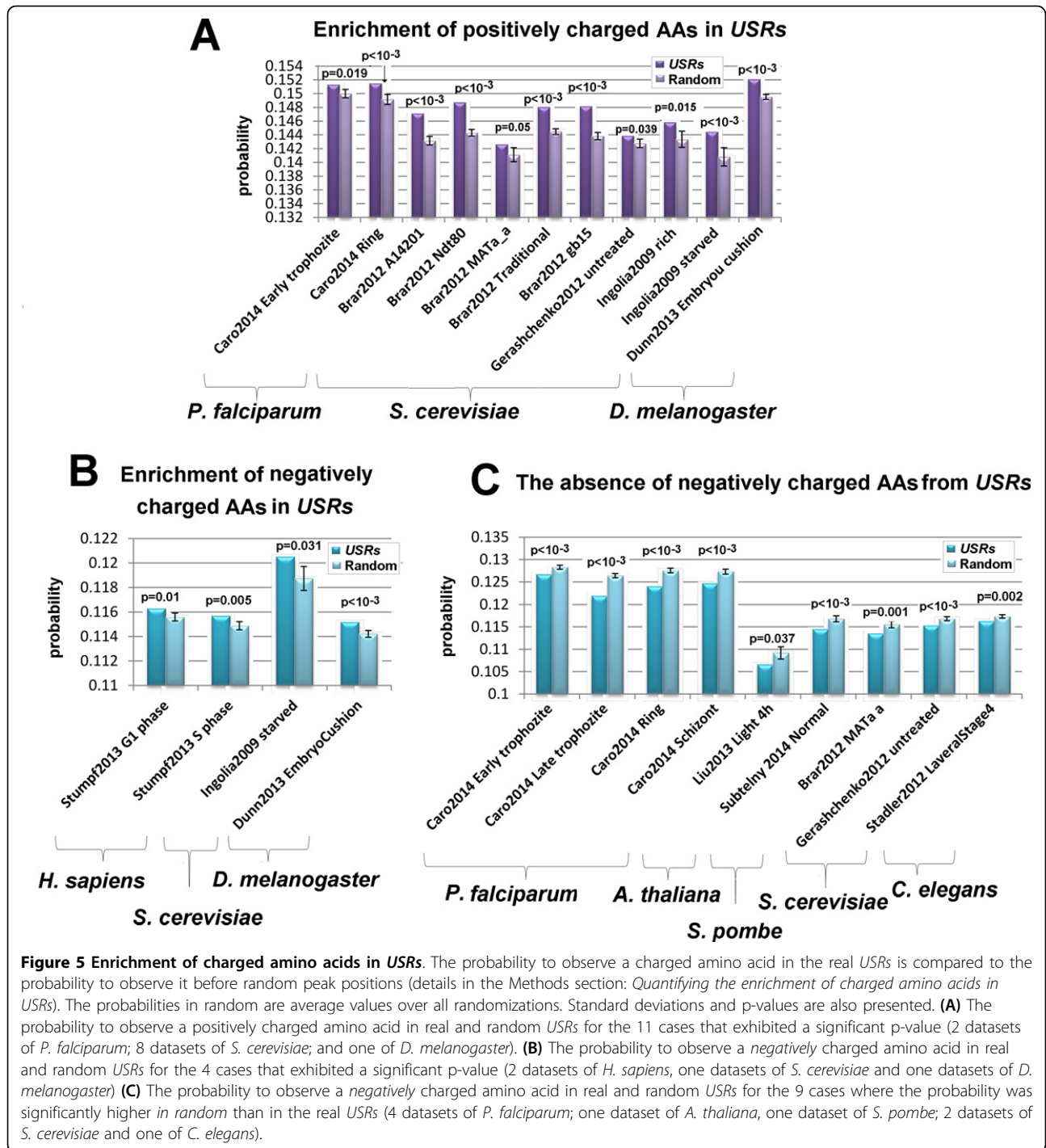


Figure 5 Enrichment of charged amino acids in USRs. The probability to observe a charged amino acid in the real USRs is compared to the probability to observe it before random peak positions (details in the Methods section: *Quantifying the enrichment of charged amino acids in USRs*). The probabilities in random are average values over all randomizations. Standard deviations and p-values are also presented. **(A)** The probability to observe a positively charged amino acid in real and random USRs for the 11 cases that exhibited a significant p-value (2 datasets of *P. falciparum*; 8 datasets of *S. cerevisiae*; and one of *D. melanogaster*). **(B)** The probability to observe a negatively charged amino acid in real and random USRs for the 4 cases that exhibited a significant p-value (2 datasets of *H. sapiens*, one datasets of *S. cerevisiae* and one datasets of *D. melanogaster*) **(C)** The probability to observe a negatively charged amino acid in real and random USRs for the 9 cases where the probability was significantly higher in random than in the real USRs (4 datasets of *P. falciparum*; one dataset of *A. thaliana*, one dataset of *S. pombe*; 2 datasets of *S. cerevisiae* and one of *C. elegans*).

prevent stalling; this may be related to charge cancellation with possible positively charged amino acids that co-appear in proximity in the exit tunnel.

Although we discuss the stalling effect of each amino acid on ribosome stalling, we do not claim that the stalling is manipulated by a specific mechanism. In fact, the explanation regarding the exact type of interaction between these amino acids and the ribosome and the

reason they differ across the tree of life is an open question for future studies.

The reported results support the conjecture that the amino acids composition of the nascent peptide affects the ribosomal translation speed and might even cause ribosomal arrest. Thus, this finding suggests a complex interaction between the protein co-translational folding, protein amino acid content and ribosomal elongation

speed: the translated amino acids affect translation speed which may affect protein folding. Thus, we believe that there is a co-evolution among these variables.

The fact that different stalling amino acids were reported for the different analyzed organisms may suggest that the biochemical properties of the exit tunnel vary along the tree of life and/or in different conditions [61-64]. This finding also provides important insights about heterologous gene expression: the expression of the same protein in different organisms may affect its translation rate simply due to the different nature of the interactions between the protein amino acids and the ribosomal exit tunnels in new organisms. This fact can explain why the topic of heterologous gene expression is often very challenging and why synonymous manipulation on the protein alone is not always sufficient for solving problems in this field.

Finally, as a future research it would be interesting to generalize the results reported here by estimating the effect of short peptides and sets of amino acids (not necessarily neighbor amino acids) in the *RET* on ribosomal halting. For example, since stalling peptides interfere with translation, they are expected to be selected against to improve translational efficiency. Thus, it would be interesting to examine the relation between the stalling effect of these peptides and their representation in the proteome. However, this mission is statistically challenging due to the exponential increase in the number of sets of amino acids compositions with more than one amino acid.

Conclusions

In this work, we performed a multiple ribosome profiling datasets analysis to understand the effect of different amino acid on ribosome stalling. The reported results support a few conjectures: various amino acids interact with the ribosomal exit tunnel; the nature of these interactions is organism/condition specific and the nascent translated peptide tends to have more interactions with the beginning of the exit tunnel (close to the p-site).

Methods

Coding Sequences Data

Coding sequences of all analyzed organisms were retrieved from the UCSC genome browser (<http://genome-euro.ucsc.edu>).

Ribo-seq and mRNA-seq data

Ribo-seq and mRNA-seq data used in this study are based on the following experiments: Ingolia *et al.* 2009 [37], Brar *et al.* 2012 [46] and Gerashchenko *et al.* 2012 [43] for *S. cerevisiae*; Subtelny *et al.* 2014 [65] for *S. pombe*; Stadler and Fire 2011 [66] and Stadler *et al.* 2012 [67] for *C. elegans*; Ingolia *et al.* 2011 [40] for *M. musculus*; Stumpf *et al.* 2013 [60] for *H. sapiens*; Dunn *et al.* 2013 [59] for

D. melanogaster; Caro *et al.* 2014 [68] for *P. falciparum*; Liu *et al.* 2013 [69] for *A. thaliana* and Schrader *et al.* 2014 [70] for *C. crescentus*. Ribosomal footprint reads of each experiment have been uniquely mapped to the corresponding genome by Michel *et al.* 2014 [71] and were retrieved from the GWIPS-viz database (<http://gwips.ucc.ie>).

Mapping ribosomal footprints to genomic positions

The specific genomic position assigned to each read represents the location of the ribosomal A-site on the mRNA. In GWIPS-viz, the genomic coordinate of each read has been determined differently for eukaryotes and prokaryotes. For the eukaryotic fragments, in which the 5' end of the footprint is sufficient to carry the positional information [37,40], an off-set of 15 nucleotides from the 5' end of the fragment was used. Prokaryotic fragments, in contrast, varied between 25 and 40 nucleotides in length, mostly as a result of the specificity of micrococcal nuclease and thus, a weighted centered approach implemented by Oh *et al.*, 2011 [47] was used to indicate the putative location of the ribosomal A-site. Specifically, 12 nucleotides were trimmed from each end of the prokaryotic fragment and the remaining residues were given a score of $1/N$, where N equals the number of positions leftover after discarding the 5' and 3' ends, and blurring the signal across the central residues.

Data filtering

The density of ribosome footprints is significantly elevated in the beginning of the gene due to a combination of biological phenomena and biases [37,40,44,72]. Thus, the first 20 codons were excluded from all aspects of the analysis described in this study. In addition, to account for biases related to sparse coverage, genes' profiles with fewer than 40 percent non-zero read counts were further filtered.

RD peaks definition

While ribosome profiling data is given at nucleotide resolution, our analysis is based on codons. Thus, we averaged the read counts at each three bases corresponding to codons triplets to get the density profile at codon resolution. Then, in order to define peak positions in a given profile, we calculated the average read count (excluding zeros and the first 20 codons) and consider positions that exceed the average by 4 standard deviations as peaks. This definition was chosen empirically by ensuring that the total *USRs* in a protein sequence will cover at the most 20 percent of its length (otherwise it is not possible to randomize these sequences).

Accounting for biases in mRNA-seq data and coverage differences

Theoretically, mRNA-seq read counts along a specific transcript should be uniformly distributed. In practice,

due to various biases, this is not the case and the read counts obtained by mRNA-seq differ along the transcript (yet with small deviation from the average read count of the transcript relatively to ribo-seq read counts).

Recently Artieri and Fraser [10], established a robust methodology to account for such biases which includes the normalization of ribo-seq data by mRNA-seq data. Similarly to their approach, we first scaled each profile (ribo and mRNA) by the gene coverage and then, calculated the RD/mRNA ratio for each codon.

We compared the results obtained by the normalized data (*i.e.* the RD/mRNA ratio) with those obtained by the RD data without the normalization. The amino acids classifications based on the non-normalized data (Additional File 2), clearly produces more false signals compared to those obtained by the normalized data (Figure 3).

The robustness of the reported results to a stricter threshold of coverage data

A general problem in large scale analyses is finding a work point where the signal to noise ratio is optimal. Specifically, in the case of ribosome profiling data, there is a tradeoff between a high-coverage demand (which is necessary for understanding the global behavior across the transcript) and a large number of genes (which strengthens the statistical power of the reported signal). In our case, we analyzed only genes with at least 40% non-zeros read counts (Methods). In order to benchmark this definition, we compared the results obtained by this definition against a stricter one (for example a threshold of 60% non-zero read counts in each profile). We observed no cases where amino acid changed the direction of significance, demonstrating the robustness of the reported results. The classified amino acids for the stricter threshold are presented in Additional File 3.

The effect of rare codons on the reported results

In order to show that the reported results cannot be trivially explained by the use of rare codons, we calculated Spearman's rank correlation coefficient between the probability that each codon occupies the P-site at peak positions and its corresponding tRNA adaptation index (*tAI*). The *tAI* is a widely used measure of the adaptation of codons to the tRNA pool of the organism [2], thus, it provides an information regarding the nominal translation rate of codons. The tRNA gene copy numbers used for the *tAI* calculation were retrieved from the Genomic tRNA database [73] (<http://gtrnadb.ucsc.edu>). The correlations for each of the nine analyzed organisms are presented in Additional File 1.

Merging all datasets of the organism into one aggregate

An aggregate that is based on all analyzed datasets of the organism was generated in two steps: First, to cancel

the effect of different coverage between datasets, we sum the read counts over all profiles to get the total number of read counts in the experiment; then, we normalized the read counts in each dataset by its total number of reads. Second, we averaged the normalized profiles of each gene to get the final aggregate. This was done for both, the ribo-seq and the mRNA-seq data. Finally, the normalization by gene coverage and mRNA-seq discussed in the previous section was performed on the aggregate dataset. Aggregate datasets were used to generate Figure 1D, 5 and Additional File 1.

Controlling for translational pausing driven by Shine-Dalgarno-like sequences

It is known that in bacteria hexanucleotide sequences that resemble Shine-Dalgarno (*SD*) features within coding sequences can cause translational pausing due to hybridization between the mRNA and the 16S ribosomal RNA of the ribosome [50]. We have defined a *SD* sequence as a hexanucleotide sequence which contains up to one substitution relative to the canonical *SD* (GGAGGU). Specifically, the optimal spacing between the 3' end of the anti-*SD* sequence and the ribosomal A-site is 8-11 nucleotides [74]. Thus, we have excluded peaks that contain a *SD* sequence 8, 9, or 11 nucleotides upstream from the peak position.

Quantifying the enrichment of charged amino acids in USRs

In order to understand whether positively charged amino acids (Lysine, arginine and histidine) tend to stall the ribosome via interaction with the *RET*, we quantified their tendency to appear before RD peaks (*i.e.*, in *USRs*). Traversing all peaks in all genes, we gave each peak a binary score: +1, if at least one positively charged amino acid (any of the three) appears in the *USR* of the peak and 0 if none of the three is observed. Finally, we summed up all peaks to get the total score (statistics) of the positively charged amino acids for the entire *USRs* ($SCORE_{pos_AA,USRs}$); then, these value were normalized by the number of peaks to get the empirical probability. In order to quantify the significance of the score, we generated a null model by randomly draw the positions of the peaks maintaining the number of peaks in each gene. Random *USRs* are equivalently the 31 amino acids sequences upstream from each random peak position. The score of the positively charged amino acids ($SCORE_{pos_AA,random}$) was calculated based on the random peaks. The process was repeated 1000 times.

The empirical p-value that determines the extent to which the frequency of occurrence of positively charged amino acids is higher in real *USRs* than in random was calculated by:

$$p_{pos_AA,USRs} = \frac{\text{number of times}(score_{pos_AA,random} \geq score_{pos_AA,USRs})}{1000}$$

$p_{pos_AA,USRs} < 0.05$ indicates a significant enrichment of positively charged amino acids in *USRs*.

Similarly, we defined a p-value for the enrichment of negatively charged amino acids (Glutamate or aspartic acid) in *USRs*. A score of +1 was given to a peak if at least one negatively charged amino acid (any of the two) appears in the *USR* of the peak. The empirical p-value was calculated by:

$$p_{neg_AA,USRs} = \frac{\text{number of times}(score_{negAA,random} \geq score_{negAA,USRs})}{1000}$$

For each type of charge, the scores do not change if both positive/negative charge appear in the *USR* (e.g., if we perform the test for the negatively charged amino acid, we will give a score of +1 to every *USR* which included any negatively charged amino acid even if a positively charged amino acid also appears). Allowing both types of charge is based on the following rationale: First, the statistical power of the test is higher (since we do not omit *USRs*). Second, we believe that the interactions between positive or negative amino acid and the exit tunnel occur in certain region of the exit tunnel; since these regions may change in different conditions/organisms or during the translation of a certain mRNA, a positively/negatively charged amino acid may affect the ribosomal movement even if there is an additional positively/negatively charged amino acid in the *USR*. Finally, this definition is more conservative since the calculated p-value might be higher (less significant) in cases of charge cancellation (i.e., when the positive and negative charge in the *USR* cancel each other effect).

Quantifying the enrichment of each amino acid in the *USRs*

Similarly to the approach described in the previous paragraph, we have quantified the tendency of each single amino acid to stall the ribosome based on its occurrence in the *USRs*. For each amino acid, we traverse all peaks in all genes and assign each peak a binary score: +1 if the amino acid in question appears in the *USR* corresponding to the peak and 0 if it is not. Finally, we sum over all peaks to get $score_{AA(i),USRs}$, the total score of the amino acid for the entire *USRs* of the proteome. For the null model we randomly draw the positions of the peaks, while maintaining the number of peaks in each profile to be identical to the actual profile, and calculate $score_{AA(i),random}$ to each amino acid. The process was repeated 1000 times.

The p-value for the *i*-th amino acid is defined by:

$$p_{AA(i),USRs} = \frac{\text{number of times}(score_{AA(i),random} \geq score_{AA(i),USRs})}{1000}$$

$p_{AA(i),USRs} < 0.05$ indicates a significant enrichment of the *i*-th amino acid in the *USRs*. To control the False Discovery Rate (FDR), we performed a multiple testing correction on the resultant p-values (based on the Benjamini-Hochberg procedure [75]).

Additional material

Additional File 1: The correlation between tAI and P-site occupation probability at peak positions.

The results are presented per organism based on an aggregate that merges all analyzed datasets of the organism (see details in the Methods section: *Merging all datasets of the organism into one aggregate*). The probability at the x-axis represents the probability that each of the 61 sense codons occupies the P-site at peak positions. Spearman's rank correlation coefficient (ρ) and a corresponding p-value (p) are to the upper right hand corner of each figure.

Additional File 2: Amino acids classifications based on ribo-seq data only.

The figure is based on ribo-seq profiles which do not include the normalization by mRNA-seq data. Each amino acid was classified as significantly stalling (red), significantly non-stalling (green) or insignificant (black) according to the frequency of its codons in the *USRs*. Stalling amino acids that passed FDR at the 0.05 level are marked with asterisk and those that passed FDR at the 0.1 level are marked by black dots. All analyzed datasets are listed to the left. Thick horizontal white lines are plotted to separate the different organisms. A color bar with the different significance levels is provided to the right.

Additional File 3: The results of a stricter threshold for the sparse data filtering.

The figure is based on RD/mRNA profiles with at least 60% non-zero read counts (see details in the Methods section: *The robustness of the reported results to a stricter threshold of coverage data*). Each amino acid was classified as significantly stalling (red), significantly non-stalling (green) or insignificant (black) according to the frequency of its codons in the *USRs*. Stalling amino acids that passed FDR at the 0.05 level are marked with asterisk and those that passed FDR at the 0.1 level are marked by black dots. All analyzed datasets are listed to the left. Thick horizontal white lines are plotted to separate the different organisms. A color bar with the different significance levels is provided to the right.

Abbreviations

RET: Ribosomal Exit Tunnel; RD: Ribosomal Density; USR: Upstream Stalling Region; AA: Amino Acid; SD: Shine-Dalgarno; FDR: False Discovery Rate; tAI: tRNA Adaptation Index.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

Conceived and designed the experiments: RS TT. Analyzed the data: RS TT. Wrote the paper: RS TT.

Acknowledgements

This study was supported in part by a fellowship from the Edmond J. Safra Center for Bioinformatics at Tel-Aviv University.

Declarations

The publication costs were funded by Tel Aviv University resources. This article has been published as part of *BMC Genomics* Volume 16 Supplement 10, 2015: Proceedings of the 13th Annual Research in Computational Molecular Biology (RECOMB) Satellite Workshop on Comparative Genomics: Genomics. The full contents of the supplement are available online at <http://www.biomedcentral.com/bmcgenomics/supplements/16/S10>.

Authors' details

¹Department of Biomedical Engineering, Tel Aviv University (TAU), Tel Aviv, Israel. ²The Sagol School of Neuroscience, Tel-Aviv University (TAU), Tel-Aviv, Israel.

Published: 2 October 2015

References

1. Kanaya S, Yamada Y, Kinouchi M, Kudo Y, Ikemura T: **Codon usage and tRNA genes in eukaryotes: correlation of codon usage diversity with translation efficiency and with CG-dinucleotide usage as assessed by multivariate analysis.** *Journal of molecular evolution* 2001, **53**(4-5):290-298.
2. dos Reis M, Savva R, Wernisch L: **Solving the riddle of codon usage preferences: a test for translational selection.** *Nucleic acids research* 2004, **32**(17):5036-5044.
3. Gustafsson C, Govindarajan S, Minshull J: **Codon bias and heterologous protein expression.** *Trends in biotechnology* 2004, **22**(7):346-353.
4. Sørensen MA, Kurland C, Pedersen S: **Codon usage determines translation rate in *Escherichia coli*.** *Journal of molecular biology* 1989, **207**(2):365-377.
5. Varenne S, Buc J, Llobes R, Lazdunski C: **Translation is a non-uniform process: effect of tRNA availability on the rate of elongation of nascent polypeptide chains.** *Journal of molecular biology* 1984, **180**(3):549-576.
6. Dana A, Tuller T: **The effect of tRNA levels on decoding times of mRNA codons.** *Nucleic acids research* 2014, **42**(14):9171-9181.
7. Kato M, Nishikawa K, Uritani M, Miyazaki M, Takemura S: **The difference in the type of codon-anticodon base pairing at the ribosomal P-site is one of the determinants of the translational rate.** *Journal of biochemistry* 1990, **107**(2):242-247.
8. Thomas LK, Dix DB, Thompson RC: **Codon choice and gene expression: synonymous codons differ in their ability to direct aminoacylated-transfer RNA binding to ribosomes in vitro.** *Proceedings of the National Academy of Sciences* 1988, **85**(12):4242-4246.
9. Sabi R, Tuller T: **Modelling the Efficiency of Codon-tRNA Interactions Based on Codon Usage Bias.** *DNA Research* 2014, **21**(5):511-526.
10. Artieri CG, Fraser HB: **Accounting for biases in riboprofiling data indicates a major role for proline in stalling translation.** *Genome research* 2014, gr. 175893.175114.
11. Pavlov MY, Watts RE, Tan Z, Cornish VW, Ehrenberg M, Forster AC: **Slow peptide bond formation by proline and other N-alkylamino acids in translation.** *Proceedings of the National Academy of Sciences* 2009, **106**(1):50-54.
12. Muto H, Ito K: **Peptidyl-prolyl-tRNA at the ribosomal P-site reacts poorly with puromycin.** *Biochemical and biophysical research communications* 2008, **366**(4):1043-1047.
13. Tuller T, Waldman YY, Kupiec M, Ruppin E: **Translation efficiency is determined by both codon bias and folding energy.** *Proceedings of the National Academy of Sciences* 2010, **107**(8):3645-3650.
14. Nackley A, Shabalina S, Tchivileva I, Satterfield K, Korchyński O, Makarov S, Maixner W, Diatchenko L: **Human catechol-O-methyltransferase haplotypes modulate protein expression by altering mRNA secondary structure.** *Science* 2006, **314**(5807):1930-1933.
15. Hall MN, Gabay J, Débarbouillé M, Schwartz M: **A role for mRNA secondary structure in the control of translation initiation** 1982.
16. Tuller T, Veksler-Lublinsky I, Gazit N, Kupiec M, Ruppin E, Ziv-Ukelson M: **Composite effects of gene determinants on the translation speed and density of ribosomes.** *Genome Biol* 2011, **12**(11):R110.
17. Johansson M, leong K-W, Trobro S, Strazewski P, Åqvist J, Pavlov MY, Ehrenberg M: **pH-sensitivity of the ribosomal peptidyl transfer reaction dependent on the identity of the A-site aminoacyl-tRNA.** *Proceedings of the National Academy of Sciences* 2011, **108**(1):79-84.
18. Charneski CA, Hurst LD: **Positively charged residues are the major determinants of ribosomal velocity.** *PLoS biology* 2013, **11**(3):e1001508.
19. Lu J, Deutsch C: **Electrostatics in the ribosomal tunnel modulate chain elongation rates.** *Journal of molecular biology* 2008, **384**(1):73-86.
20. Nakatogawa H, Ito K: **The ribosomal exit tunnel functions as a discriminating gate.** *Cell* 2002, **108**(5):629-636.
21. Berisio R, Schluenzen F, Harms J, Bashan A, Auerbach T, Baram D, Yonath A: **Structural insight into the role of the ribosomal tunnel in cellular regulation.** *Nature Structural & Molecular Biology* 2003, **10**(5):366-370.
22. Woolhead CA, McCormick PJ, Johnson AE: **Nascent membrane and secretory proteins differ in FRET-detected folding far inside the ribosome and in their exposure to ribosomal proteins.** *Cell* 2004, **116**(5):725-736.
23. Woolhead CA, Johnson AE, Bernstein HD: **Translation arrest requires two-way communication between a nascent polypeptide and the ribosome.** *Molecular cell* 2006, **22**(5):587-598.
24. Ban N, Nissen P, Hansen J, Moore PB, Steitz TA: **The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution.** *Science* 2000, **289**(5481):905-920.
25. Nissen P, Hansen J, Ban N, Moore PB, Steitz TA: **The structural basis of ribosome activity in peptide bond synthesis.** *Science* 2000, **289**(5481):920-930.
26. Ménétret J-F, Neuhof A, Morgan DG, Plath K, Radermacher M, Rapoport TA, Akey CW: **The structure of ribosome-channel complexes engaged in protein translocation.** *Molecular cell* 2000, **6**(5):1219-1232.
27. Beckmann R, Spahn CM, Eswar N, Helmers J, Penczek PA, Sali A, Frank J, Blobel G: **Architecture of the protein-conducting channel associated with the translating 80S ribosome.** *Cell* 2001, **107**(3):361-372.
28. Kaiser C, Goldman D, Tinoco I, Bustamante C: **The Ribosome Modulates Nascent Protein Folding.** *Biophysical Journal* 2012, **102**(3):68a.
29. Kosolapov A, Deutsch C: **Tertiary interactions within the ribosomal exit tunnel.** *Nature structural & molecular biology* 2009, **16**(4):405-411.
30. Ramu H, Mankin A, Vazquez-Laslop N: **Programmed drug-dependent ribosome stalling.** *Molecular microbiology* 2009, **71**(4):811-824.
31. Yanagitani K, Kimata Y, Kadokura H, Kohno K: **Translational pausing ensures membrane targeting and cytoplasmic splicing of XBP1u mRNA.** *Science* 2011, **331**(6017):586-589.
32. Wei J, Wu C, Sachs MS: **The arginine attenuator peptide interferes with the ribosome peptidyl transferase center.** *Molecular and cellular biology* 2012, **32**(13):2396-2406.
33. Ito K, Chiba S, Pogliano K: **Divergent stalling sequences sense and control cellular physiology.** *Biochemical and biophysical research communications* 2010, **393**(1):1-5.
34. Woolstenhulme CJ, Parajuli S, Healey DW, Valverde DP, Petersen EN, Starosta AL, Guydosh NR, Johnson WE, Wilson DN, Buskirk AR: **Nascent peptides that block protein synthesis in bacteria.** *Proceedings of the National Academy of Sciences* 2013, **110**(10):E878-E887.
35. Tenson T, Ehrenberg M: **Regulatory nascent peptides in the ribosomal tunnel.** *Cell* 2002, **108**(5):591-594.
36. Lovett PS, Rogers EJ: **Ribosome regulation by the nascent peptide.** *Microbiological reviews* 1996, **60**(2):366-385.
37. 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):218-223.
38. Dana A, Tuller T: **Determinants of translation elongation speed and ribosomal profiling biases in mouse embryonic stem cells** 2012.
39. Dana A, Tuller T: **The effect of tRNA levels on decoding times of mRNA codons.** *Nucleic acids research* 2014, **42**(14):9171-9181.
40. Ingolia NT, Lareau LF, Weissman JS: **Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes.** *Cell* 2011, **147**(4):789-802.
41. Ingolia NT: **Ribosome profiling: new views of translation, from single codons to genome scale.** *Nature Reviews Genetics* 2014, **15**(3):205-213.
42. Zur H, Tuller T: **Strong association between mRNA folding strength and protein abundance in *S. cerevisiae*.** *EMBO reports* 2012, **13**(3):272-277.
43. Gerashchenko MV, Lobanov AV, Gladyshev VN: **Genome-wide ribosome profiling reveals complex translational regulation in response to oxidative stress.** *Proceedings of the National Academy of Sciences* 2012, **109**(43):17394-17399.
44. Tuller T, Carmi A, Vestsigian K, Navon S, Dorfan Y, Zaborske J, Pan T, Dahan O, Furman I, Pilpel Y: **An evolutionarily conserved mechanism for controlling the efficiency of protein translation.** *Cell* 2010, **141**(2):344-354.
45. Bazzini AA, Lee MT, Giraldez AJ: **Ribosome profiling shows that miR-430 reduces translation before causing mRNA decay in zebrafish.** *Science* 2012, **336**(6078):233-237.
46. 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):552-557.
47. Oh E, Becker AH, Sandkci A, Huber D, Chaba R, Gloge F, Nichols RJ, Typas A, Gross CA, Kramer G: **Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo.** *Cell* 2011, **147**(6):1295-1308.
48. 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 reports* 2014, **8**(5):1365-1379.
49. Ingolia NT, Brar GA, Rouskin S, McGeachy AM, Weissman JS: **Genome-Wide Annotation and Quantitation of Translation by Ribosome Profiling.** *Current Protocols in Molecular Biology* 2013, **4**(18):11-14, 18. 19.
50. Li G-W, Oh E, Weissman JS: **The anti-Shine-Dalgarno sequence drives translational pausing and codon choice in bacteria.** *Nature* 2012, **484**(7395):538-541.

51. Juntawong P, Girke T, Bazin J, Bailey-Serres J: **Translational dynamics revealed by genome-wide profiling of ribosome footprints in Arabidopsis.** *Proceedings of the National Academy of Sciences* 2014, **111**(1):E203-E212.
52. Lu J, Deutsch C: **Secondary structure formation of a transmembrane segment in Kv channels.** *Biochemistry* 2005, **44**(23):8230-8243.
53. Gabashvili IS, Gregory ST, Valle M, Grassucci R, Worbs M, Wahl MC, Dahlberg AE, Frank J: **The polypeptide tunnel system in the ribosome and its gating in erythromycin resistance mutants of L4 and L22.** *Molecular cell* 2001, **8**(1):181-188.
54. Wu C, Wei J, Lin P-J, Tu L, Deutsch C, Johnson AE, Sachs MS: **Arginine changes the conformation of the arginine attenuator peptide relative to the ribosome tunnel.** *Journal of molecular biology* 2012, **416**(4):518-533.
55. Shine J, Dalgarno L: **The 3'-terminal sequence of Escherichia coli 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites.** *Proceedings of the National Academy of Sciences* 1974, **71**(4):1342-1346.
56. Lu J, Kobertz WR, Deutsch C: **Mapping the electrostatic potential within the ribosomal exit tunnel.** *Journal of molecular biology* 2007, **371**(5):1378-1391.
57. Dana A, Tuller T: **Determinants of translation elongation speed and ribosomal profiling biases in mouse embryonic stem cells.** *PLoS computational biology* 2012, **8**(11):e1002755.
58. Gerashchenko MV, Gladyshev VN: **Translation inhibitors cause abnormalities in ribosome profiling experiments.** *Nucleic acids research* 2014, gku671.
59. Dunn JG, Foo CK, Belletier NG, Gavis ER, Weissman JS: **Ribosome profiling reveals pervasive and regulated stop codon readthrough in Drosophila melanogaster.** *Elife* 2013, **2**.
60. Stumpf CR, Moreno MV, Olshen AB, Taylor BS, Ruggero D: **The translational landscape of the mammalian cell cycle.** *Molecular cell* 2013, **52**(4):574-582.
61. Spahn CM, Beckmann R, Eswar N, Penczek PA, Sali A, Blobel G, Frank J: **Structure of the 80S ribosome from Saccharomyces cerevisiae-tRNA-ribosome and subunit-subunit interactions.** *Cell* 2001, **107**(3):373-386.
62. Gerbi S: **Expansion segments: regions of variable size that interrupt the universal core secondary structure of ribosomal RNA.** *Ribosomal RNA structure, evolution, processing, and function in protein biosynthesis* 1996, 71-87.
63. Gomez-Lorenzo MG, Spahn CM, Agrawal RK, Grassucci RA, Penczek P, Chakraburty K, Ballesta JP, Lavandera JL, Garcia-Bustos JF, Frank J: **Three-dimensional cryo-electron microscopy localization of EF2 in the Saccharomyces cerevisiae 80S ribosome at 17.5 Å resolution.** *The EMBO journal* 2000, **19**(11):2710-2718.
64. Bailey-Serres J, Freeling M: **Hypoxic stress-induced changes in ribosomes of maize seedling roots.** *Plant Physiology* 1990, **94**(3):1237-1243.
65. Subtelny AO, Eichhorn SW, Chen GR, Sive H, Bartel DP: **Poly (A)-tail profiling reveals an embryonic switch in translational control.** *Nature* 2014, **508**(7494):66-71.
66. Stadler M, Fire A: **Wobble base-pairing slows in vivo translation elongation in metazoans.** *Rna* 2011, **17**(12):2063-2073.
67. Stadler M, Artilles K, Pak J, Fire A: **Contributions of mRNA abundance, ribosome loading, and post-or peri-translational effects to temporal repression of C. elegans heterochronic miRNA targets.** *Genome research* 2012, **22**(12):2418-2426.
68. Caro F, Ahnyong V, Betegon M, DeRisi JL: **Genome-wide regulatory dynamics of translation in the Plasmodium falciparum asexual blood stages.** *Elife* 2014, **3**:e04106.
69. Liu M-J, Wu S-H, Wu J-F, Lin W-D, Wu Y-C, Tsai T-Y, Tsai H-L, Wu S-H: **Translational landscape of photomorphogenic Arabidopsis.** *The Plant Cell* 2013, **25**(10):3699-3710.
70. Schrader JM, Zhou B, Li G-W, Lasker K, Childers WS, Williams B, Long T, Crosson S, McAdams HH, Weissman JS: **The coding and noncoding architecture of the Caulobacter crescentus genome.** 2014.
71. Michel AM, Fox G, Kiran AM, De Bo C, O'Connor PB, Heaphy SM, Mullan JP, Donohue CA, Higgins DG, Baranov PV: **GWIPS-viz: development of a ribo-seq genome browser.** *Nucleic acids research* 2014, **42**(D1):D859-D864.
72. Tuller T, Zur H: **Multiple roles of the coding sequence 5' end in gene expression regulation.** *Nucleic acids research* 2015, **43**(1):13-28.
73. Chan PP, Lowe TM: **GtRNAdb: a database of transfer RNA genes detected in genomic sequence.** *Nucleic acids research* 2009, **37**(suppl 1):D93-D97.
74. Chen H, Bjercknes M, Kumar R, Jay E: **Determination of the optimal aligned spacing between the Shine-Dalgarno sequence and the translation initiation codon of Escherichia coli mRNAs.** *Nucleic acids research* 1994, **22**(23):4953-4957.
75. Benjamini Y, Hochberg Y: **Controlling the false discovery rate: a practical and powerful approach to multiple testing.** *Journal of the Royal Statistical Society Series B (Methodological)* 1995, 289-300.
76. Letunic I, Bork P: **Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy.** *Nucleic acids research* 2011, gkr201.
77. Letunic I, Bork P: **Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation.** *Bioinformatics* 2007, **23**(1):127-128.

doi:10.1186/1471-2164-16-S10-S5

Cite this article as: Sabi and Tuller: A comparative genomics study on the effect of individual amino acids on ribosome stalling. *BMC Genomics* 2015 **16**(Suppl 10):S5.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

