An Evolutionarily Conserved Mechanism for Controlling the Efficiency of Protein Translation

Tamir Tuller, Asaf Carmi, Kalin Vestsigian, Sivan Navon, Yuval Dorfan, John Zaborske, Tao Pan, Orna Dahan, Itay Furman, Yitzhak Pilpel

Cell, Volume 141, Issue 2, 344-354, 16 April 2010

Triinu Kõressaar,

Seminar in bioinformatics

TARTU 2010

Introduction

mRNA translation is controlled at multiple stages and by diverse mechanisms (initiation, mRNA secondary structure, tRNA pool)

The abundance of tRNAs that correspond to the different codons in a gene was suggested to determine the speed and accuracy of translation (codons with high and low efficiency)

The Question

Studies that gauge translation efficiency have mostly considered average codon usage over entire genes (do not consider the order in which codons with low and high translation efficiency appear along the transcript)

The order of high-efficiency and low-efficiency codons along transcripts could govern the process of translation (multiple ribosomes are often simultaneously loaded on a given transcript).

Speed and processivity of translation, the overall cost of protein production in cells.

A Universally Conserved Translation Efficiency Profile

tRNA-adaptation index (tAI) to evaluate translation efficiency at each codon

The tAI measure of an entire gene is defined as the (geometric) average of tRNA availability values over all the codons in the gene

For each codon, the tAI considers the availability of the tRNA with the perfectly matched anticodon along with weighted contributions from imperfect codon-anticodon pairs, reflecting wobble interactions

Local tAI for each codon along the mRNA

Let n_i be the number of tRNA isoacceptors recognizing codon i. Let $tGCN_{ij}$ be the copy number of the jth tRNA that recognizes the ith codon, and let S_{ij} be the selective constraint on the efficiency of the codon-anticodon coupling.

We define the ***absolute adaptiveness***, W_i, for each codon i as

$$W_i = \sum_{j=1}^{n_i} (1 - S_{ij}) tGCN_{ij}.$$

The averaged translation efficiency profiles of all the genes in a given genome

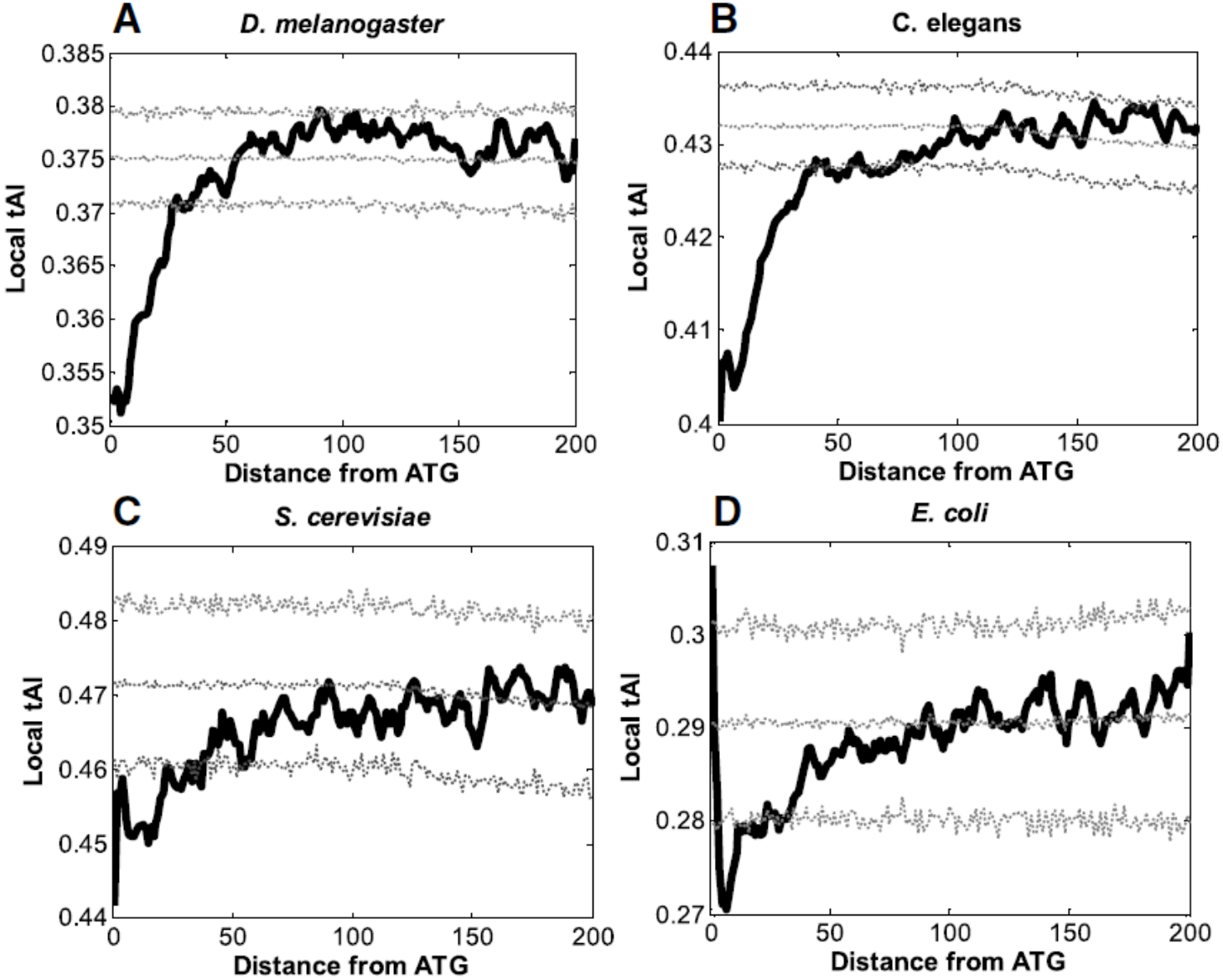
*27 genomes are analyzed
*representatives from all three domains of life

•all genes were lined up according to their start (stop) codon
•an average local tAI value at each position was calculated

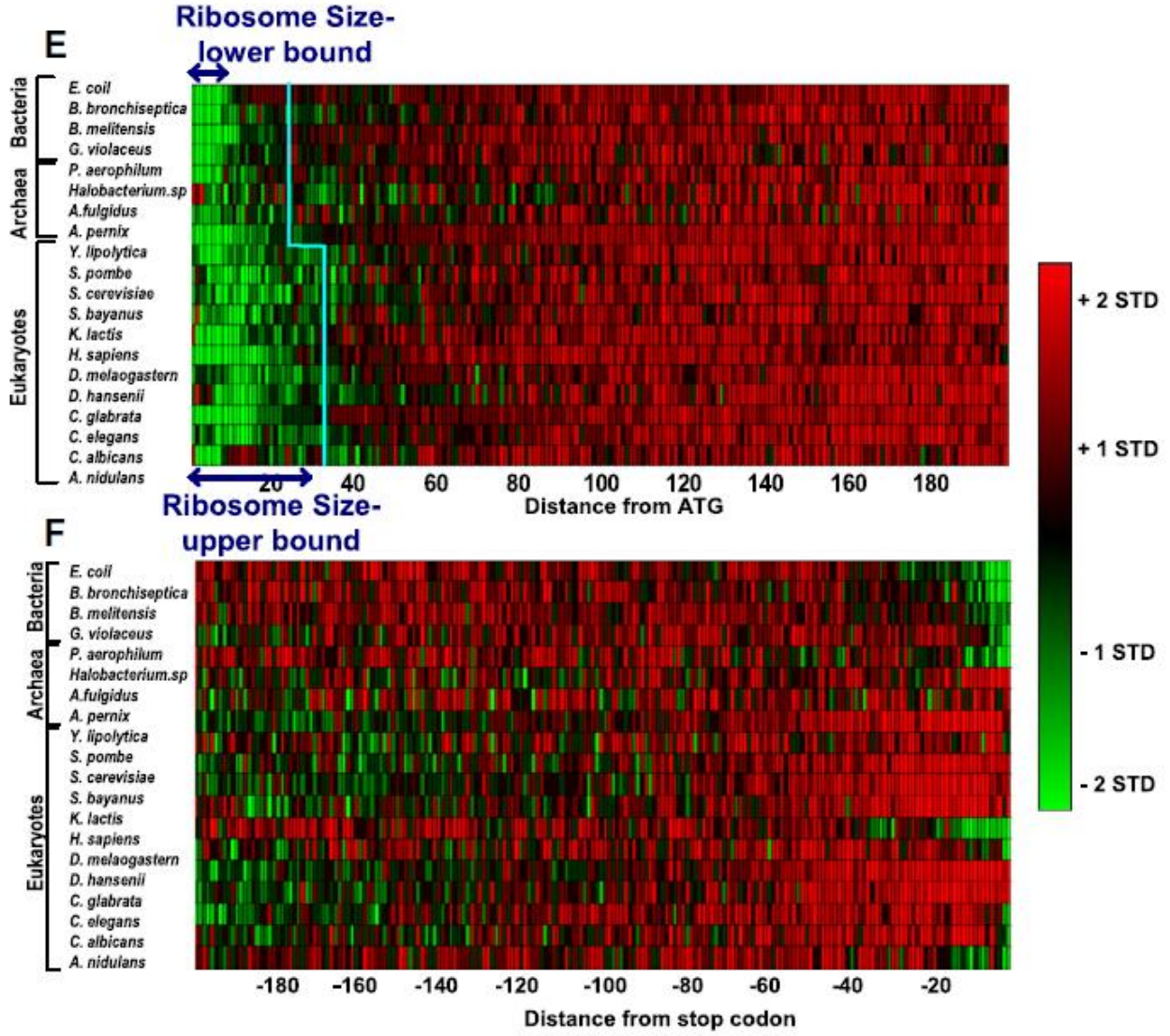
The averaged translation efficiency profile reveals several remarkably conserved features:

1. Translation starts with relatively low efficiency codons for about the first 30–50 positions (‘‘low-efficiency ramp’’ or the ‘‘ramp’’)

2. The ramp region is followed by a plateau with 5%–10% higher translation efficiency on a genome average



(A–D) Averaged translation efficiency profile for the first 200 codons in *D. melanogaster* (A), *C. elegans* (B), *S. Cerevisiae* (C), and *E. coli* (D). Note the different span of values in each subplot. Each figure contains the averaged tAI profile (black) and the randomized profile ±3 standard deviations (gray)

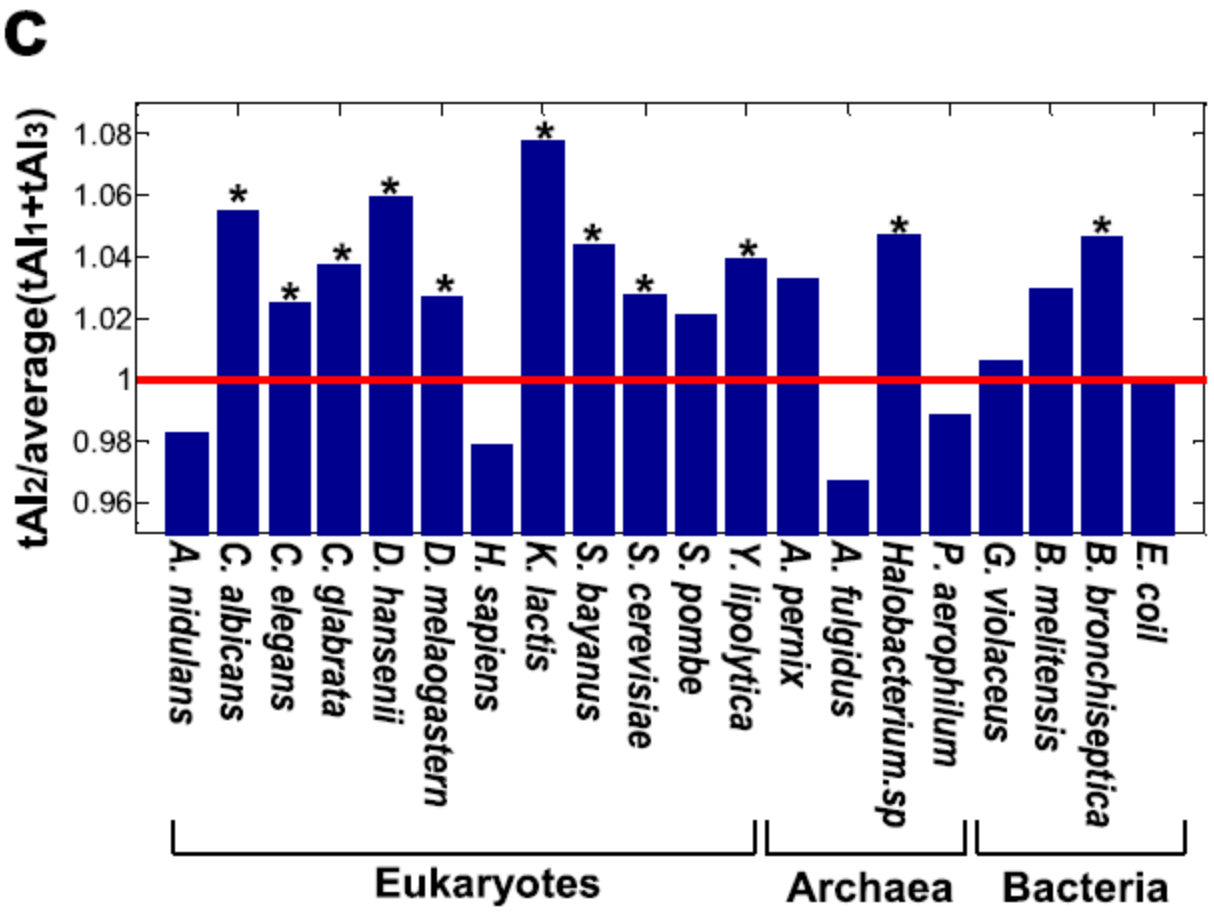


The translation efficiency profiles in various organisms for the start/end codon line-up Each row describes the translation efficiency profile of a different organism and each pixel describes a codon. Green denotes lower tAI, whereas red denotes higher tAI. The blue vertical line (E) denotes the means of the length of the ramp in prokaryotes/eukaryotes; the ratio between the means of these regions in eukaryotes/prokaryotes (34.5/24 = 1.43) may correspond to a difference in the size of the footprinted region of the eukaryotic and prokaryotic ribosomes on transcripts.

The averaged translation efficiency profile reveals several remarkably conserved features:

3. A clear outlier in the ramp is the second codon position, which follows the initiating methionine that shows high efficiency compared to its neighboring codons in the majority of the species

This might support a fast release and recycling of the initiating methionine tRNA



The local tAI of the second codon from the ATG codon divided by the mean tAI of the first and the third codons for various organisms. Organisms where this ratio is significantly high are marked with an asterisk (empirical p value < 0.05/20; by comparison to the distribution of the ratios between the tAI of codon i and the mean tAI of codons i-1 and i+1 over all the codons along the translation efficiency profile; the p value is the fraction of positions with lower/higher ratio). In most of the organisms (14 out of 20) this ratio is larger than one (the red line), in 11 organisms this ratio was significantly high; the ratio was not significantly low in any of the analyzed organisms.

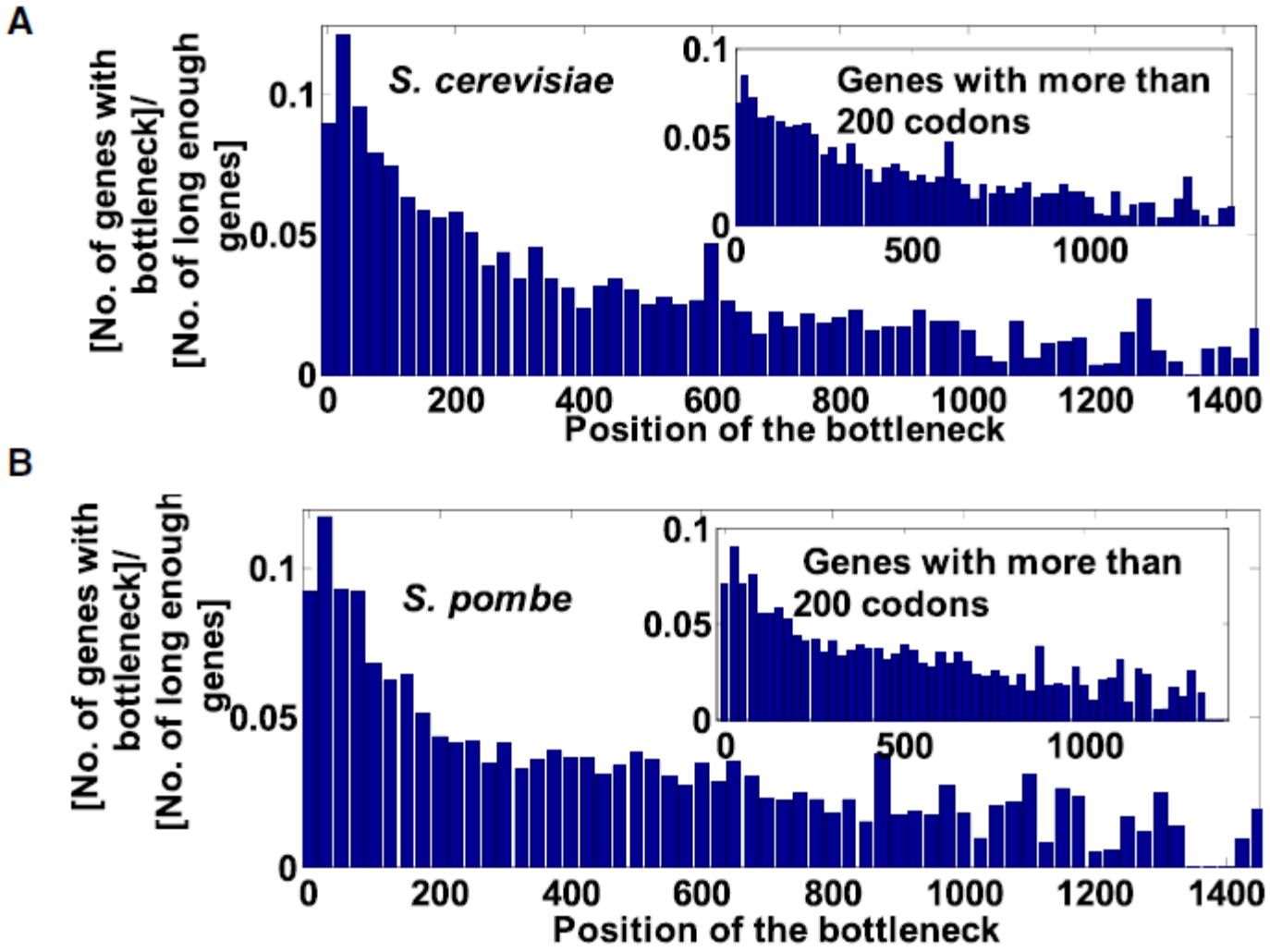
Inspecting single genes

•**“Bottleneck” region** in each individual gene—a sequence window of 15 codons in length

•Represents the length of the ribosome footprint region on mRNAs, with the highest averaged values of 1/(local tAI) (i.e., 15 codons with the longest dwell time in a gene)

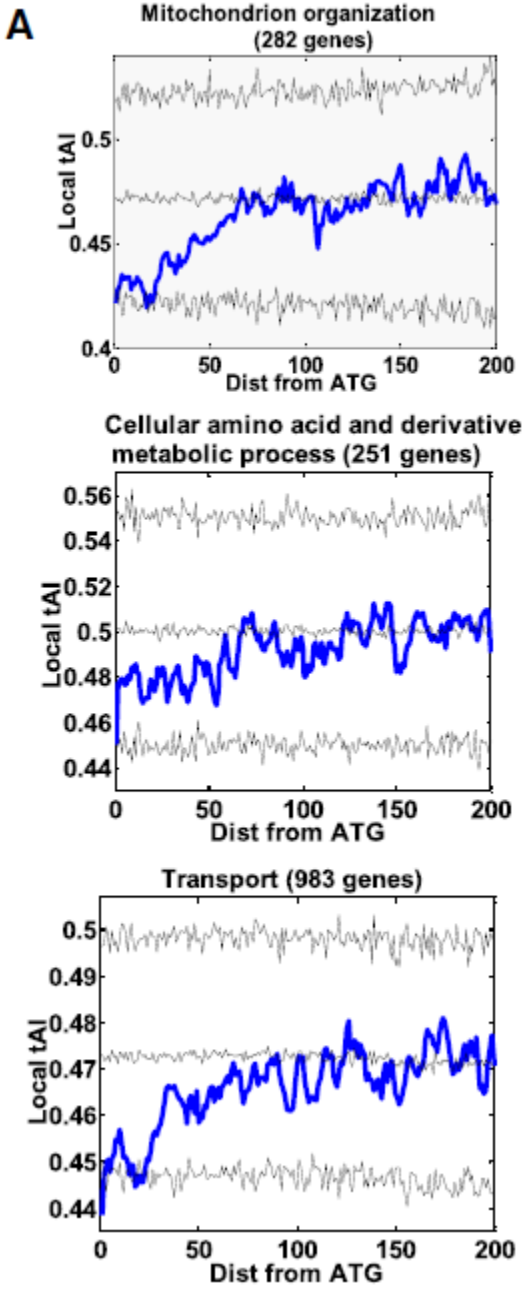
•Two distantly related yeast species, *S. cerevisiae* and *S. pombe* were analyzed

•Both distributions show a consistent picture—a clear tendency to have the bottleneck relatively early along the genes



Bottlenecks in Translation Efficiency Tend to Be Localized Close to mRNAs 50 Ends The distribution of the positions of the bottlenecks in S. cerevisiae (A) and S. pombe (B). For each bottleneck position, the number of genes with a bottleneck in that position was normalized by dividing it by the number of genes whose length extends beyond that position. The distribution is similar also when considering only genes with more than 200 codons (inset).

Genes that share a biological function

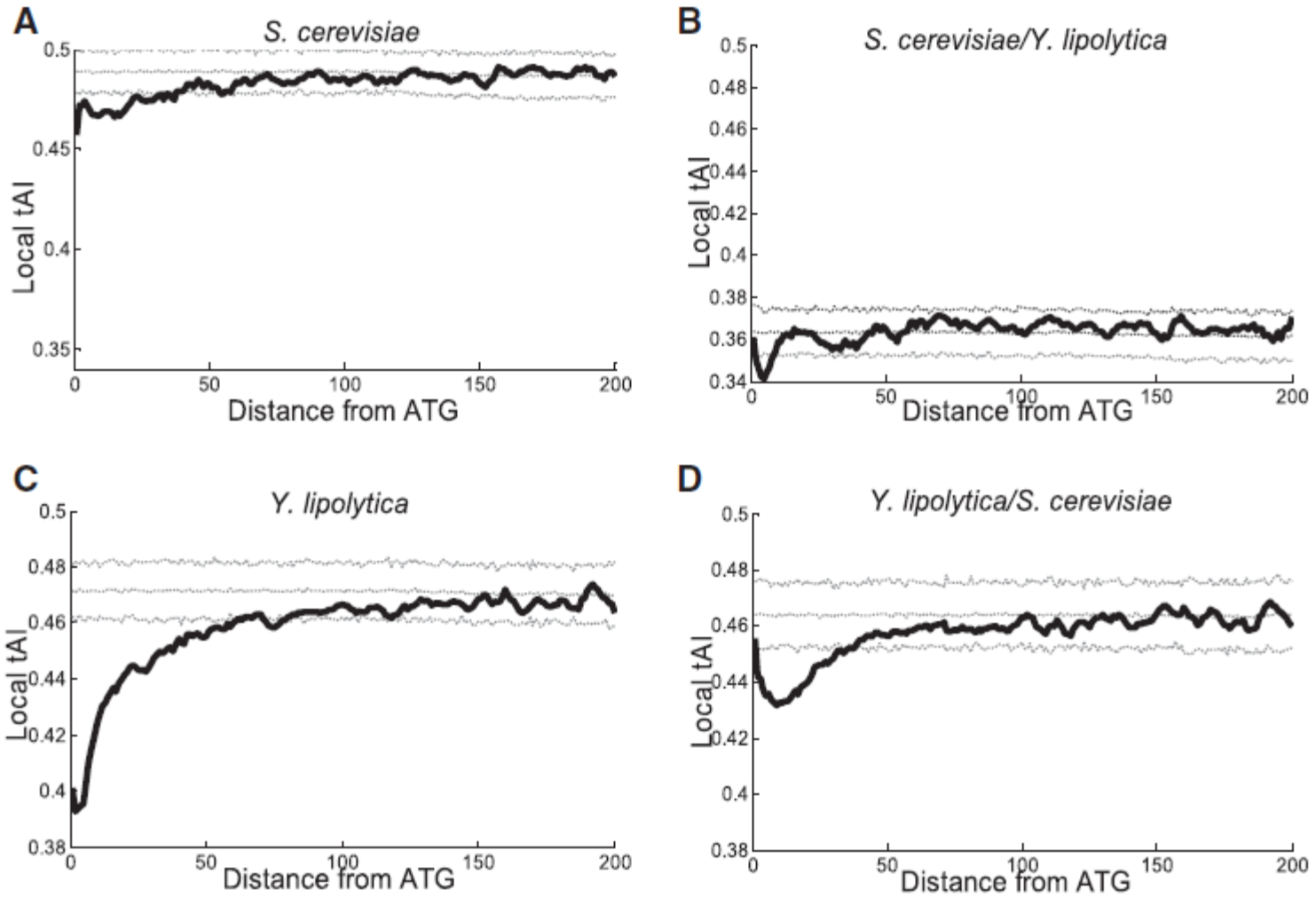


The Universal Translation Efficiency Profile Is under Selection

The translation efficiency profile is highly conserved in evolution, but is the profile under direct selection?

The observed profile is conserved merely because the tRNA pool and codon biases are sufficiently conserved?

By-product of a putative position-dependent variation in the GC content along genes?



Hybrid Analysis Indicates Selection for Coevolution of tRNA Pools and Genes Sequences to Preserve the Ramp Translation efficiency profiles with native and nonnative tRNA pools for start codon line-up.
(A) The translation efficiency profile of *S. cerevisiae*.
(B) The translation efficiency profile of *S. cerevisiae* using *Y. lipolytica* tRNA pool.
(C) The translation efficiency profile of *Y. lipolytica*.
(D) The translation efficiency profile of *Y. lipolytica* using *S. cerevisiae* tRNA pool.
The black bolded line represents the actual calculated tAI profile; the gray lines represent the mean ± 3 standard deviations of the tAI profiles of randomized sets of gene.

The Universal Translation Efficiency Profile Is under Selection

The translation efficiency profile is highly conserved in evolution, but is the profile under direct selection?

The observed profile is conserved merely because the tRNA pool and codon biases are sufficiently conserved?

By-product of a putative position-dependent variation in the GC content along genes?

Codon-tRNA Adaptation May Determine Translation Speed and Ribosome Density along Transcripts

What actual physical or biochemical quantity is encoded by the translation efficiency profile?

Values of the local tAI determine the local speed of movement of the translating ribosome through each codon along mRNAs

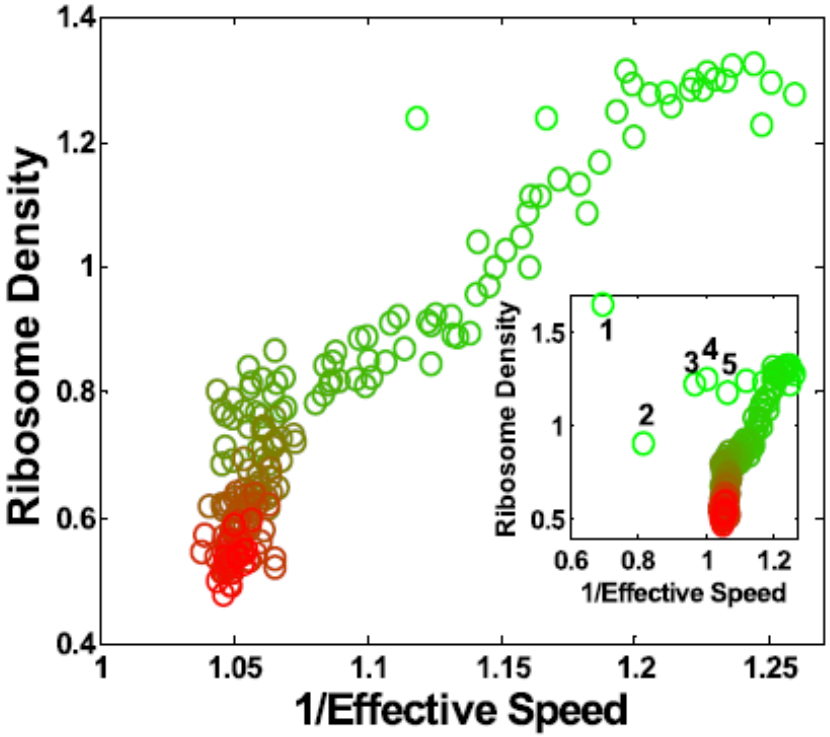
Density of ribosomes

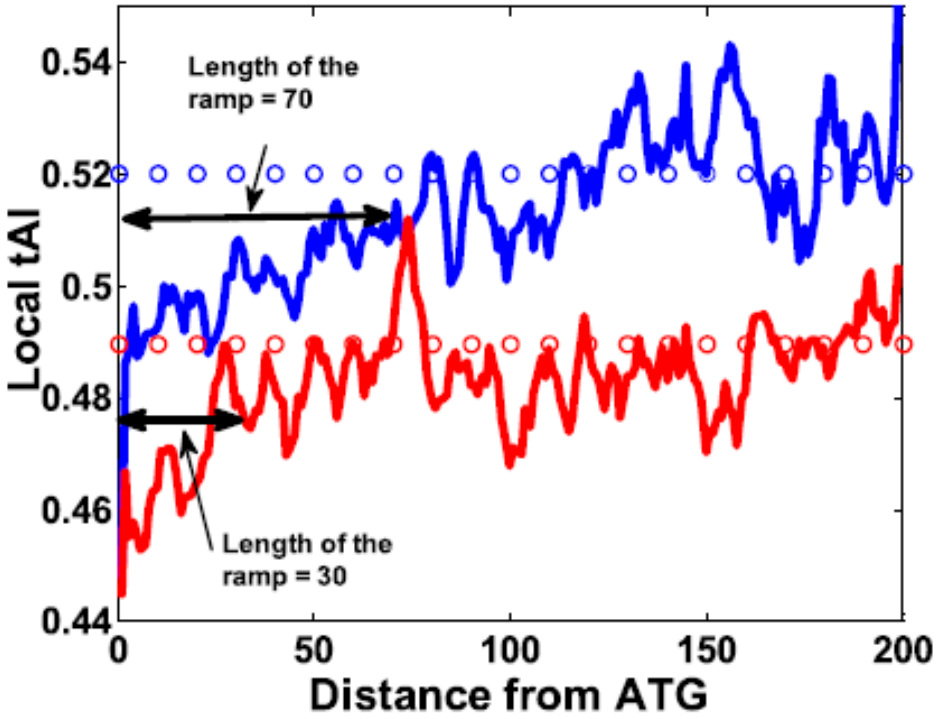
Codon-tRNA Adaptation May Determine Translation Speed and Ribosome Density along Transcripts

Ribosome densities at a single base resolution is measured genome-wide for thousands of transcripts in the *S. cerevisiae* genome (Ingolia et al., 2009)

Experimentally measured ribosomal density with the translation efficiency profile reveals similarity (Pearson correlation $r = 0.5749$; $p < 10^{28}$)

‘‘Effective speed profile’’ - at higher initiation rates, ribosome jamming (simulation of ribosome movement)





Experimentally Measured Ribosome Density Negatively Correlates with Computed Translation Efficiency

(left) Correlation between experimentally measured ribosomal density (Ingolia et al., 2009) and the reciprocal of the simulated speed profile when not considering the first five codons (which are outliers) and when considering all the codons (the subfigure at the lower right corner). Dots are color coded according to codon location along genes, with the greenest dots representing codons that are close to the ATG, and codons that are farthest away in red. The density and speed profiles were obtained by averaging the profiles at each position of the genes in the *S. cerevisiae* genome. The speed profile was obtained by simulating ribosomal scan of all the transcripts in this species. The Pearson correlation between density and 1/speed is 0.93 ($p < 10^{75}$). The correlation between density and the reciprocal of original ‘nominal’’ translation efficiency profile is lower, r = 0.5749 ($p < 10^{28}$).

(right) The translation efficiency profile of genes with the top and the lowest ribosomal density distribution. As can be seen, the extent of ramping decreases at lowly dense genes.

Highlights

The efficiency of translation is universally lower across the first ~50 codons

Evolutionary forces act to maintain this profile of translation elongation speed

The profile is predictive of ribosome density for translation of yeast genes

The ramp in efficiency may contribute to fitness and managing the cost of translation