RESEARCH COMMENTARY

Benefits of being biased!

S U T I R T H   D E Y *

*Evolutionary Biology Laboratory, Evolutionary & Organismal Biology Unit,
Jawaharlal Nehru Centre for Advanced Scientific Research,
Jakkur P.O. Box 6436, Bangalore 560 064, India*

The genetic code is degenerate: most amino acids are coded by multiple codons. However, it is known that certain codons for a given amino acid are preferred over other synonymous codons. Theoretical studies have predicted and experiments have shown that this kind of a bias can lead to higher efficiency of translation, presumably accompanied by benefits in terms of the Darwinian fitness of the organism. However, population genetics models based on gene sequence data predict that, at least for certain organisms like the fruit fly, *Drosophila melanogaster,* the fitness effects might be too small to be observed. Two recent studies (Carlini and Stephan 2003; Carlini 2004) show that even in such systems, under certain conditions, it is possible to empirically demonstrate the effects of codon bias at the phenotypic level, and the corresponding effect on at least one major component of the fitness of the organism.

The neutral theory of molecular evolution assumes that synonymous mutations do not affect fitness (King and Jukes 1969). This implies that all the different codons that code for the same amino acid are expected to be equally abundant in a large sample of genes (Powell and Moriyama 1997). However, it is known that, in most genes, certain codons are found more frequently than their synonymous counterparts: a phenomenon often termed as codon bias. This naturally leads to the question as to whether codon bias actually has any selective advantage, and several lines of evidence suggest that this is indeed the case. Among all the synonymous codons, the most often used codons are generally the ones whose matching tRNAs are the most abundant in the tRNA pool (Moriyama and Powell 1997). Moreover, the level of bias is generally found to be higher in genes that have high levels of expression (Duret and Muchiroud 1999), suggesting that the presence of codon bias is probably related to translation efficiency (Bulmer 1991). This might be possible in two ways that are not mutually exclusive:

(i) Presence of optimal codons might increase the speed of translation, by decreasing the time taken for attachment of the correct aminoacyl tRNA, which in turn would increase the availability of free ribosomes in the cytoplasm. This would lead to an increase in the level of proteins in a cell, particularly for genes that are expressed at very high rates (Kurland 1987).

(ii) It has been shown that unpreferred codons can increase the frequency of incorporation of incorrect amino acids to the nascent polypeptide (Precup and Parker 1987). This reduces the fidelity of protein synthesis, thus enhancing the metabolic costs incurred in terms of nonfunctional/misfunctional proteins. Hence, codon bias leading to translational accuracy is expected to be under positive selection, particularly in highly expressed genes (Akashi 1994).

It has been shown empirically that the usage of optimal codons can increase the rate of translation and hence presumably affect the fitness of an organism. For example, Sörensen *et al.* (1989) introduced small DNA fragments containing infrequent codons into the *lacZ* gene of *E. coli*, and measured the translation rate of the mRNA transcripts. They found that usage of the biased codons actually increased the translation rate up to six times. However, owing to the multitude of tissue types that can vary in their gene expression and codon usage patterns, extending such works *in vivo* for multicellular organisms was a difficult proposition (Powell and Moriyama 1997). Since synonymous mutations do not change the final protein product, selection for optimal codons is thought to be fairly weak. Using different methodologies, Akashi (1995, 1996) and McVean and Viera (2001) came to the conclusion that the strength of selection (given as the product of the effective population size, N_e and the selection coefficient, *s*) for any synonymous codon substitution in *D. melanogaster* would be almost negligible, implying that the difference in fitness arising out of preferential usage of optimal codons was likely to be very low. Thus, empirical measurements of the effects of codon bias on the fitness of an organism appeared to be difficult, if not impossible.

However, Carlini and Stephan (2003) point out that these estimates are actually averaged over many codon families across a large subset of genes. They argue that since selection coefficients are not the same for all codon families, it might be possible to empirically observe the effects of codon bias by manipulating codon families that have a high level

*E-mail: dey@jncasr.ac.in.

of bias. For this purpose, they chose to manipulate the heavily biased, six-fold degenerate, leucine codon family of alcohol dehydrogenase gene (*Adh*) of *D. melanogaster*. In nature, these flies are found on rotting fruits containing high levels of alcohol, and *Adh* is known to be one of the most highly expressed genes in these organisms (Carlini 2004). Carlini and Stephan (2003) derived three different *Adh* constructs in which 1,6 and 10 codons coding for leucine had been mutated to the unpreferred leucine codon CTA. It is known from prior work that normally the frequency of CTA in *Adh* alleles of *D. melanogaster* is 0 (Nakamura *et al.* 2000). These three constructs, along with the wild type *Adh* allele from which they were derived, were introduced separately into the flies through germline transformations. The flies chosen for the purpose had a splicing defect due to which they could not make any alcohol dehydrogenase (ADH) of their own and hence, any ADH found in the transformed flies would have to be produced by the introduced genes. Thus, there were four lines containing 0 (control), 1, 6 and 10 unpreferred leucine codons and the ADH activity was measured for each line. It was found that, on an average, the introduction of each unpreferred codon led to an approximately 2% drop in ADH activity. There were no appreciable differences in the folding free energies, which indicated no major changes in the secondary structure of the mRNA transcript. Thus the observed differences in the protein activity were directly attributable to the incorporation of unpreferred codons. This shows that, at least in certain extreme cases, the effects of codon bias can indeed be empirically measured at the phenotypic level.

However, natural selection operates on organisms and not on protein levels. In other words, for codon bias to be evolutionarily favored, the possessors of the biased genes must enjoy some fitness advantages. For example, it is possible that the reduction in ADH activity in the flies containing unpreferred substitutions was not enough to produce any observable change in the alcohol tolerance of these flies. In such a case, natural selection would fail to observe the deleterious effects of the unpreferred codons and they would tend to persist in the population. Thus, the next logical step was to confirm if there were indeed some differences in the fitness (as defined by tolerance to alcohol) among these lines. Carlini (2004) measured the alcohol tolerance of the above mentioned lines under different concentrations of ethanol. He found that despite considerable variations within the lines in response to different concentrations of ethanol, a sort of pattern was discernible. The lines with a higher number of unpreferred substitutions were more sensitive to ethanol compared to the controls and the lines with less unpreferred substitutions. For example, the average LD_{50} (ethanol concentration at which 50% of the flies would die) of the lines with 10 codon substitutions was significantly less than the controls and the lines with one substitution. A positive correlation between the LD_{50} of the lines and the corresponding ADH activity

implies that the ethanol tolerance was probably related to the levels of ADH in the flies.

Although these studies indicate that the effects of unpreferred codon substitutions are measurable, they do not contradict the estimations from population genetics theories. As Carlini (2004) points out, these studies were performed on one of the most biased codon families in a very highly expressed gene that is crucial to the fitness of the organism. On the other hand, the existing predictions (Akashi 1995, 1996; McVean and Viera 2001) were based on all codon families and averaged over a number of genes, and therefore, the true value of selection coefficient against unpreferred codons is expected to be fairly low. Moreover, the intensity of selection in nature is expected to be lower than that under the assay conditions, which again would reduce the strength of selection favoring codon bias. However, these two recent studies (Carlini and Stephan 2003; Carlini 2004) are noteworthy as they underline the fact that the effects of codon bias on fitness can be empirically measured atleast in some systems, and open up several possibilities for future work. These include studying other codon families, genes somewhat less expressed, and replacing unpreferred codons with preferred ones (Carlini and Stephan 2003). The fact that they are lucidly written is an added attraction of the two papers.

References

Akashi H. 1994 Synonymous codon usage in *Drosophila melanogaster:* natural selection and translational accuracy. *Genetics* **136**, 927–935.

Akashi H. 1995 Inferring weak selection from patterns of polymorphism and divergence at 'silent' sites in *Drosophila* DNA. *Genetics* **139**, 1067–1076.

Akashi H. 1996 Molecular evolution between *Drosophila melanogaster* and *D. simulans:* reduced codon bias, faster rates of amino acid substitution, and larger proteins in *D. melanogaster*. *Genetics* **144**, 1297–1307.

Bulmer M. 1991 The selection-mutation-drift theory of synonymous codon usage. *Genetics* **129**, 897–907.

Carlini D. B. 2004 Experimental reduction of codon bias in the *Drosophila* alcohol dehydrogenase gene results in decreased ethanol tolerance of adult flies. *J. Evol. Biol.* **17**, 779–785.

Carlini, D.B. and Stephan W. 2003 *In vivo* introduction of unpreferred synonymous codons into the *Drosophila Adh* gene results in reduced levels of ADH protein. *Genetics* **163**, 239–243.

Duret L. and Mouchiroud D. 1999 Expression pattern and, surprisingly, gene length shape codon usage in *Caenorhabditis, Drosophila,* and *Arabidopsis. Proc. Natl. Acad. Sci. USA* **96**, 4482–4487.

King J. L. and Jukes T. H. 1969 Non-Darwinian evolution. *Science* **164**, 788–798.

Kurland C. G. 1987 Strategies for efficiency and accuracy in gene expression. 1. The major codon preference: a growth optimization strategy. *Trends Biochem. Sci.* **12**, 126–128.

McVean G. A. T. and Vieira J. 2001 Inferring parameters of mutation, selection, and demography from patterns of synonymous site evolution in *Drosophila. Genetics* **157**, 245–257.

Moriyama E. N. and Powell J. R. 1997 Codon usage bias and tRNA abundance in *Drosophila. J. Mol. Evol.* **45**, 514–523.

Nakamura Y., Gojobori, T. and Ikemura, T. 2000 Codon usage tabulated from the international DNA sequence database: status for the year 2000. *Nucleic Acids Res.* **28**, 292.

Powell J. R. and Moriyama E. N. 1997 Evolution of codon usage bias in *Drosophila. Proc. Natl. Acad. Sci. USA* **94,** 7784–7790.

Precup J. and Parker J. 1987 Missense misreading of asparagine codons as a function of codon identity and context. *J. Biol. Chem.* **262**, 11351–11356.

Sörensen M. A., Kurland C. G. and Pedersen S. 1989 Codon usage determines translation rate in *Escherichia coli. J. Mol. Biol.* **207**, 365–377.