Codon usage bias and nuclear mRNA concentration: Correlation vs. causation

Wenfeng Qian (钱文峰)[a,b,1] and Jianzhi Zhang[c,1]

Much confusion in genome-wide studies results from mistakenly interpreting correlation as causation. Zhao et al. (1) observe a positive correlation between the codon bias index (CBI)—the extent to which a gene uses preferred synonymous codons—and the nuclear messenger RNA (mRNA) concentration among *Neurospora crassa* genes. This correlation could have originated from the processes of 1) high CBIs causing high mRNA levels, 2) high mRNA levels causing high CBIs, and/or 3) a third factor concordantly altering the CBI and mRNA level of a gene. Considering only process 1, Zhao et al. deduce that "codon usage broadly influences mRNA levels through transcription." Below, we explain why processes 2 and 3 are likely, and why Zhao et al.'s conclusion is unwarranted.

Multiple studies have inferred that preferred codons are decoded more accurately than unpreferred codons (2, 3). Because translations of mRNAs of higher concentrations should generate more mistranslation-induced toxic protein misfolding (3) and misinteraction (4), natural selection against the use of unpreferred codons is stronger in mRNAs of higher concentrations. This differential selection results in higher CBIs of more highly transcribed genes—an instance of process 2.

Using preferred codons in an mRNA generally expedites decoding on the ribosome (5) and boosts the cellular translational efficiency under limited ribosomes (6). This benefit is greater when preferred codons are in high-concentration mRNAs than when they are in low-concentration mRNAs, because the former occupy more ribosomes. Consequently, selection for translational efficiency should cause highly transcribed genes to acquire high CBIs in evolution—another instance of process 2.

Translational elongation can sometimes become the rate-limiting step of protein synthesis (7). Hence, the cellular demand for more proteins from a gene can select for both higher transcription and higher translational efficiency (thereby higher CBI), creating process 3. Additionally, preferred codons can stabilize mRNAs in a translation-dependent manner (8, 9). So, selection for a higher mRNA concentration of a gene favors both mutations in the promoter that enhance transcription and synonymous mutations from unpreferred to preferred codons that reduce mRNA degradation, creating a positive correlation between nuclear mRNA level and CBI via process 3.

Zhao et al. (1) claim to have identified 18 genes that "mediate the codon usage effects on transcription," but these genes may simply globally disturb mRNA levels upon deletion, uncoupling the relationship between the mRNA level and CBI observed in wild-type cells regardless of the specific processes that established the relationship in the first place. Therefore, findings about these genes are the consequences of the screening method and may not inform whether and how CBI influences transcription. Furthermore, although the same group previously demonstrated that using preferred codons enhanced transcriptions of two reporter genes independently of the GC content (10), the differential effect of deleting *set-2* on the mRNA levels of these reporters in the present study (1) could still be due to the GC content instead of codon usage per se. In summary, Zhao et al. (1) fail to establish an influence of codon usage on transcription at the genomic scale, let alone its general mechanism of genetic regulation.

1 F. Zhao *et al.*, Genome-wide role of codon usage on transcription and identification of potential regulators. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2022590118 (2021).
2 H. Akashi, Synonymous codon usage in *Drosophila melanogaster*: Natural selection and translational accuracy. *Genetics* **136**, 927–935 (1994).

[a]State Key Laboratory of Plant Genomics, Institute of Genetics and Developmental Biology, Innovation Academy for Seed Design, Chinese Academy of Sciences, Beijing 100101, China; [b]University of Chinese Academy of Sciences, Beijing 100049, China; and [c]Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109

Author contributions: W.Q. and J.Z. wrote the paper.
The authors declare no competing interest.

[1]To whom correspondence may be addressed. Email: wfqian@genetics.ac.cn or jianzhi@umich.edu.
Published May 3, 2021.

3 D. A. Drummond, C. O. Wilke, Mistranslation-induced protein misfolding as a dominant constraint on coding-sequence evolution. *Cell* **134**, 341–352 (2008).
4 J. R. Yang, B. Y. Liao, S. M. Zhuang, J. Zhang, Protein misinteraction avoidance causes highly expressed proteins to evolve slowly. *Proc. Natl. Acad. Sci. U.S.A.* **109**, E831–E840 (2012).
5 D. E. Weinberg *et al.*, Improved ribosome-footprint and mRNA measurements provide insights into dynamics and regulation of yeast translation. *Cell Rep.* **14**, 1787–1799 (2016).
6 W. Qian, J. R. Yang, N. M. Pearson, C. Maclean, J. Zhang, Balanced codon usage optimizes eukaryotic translational efficiency. *PLoS Genet.* **8**, e1002603 (2012).
7 D. Chu *et al.*, Translation elongation can control translation initiation on eukaryotic mRNAs. *EMBO J.* **33**, 21–34 (2014).
8 A. Radhakrishnan *et al.*, The DEAD-Box protein Dhh1p couples mRNA decay and translation by monitoring codon optimality. *Cell* **167**, 122–132.e9 (2016).
9 S. Chen *et al.*, Codon-resolution analysis reveals a direct and context-dependent impact of individual synonymous mutations on mRNA level. *Mol. Biol. Evol.* **34**, 2944–2958 (2017).
10 Z. Zhou *et al.*, Codon usage is an important determinant of gene expression levels largely through its effects on transcription. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E6117–E6125 (2016).