Strong association between mRNA folding strength and protein abundance in *S. cerevisiae*

Hadas Zur[1] & Tamir Tuller[2+]

[1]School of Computer Science, and [2]Faculty of Engineering, Department of Biomedical Engineering, Tel Aviv University, Ramat Aviv, Israel

One of the open questions in regulatory genomics is how the efficiency of gene translation is encoded in the coding sequence. Here we analyse recently generated measurements of folding energy in *Saccharomyces cerevisiae*, showing that genes with high protein abundance tend to have strong mRNA folding (mF; $R = 0.68$). mF strength also strongly correlates with ribosomal density and mRNA levels, suggesting that this relation at least partially pertains to the efficiency of translation elongation, presumably by preventing aggregation of mRNA molecules.

INTRODUCTION

Understanding gene expression, and specifically how the efficiency of this process is correlated or encoded in the coding regions and untranslated regions, has been the topic of dozens of papers in recent years [1–5].

The abundance level of a protein is related to its mRNA levels, its translation rate and its degradation rate. Specifically, if we assume constant mRNA levels, the translation rate should have a positive effect on the protein abundance (PA), while the degradation rate should have a negative effect on PA (for example, see [6]). Expressly, it was suggested that translation and thus PA is correlated with adaptation to the transfer RNA (tRNA) pool [7], weak mRNA folding (mF) at the beginning of the open reading frame (ORF) [8], ORF length [9], GC content [10] and various ancillary features of the 5′ untranslated region (UTR) [1].

In addition, it was found that highly expressed genes tend to evolve at a slower rate [11], and to have more protein–protein interactions [12].

Recently, a new technology for measuring folding strength of RNA sequences at single-nucleotide resolution was developed

[13]. The product of this method, named the Parallel Analysis of RNA Structure (PARS) score, includes the estimated ratio between the probability that each nucleotide (nt) in the transcript is in a double-stranded conformation and the probability that it is in a single-stranded conformation. The PARS score was computed *in vitro* for transcripts devoid of any ribosomes. As mF is a main feature of a transcript, it might affect its translation rate, or might be related to its PA in a non-causal way (for example, via its relation to the mRNA levels).

In this study, we used the availability of such a new tool to analyse the relationship between mF strength and PA.

RESULTS

The mF strength of a transcript (or a part of it) was defined as the mean PARS score over the sequence; higher values of this measure correspond to stronger folding. We found the correlation between PA (the mean of four data sets, Methods) and mF strength to be 0.68 ($P = 10^{-200}$; Fig 1A); thus, except for measures of codon bias (for example, the tRNA Adaptation Index (tAI); Methods), the mF strength is the feature with the highest known correlation to PA (Fig 1B). Among the analysed features we included amino acid frequencies (which are known to correlate with the expression levels [3,14]), GC content and the ORF length. Distinctively the correlation between mF strength and PA is slightly higher than the correlation between mRNA levels and PA (Fig 1B).

The correlation remains significant when controlling for mRNA levels (Fig 1C), and (again excluding codon bias) mF strength is the feature with the highest correlation to PA given mRNA levels (Fig 1C). When correlating mF strength with mRNA levels the correlation ($r = 0.695$; $P < 10^{-200}$) is higher than any other feature of the coding sequence.

In addition, a significant correlation was found between predicted local mF energy (Methods) and PA (Fig 1D). When we performed several regression analysis between PA and various variables including mF strength, amino acid frequencies, mRNA levels, codon bias (tAI, Methods), GC content, protein half-life and gene length, we found that mF strength has significant effect on PA even when considering all the other variables ($P = 8.9 \times 10^{-17}$ for mF strength; total correlation of the regressor $r = 0.84$ ($P < 10^{-200}$)); thus, the correlation between mF strength and PA cannot be explained solely by the aforementioned variables.

[1]School of Computer Science, and
[2]Faculty of Engineering, Department of Biomedical Engineering, Tel Aviv University, Ramat Aviv, Israel
+Corresponding author. Tel: +972 3 6405836; Fax: +972 3 6407939;
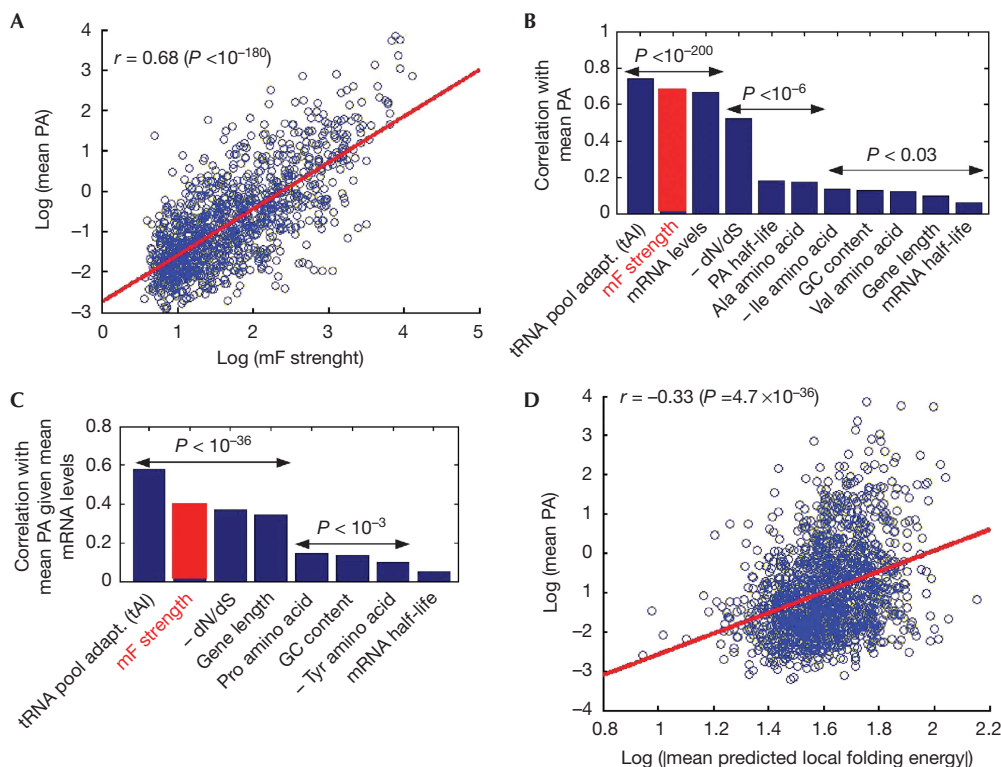E-mail: tamirtul@post.tau.ac.il

Fig 1 | Strong association between mRNA folding (mF) strength and protein abundance (PA). (A) A dot-plot of PA versus mean mF strength of the transcript. **(B)** Correlation of various features of the coding sequence with PA (features with the top correlations) in comparison to the correlation with the mF strength. **(C)** Correlation of various features of the coding sequence with PA given mRNA levels (features with the top partial correlations), in comparison to the correlation with the mF strength. **(D)** Correlation between PA and mean absolute levels of predicted local folding energy along the open reading frame . Note that the mean predicted local folding energy is negative but we show the log of its absolute levels. tAI, tRNA Adaptation Index.

It is important to emphasize that previous papers in the field have reported different relations between mF and PA, and were based on predictions of local mF and not real measurements of mF. Specifically, it was shown that there is selection for weak folding at the beginning of the coding sequence [7,15], probably as strong mF in this region decreases the efficiency of ribosomal binding; furthermore, it was shown that there is selection for strong predicted mF along the coding sequence (however there is no correlation between this measure and expression levels) [16].

To understand the observed correlation between mF strength and PA we performed several more tests. First, we found that there is a strong correlation between mF strength and ribosomal density (RD; Methods, $r = 0.52$; $P < 10^{-180}$), and with the product of RD and mRNA levels ($r = 0.75$; $P < 10^{-180}$), suggesting that the observed correlation is related to the process of ribosomal translation (Fig 2A; see regression analysis and comparison to other features in the Methods and supplementary Tables online).

Second, it is possible that the observed correlation is due to the folding in specific regions of the transcript (for example, it was found [7,8,15] that there is selection for weak folding only at the beginning of the coding sequence, possibly to improve ribosomal binding). Thus, we computed the correlations between gene expression levels and mF strength for different segments of the transcript. We found that the observed correlation is higher when

considering only the coding sequence ($r = 0.674$; $P < 10^{-200}$) versus when considering the UTRs (5′UTR: $r = 0.374$ with $P = 1.13 \times 10^{-36}$; 3′UTR: $r = 0.384$ with $P = 1.5 \times 10^{-40}$), and that the mF strength of different segments of the coding sequence have similar correlation with PA (Fig 2B; similar results were obtained when we performed correlations with RD or with the product of RD and mRNA levels; see supplementary Tables online). Thus, the observed correlation is related to the ORF and not due to a specific part of it.

In addition, we found that the ratio between the mean predicted local folding energy of genes with high PA and the mean predicted local folding energy of genes with low PA is higher than in randomized genomes (which maintain the amino acid content of each gene and the codon bias of the original genome; see Methods; $P < 0.05$; similar results were obtained for groups of genes with high/low RD and mRNA levels, Methods); thus, the stronger folding energy of genes with high PA is probably under selection, and is not due to amino acid bias.

DISCUSSION

In this study we endeavour to elucidate the reason for the observed strong correlation between mF strength and gene expression. To this end, we discerned evidence that supports the conjecture that the observed correlation between mF strength and
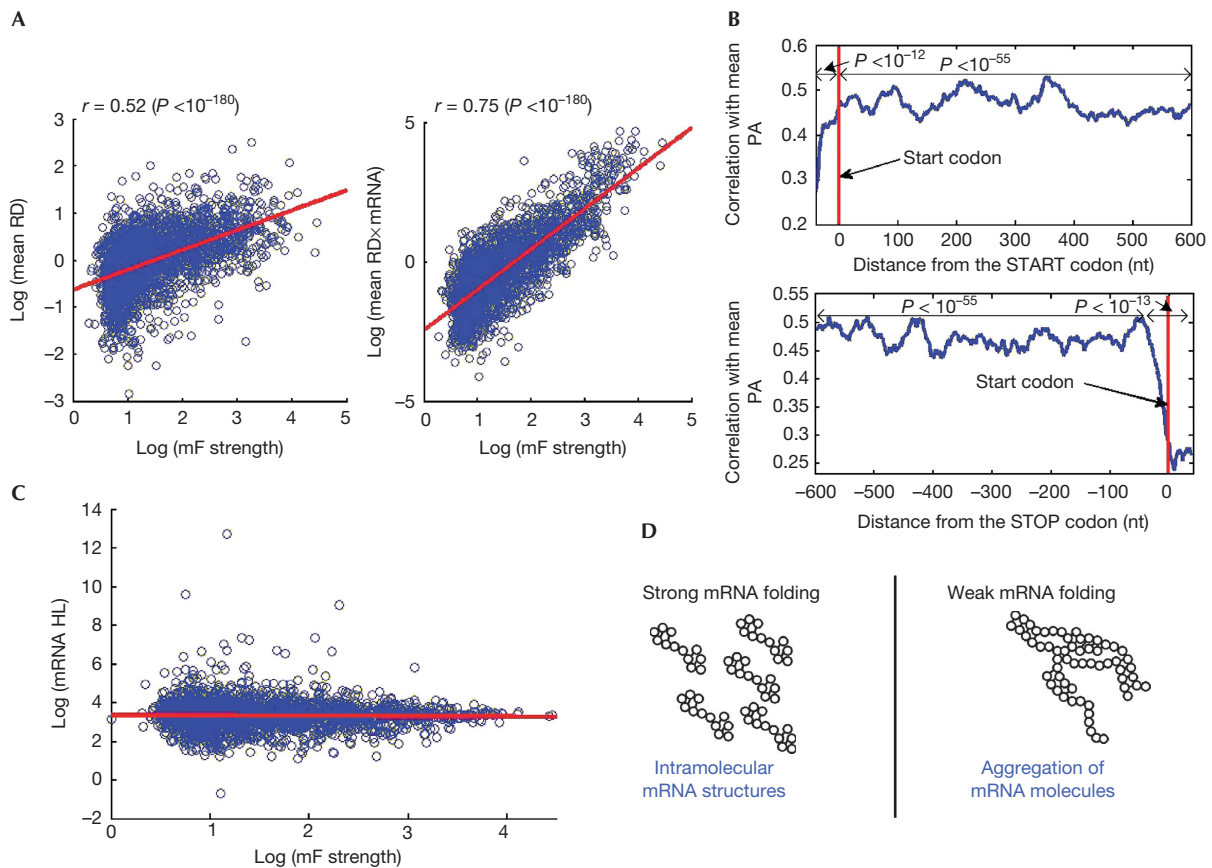
Fig 2 | Possible explanation for the strong correlation between mRNA folding (mF) strength and protein abundance (PA). (A) Correlation between mF strength with ribosomal density (RD) and with the product of RD and mRNA levels. **(B)** Correlation between the mF strength of different segments of the coding sequence and PA (similar results were obtained for correlation with mean RD, and mean RD × mean mRNA levels—see supplementary Figs S1–S4 online). **(C)** There is no correlation between mF strength and mRNA half-life (HL). **(D)** An illustration of a possible explanation for the relation between mF strength and PA.

expression levels (mRNA levels or PA) is not due to an artefact and/or bias: First, similar results were obtained based on the computational predictions of mF, and not only for the PARS score. Second, the results remain significant also when controlling for various features of the transcript, which might explain the observed correlation between the PARS score and PA (for example, amino acid bias, GC content and dozens of extra features).

Third, a bias in the PARS score is expected to be related to gene mRNA levels (and not their PA), as it is based on analysis of mRNA molecules. However, the observed correlation between PA and mF strength remains significant even when we controlled for mRNA levels.

In addition, Kertesz *et al* [13] performed several tests to demonstrate that their method is without bias. Specifically they showed that: (A) the RNase cleavage, adaptor ligation and complementary DNA conversion do not introduce significant sequence biases, (B) the protocol has a very small bias towards particular regions along the transcript; (C) the protocol captures RNA fragments in proportion to their abundance in the initial pool; (D) the signals generated by RNase V1 (that measures base-pair conformation) are highly distinct from those generated by S1 nuclease (that measures single-stranded RNA); (E) the RNA

structures measured by PARS are similar to those obtained with traditional footprinting, reported structures of yeast coding and non-coding RNAs, and with computational predictions of RNA structure.

How does mF strength contribute to translation elongation and/or gene expression efficiency? It is important to emphasize that the correlations reported in this study do not suffice for determining the causality of the relation. We propose two general plausible explanations for the reported phenomenon: (1) Higher expressed genes are selected for stronger folding that not necessarily improves directly, and/or in a causal way their expression levels and/or translation efficiency. (2) The stronger folding of a gene improves its expression or translation levels in a causal/direct way.

We offer several explicit explanations for the observed relations.

One possibility is that it increases mRNA half-life. However, the correlation between mF strength and mRNA half-life is very low ($r = -0.0607$; $P = 0.026$; Fig 2C) in comparison to the correlations between mF strength and gene expression reported above, even when controlling for RD (to control for the possibility that higher ribosomal densities might increase mRNA half-life [17]; $r = -0.017$ with $P = 0.39$). Thus, it seems that mRNA

degradation, which is usually regulated via the ends of the transcripts [16,18], is not a main explication for the observed correlation between PA or mRNA levels and mF strength.

Another possibility is that the observed correlation might be related to lack of selection for weak mF (which improves elongation rates) for highly expressed genes, as such genes are typically occupied by more ribosomes that unfold their mRNA. However, the observed correlations between mRNA levels/PA and mF strength remain significant and high even when controlling for RD (r(PA, PARS|RD) = 0.5822, $P = 6.4 \times 10^{-99}$; r(mR,PARS|RD) = 0.5402, P-value = 2.49×10^{-196}), supporting the rejection of this hypothesis as a primary explanation for the observed correlation.

In addition, previous studies based on small-scale experimental and computational techniques for measuring or estimating mF, supported the conjecture that strong mF should have a negative effect on translation elongation and initiation rates [8,19,20]. Thus, strong mF probably does not improve the translation efficiency of a gene directly.

It is plausible that the observed correlation is due to lesser-understood features of gene expression.

One plausibility is related to self-folding versus aggregation of mRNA molecules. If self-intramoleculer folding is strong enough it can prevent aggregation of mRNA molecules by hybridization (illustration in Fig 2D; for example, see ref. 21 regarding competition between self-folding and aggregation of mRNA molecules). Aggregation of mRNA molecules should decrease the translation efficiency of the genes involved in it. Therefore, under this hypothesis, the tendency to aggregate should be more deleterious for genes with higher mRNA levels and for genes that are under selection for higher translation rates. Thus, under this hypothesis, these genes are reasonably under stronger selection for increased intramoleculer folding.

It is also plausible that the observed correlation is due to (not necessarily mutually exclusive) facets of the gene translation process (or more generally the gene expression process) that are yet unknown. For example, this correlation might be related to the biophysical and biomechanical features of gene translation [19,22], which have higher efficiency and/or fidelity when the mF is higher. The answers related to these aspects of translation can be gained based on molecular dynamic simulations coupled with real-time high-resolution measurements of single ribosomes' translation rates (for example, see ref. 23).

We propose the following experiment for determining the dominant mechanism that associates mF strength and expression levels: Generate a highly expressed library of a heterologous green fluorescent protein(s) with different codon bias and fluorescent mRNA molecules, while maintaining the protein amino acid sequence. For each version of the protein, the self-folding (for example, according to a version of the PARS approach), the tendency to aggregate (for example, by measuring and analysing the spatial mRNA molecule florescence distribution; see [24] for an example of how mRNA molecules can be fluoresced), the mRNA level, the PA (for example, by the florescent levels of the green fluorescent protein) and growth rate should be measured. Such an experiment will enable calculating correlations between PA or mRNA levels, and mF strength even when controlling for aggregation levels, in addition to computing the correlations between aggregation levels and growth rate.

METHODS

Data used. We used the following data in our analysis:

PARS score at a single-nucleotide resolution. The large-scale data of RNA structure was downloaded from the study by Kertesz *et al* [13].

In our analysis we use the ratio of probabilities (instead of the original log ratio) of each nucleotide. For each transcript, we average across the PARS score of its nucleotides, and perform our analyses with the resultant gene vector, representing the PARS score (mF strength). A similar approach was used when we analysed subsequences/segments of a transcript.

Protein abundance. We considered four quantitative large-scale measurements of PA: large-scale data from the study by Ghaemmaghami *et al* [25], two large-scale measurements in two conditions from the study by Newman *et al* [26] and large-scale PA from the study by Lee *et al* [27]. We average across the four data sets of PA measurements (after normalizing each data set by its mean), to minimize experimental noise. Similar results were obtained when we analysed each data set separately (see supplementary Tables online).

Ribosomal densities. We consider two large-scale RD measurements (the number of ribosomes occupying the transcript divided by its length); each generated by a different technology. The first data set was generated more recently by Ingolia *et al* [28] and the second by Arava *et al* [29].

Similarly to the PA, we averaged across the two data sets of RD to reduce experimental noise. Similar results were obtained when we analysed each data set separately (see supplementary Tables online).

mRNA levels. We considered two large-scale measurements of mRNA levels from the studies by Ingolia *et al* [28] and Wang *et al* [30].

Similarly to PA and RD, we averaged across the two data sets of mRNA levels.

Coding sequences and UTRs. The coding sequences and UTRs of *Saccharomyces cerevisiae* were downloaded from the study by Kertesz *et al* [13], and the UTR lengths were based on the study of Nagalakshmi *et al* [31].

tRNA copy number. The computation of tAI is based on tRNA copy numbers (see below). These data were taken from the study by Tuller *et al* [7].

dN/dS estimations. These data were taken from the study by Wall *et al* [32].

mRNA and PA half-lives. Measurements of mRNA half-lives were taken from the study by Shalem *et al* [33], and the measurements of PA half-lives were taken from the study by Belle *et al* [34].

Correlations. All the correlations reported in this study are Spearman correlations.

Prediction of folding energy. We used the Matlab rnafold function (Matlab Bioinformatics toolbox), which predicts the folding energy of the secondary structure associated with the minimum free energy for an RNA sequence (or subsequence).

To obtain the estimation of the mean local folding of a genomic sequence, we estimated the folding energy of all the sliding windows of length 40 nt (slide of 1 nt; 40 is the approximated length of the ribosomes' footprint [35]), and averaged the resultant folding energy prediction of all the windows induced by the sequence. Specifically, mean predicted local folding energies reported in the paper were performed based on the entire

transcript of each gene (that is, including 5′UTR, ORF and 3′UTR) and for each region (5′UTR, ORF and 3′UTR) separately. The correlation between the predicted mean local folding and the PARS score across the *S. cerevisiae* ORFs is $r = -0.36$ ($P = 4.2 \times 10^{-85}$).

Randomized profiles of folding energy. To evince that the correlation between the mean predicted folding energy and PA is not due to amino acid bias (that is, supporting the conjecture that the stronger local folding energy of highly expressed genes is selected for), we performed the following randomization test:

We divided the genes to two groups: 50% of the genes with the highest PA versus 50% of the genes with the lowest PA. We computed the ratio between the mean local folding energy in the group of highly expressed genes, and in the group of lowly expressed genes.

We compared the obtained ratio correlation with the ratio observed for randomized versions of the genome. The genome was randomized in the following manner: for each amino acid in each gene, we sampled a codon while considering the genomic codon frequency/codon-bias in *S. cerevisiae* (that is, more frequent codons in the genome have a higher probability of being sampled).

Thus, the randomized genomes maintained both the amino acid content of each coding sequence, and the codon frequencies of the original genome. We compared the ratio defined above of the 20 randomized genomes to the original one. Similar analyses were performed when we divided the genes to the group with top/lowest: RD, mRNA levels and RD × mRNA (instead of PA).

In all the cases, the ratio between the mean local folding energy of the group with top values versus the group with the lowest values was higher for the original data, than all the 20 randomized genomes (that is, empirical *P*-value <0.05). Specifically, the real ratio was 2.19/4.02/10.7/12.44 standard deviations from the mean of the randomized genomes in the case of PA, RD, mRNA levels and (RD) × (mRNA levels) respectively.

tRNA Adaptation Index. The tAI [36] gauges the availability of the different tRNA molecules for each codon along an mRNA (see exact details in [36]). We obtained tAI measurements from the study by Tuller *et al* [7].

Codon Adaptation Index. Codon Adaptation Index is a codon bias measure similar to the tAI and was downloaded from yeast database (http://www.yeastgenome.org/).

Analyses preformed based on the PARS score. We performed non-parametric Spearman correlations for the mean PARS score of the entire sequence (mF strength, see above) of the 5′UTR, coding sequence and 3′UTR, respectively; the mean predicted local folding energy of the entire sequence (see above) of the 5′UTR, coding sequence and 3′UTR, respectively, with the mean PA, mean PA given mean mRNA levels (partial correlation), mean RD, the product of the mean RD and mRNA levels, and mRNA half-lives.

These correlations were compared to the correlations of the following variables with PA, mean PA given mean mRNA levels (partial correlation), mean RD, the product of the mean RD and mRNA levels: GC content, tAI, dN/dS, amino acid frequencies, the length of the coding sequence, mean mRNA levels, mRNA half-lives and protein half-lives.

mF strength of different segments of the coding sequence. We sought to investigate the relation between the mean PARS score (mF strength) of different segments of the transcript, and the PA

of the gene. To that end, we created for each gene a PARS score profile by averaging across all sliding windows of length 40 nt, taking up to 40 windows (that is, 40 nt) of the 5′UTR upstream to the start codon and 600 windows (that is, 600 nt) of the coding sequence. Positions out of the boundaries of a gene (that is, for genes with 5′UTR shorter than 40 nt and/or ORF shorter than 600 nt) were replaced with NaNs (missing values). The product of this stage is a matrix of mF strength values, with a row (of length 640) for each gene, and the *i*th column in the matrix corresponding to the *i*th window from the beginning of the transcript. We computed the correlation of each of the 640 columns with PA (Fig 2B) and RD (see supplementary Figs online).

Conversely, we performed the same analysis for 40 windows of the 3′UTR downstream of the stop codon, and 600 windows of the coding sequence upstream from the stop codon (Fig 2B; the supplementary Figs online for RD). In this case, we generated a matrix of mF strength values, with a row (of length 640) for each gene, and the *i*th column in the matrix corresponding to the *i*th window from the end of the transcript.

Regressors and correlations. To show that the correlation between mF strength and PA cannot be explained by the aforementioned additional variables alone, we performed several linear regression analysis between PA (the dependent variable) and a set of independent variables that includes the mF strength, amino acid frequencies, mRNA levels, codon bias (tAI), GC content and gene coding sequence lengths (a total of 25 variables). The output of this analysis, performed by MATLAB, shows that there is a significant positive correlation between the mF strength and PA, and that it has significant effect on PA even when considering all the other variables (as reported above). Similar results were obtained when we replaced the dependent variable with RD or with the product of the mean RD and mRNA levels: we found that mF strength has significant effect on mean RD even when considering all the other variables ($P = 8.9 \times 10^{-17}$ for mF strength; total correlation of the regressor $r = 0.67$ ($P < 10^{-200}$)); thus, the correlation between mF strength and RD cannot alternatively be explained by the aforementioned variables.

Similarly, when we performed the same several regression with the product of the mean RD and mRNA levels, we received a correlation of $r = 0.89$ ($P < 10^{-200}$; $P = 8.9 \times 10^{-17}$ for mF strength).

The feature 'product of the mean RD and mRNA levels'. In addition to the feature RD that represents the mean ribosomal 'load' of a certain mRNA of the gene, we also considered the feature (mRNA levels) × (RD). This feature represents the total ribosomal load of a certain gene; thus, in our opinion it should better correlate with coding sequence features related to the optimality of translation elongation (see also [37]).

Supplementary information is available at EMBO *reports* online (http://www.emboreports.org).

CONFLICT OF INTEREST
The authors declare that they have no conflict of interest.

REFERENCES

1. Vogel C, Abreu Rde S, Ko D, Le SY, Shapiro BA, Burns SC, Sandhu D, Boutz DR, Marcotte EM, Penalva LO (2010) Sequence signatures and mRNA concentration can explain two-thirds of protein abundance variation in a human cell line. *Mol Syst Biol* **6**: 1–9

2. Plotkin JB, Kudla G (2010) Synonymous but not the same: the causes and consequences of codon bias. *Nat Rev Genet* **12**: 32–42

3. Tuller T, Kupiec M, Ruppin E (2007) Determinants of protein abundance and translation efficiency in *S. cerevisiae*. *PLoS Comput Biol* **3**: 2510–2519

4. Gingold H, Pilpel Y (2011) Determinants of translation efficiency and accuracy. *Mol Syst Biol* **7**: 481

5. Drummond DA, Wilke CO (2009) The evolutionary consequences of erroneous protein synthesis. *Nat Rev Genet* **10**: 715–724

6. Reuveni S, Meilijson I, Kupiec M, Ruppin E, Tuller T (2011) Genome-scale analysis of translation elongation with a ribosome flow model. *PLoS Comput Biol* **7**: e1002127

7. Tuller T, Waldman YY, Kupiec M, Ruppin E (2010) Translation efficiency is determined by both codon bias and folding energy. *Proc Natl Acad Sci USA* **107**: 3645–3650

8. Kudla G, Murray AW, Tollervey D, Plotkin JB (2009) Coding-sequence determinants of gene expression in Escherichia coli. *Science* **324**: 255–258

9. Coghlan A, Wolfe KH (2000) Relationship of codon bias to mRNA concentration and protein length in *Saccharomyces cerevisiae*. *Yeast* **16**: 1131–1145

10. Lercher MJ, Urrutia AO, Pavlicek A, Hurst LD (2003) A unification of mosaic structures in the human genome. *Hum Mol Genet* **12**: 2411–2415

11. Pal C, Papp B, Hurst LD (2001) Highly expressed genes in yeast evolve slowly. *Genetics* **158**: 927–931

12. Krylov DM, Wolf YI, Rogozin IB, Koonin EV (2003) Gene loss, protein sequence divergence, gene dispensability, expression level, and interactivity are correlated in eukaryotic evolution. *Genome Res* **13**: 2229–2235

13. Kertesz M, Wan Y, Mazor E, Rinn JL, Nutter RC, Chang HY, Segal E (2010) Genome-wide measurement of RNA secondary structure in yeast. *Nature* **467**: 103–107

14. Welch M, Govindarajan S, Ness JE, Villalobos A, Gurney A, Minshull J, Gustafsson C (2009) Design parameters to control synthetic gene expression in *Escherichia coli*. *PLoS One* **4**: 1–10

15. Gu W, Zhou T, Wilke CO (2010) A universal trend of reduced mRNA stability near the translation-initiation site in prokaryotes and eukaryotes. *PLoS Comput Biol* **6**: 1–8

16. Katz L, Burge CB (2003) Widespread selection for local RNA secondary structure in coding regions of bacterial genes. *Genome Res* **13**: 2042–2051

17. Iost I, Dreyfus M (1995) The stability of *Escherichia coli* lacZ mRNA depends upon the simultaneity of its synthesis and translation. *EMBO J* **14**: 3252–3261

18. Carrier TA, Keasling JD (1997) Controlling messenger RNA stability in bacteria: strategies for engineering gene expression. *Biotechnol Prog* **13**: 699–708

19. Tuller T, Veksler-Lublinsky I, Gazit N, Kupiec M, Ruppin E, Ziv-Ukelson M (2011) Composite effects of gene determinants on the translation speed and density of ribosomes. *Genome Biol* **12**: R110

20. Babendure JR, Babendure JL, Ding JH, Tsien RY (2006) Control of mammalian translation by mRNA structure near caps. *RNA* **12**: 851–861

21. Guttal V, Bundschuh R (2006) Model for folding and aggregation in RNA secondary structures. *Phys Rev Lett* **96**: 018105

22. Jenner L, Rees B, Yusupov M, Yusupova G (2007) Messenger RNA conformations in the ribosomal E site revealed by X-ray crystallography. *EMBO Rep* **8**: 846–850

23. Uemura S, Aitken CE, Korlach J, Flusberg BA, Turner SW, Puglisi JD (2010) Real-time tRNA transit on single translating ribosomes at codon resolution. *Nature* **464**: 1012–1017

24. Haim L, Zipor G, Aronov S, Gerst JE (2007) A genomic integration method to visualize localization of endogenous mRNAs in living yeast. *Nat Methods* **4**: 409–412

25. Ghaemmaghami S, Huh WK, Bower K, Howson RW, Belle A, Dephoure N, O'Shea EK, Weissman JS (2003) Global analysis of protein expression in yeast. *Nature* **425**: 737–741

26. Newman JR, Ghaemmaghami S, Ihmels J, Breslow DK, Noble M, DeRisi JL, Weissman JS (2006) Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature* **441**: 840–846

27. Lee MV, Topper SE, Hubler SL, Hose J, Wenger CD, Coon JJ, Gasch AP (2011) A dynamic model of proteome changes reveals new roles for transcript alteration in yeast. *Mol Syst Biol* **7**: 514

28. Ingolia NT, Ghaemmaghami S, Newman JR, Weissman JS (2009) Genome-wide analysis *in vivo* of translation with nucleotide resolution using ribosome profiling. *Science* **324**: 218–223

29. Arava Y, Wang Y, Storey JD, Liu CL, Brown PO, Herschlag D (2003) Genome-wide analysis of mRNA translation profiles in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci USA* **100**: 3889–3894

30. Wang Y, Liu CL, Storey JD, Tibshirani RJ, Herschlag D, Brown PO (2002) Precision and functional specificity in mRNA decay. *Proc Natl Acad Sci USA* **99**: 5860–5865

31. Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D, Gerstein M, Snyder M (2008) The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* **320**: 1344–1349

32. Wall DP, Hirsh AE, Fraser HB, Kumm J, Giaever G, Eisen MB, Feldman MW (2005) Functional genomic analysis of the rates of protein evolution. *Proc Natl Acad Sci USA* **102**: 5483–5488

33. Shalem O, Dahan O, Levo M, Martinez MR, Furman I, Segal E, Pilpel Y (2008) Transient transcriptional responses to stress are generated by opposing effects of mRNA production and degradation. *Mol Syst Biol* **4**: 223

34. Belle A, Tanay A, Bitincka L, Shamir R, O'Shea EK (2006) Quantification of protein half-lives in the budding yeast proteome. *Proc Natl Acad Sci USA* **103**: 13004–13009

35. Ingolia NT, Ghaemmaghami S, Newman JRS, Weissman JS (2009) Genome-wide analysis *in vivo* of translation with nucleotide resolution using ribosome profiling. *Science* **324**: 218

36. dos Reis M, Savva R, Wernisch L (2004) Solving the riddle of codon usage preferences: a test for translational selection. *Nucleic Acids Res* **32**: 5036–5044

37. Tuller T, Carmi A, Vestsigian K, Navon S, Dorfan Y, Zaborske J, Pan T, Dahan O, Furman I, Pilpel Y (2010) An evolutionarily conserved mechanism for controlling the efficiency of protein translation. *Cell* **141**: 344–354