Genome Watch

Whisperings from not so silent mutations

Nicolas Grosjean and Ian K. Blaby

This month's Genome Watch explores how synonymous mutations may have wider implications for protein function than previously acknowledged.

The accessibility and reduced cost of both chemically synthesized DNA and genomic data from environmental samples has resulted in an enormous untapped resource for sustainable biotechnologies and functional characterization. While genomic data is broadly accessible to the community through a number of databases, DNA synthesis has become an invaluable tool that provides the means for functional interrogation of these genomic blueprints without ever needing to grow the organism, or indeed access the native physical DNA[1]. Consequently, this approach has become a crucial tool for investigating unculturable microorganisms that are estimated to represent 99% of microbial life.

Yet, despite advances, chemical DNA synthesis comes with some complications: not every conceivable sequence can be easily made. To help overcome synthesis violations, such as repeats and homopolymers (as well as to "codon optimize" for expression in a given host organism), codon shuffling is often employed to alter the DNA sequence without affecting the amino acid sequence. This codon shuffling is enabled by redundancy in the genetic code – i.e. most amino acids are encoded by multiple triplet codons, resulting in the existence of synonymous codons. Since the functional machinery – the amino acid sequence – remains unaltered, it seems a reasonable inference that a protein would behave identically irrespective of the encoding DNA sequence.

Jiang et al.[2] reshape our thinking on codon usage by providing evidence of an additional layer in the regulation of protein translation processes and protein folding. Through computational models, they demonstrate that synonymous mutations can affect protein structure and function. In a previous report, the use of faster translating codons resulted in a 20% decrease in the specific activity of CAT-III from *Escherichia coli*. By permuting synonymous codons to achieve theoretically fast and slow-translated CAT-III, *Jiang et al.* were able to qualitatively recapitulate previously experimentally observed changes. They extended their analysis to a wider range by virtually screening proteins to detect candidates likely sensitive to changes in the translation speed. Both slow and fast translation resulted in the generation of topological trapped entangled structures/states depending on the protein. These near-native entangled proteins display lower catalytic efficiencies than their native state counterparts and populate heterogeneous protein pools that could have noticeable impacts on cellular function and phenotype. While some of the entangled structures slowly convert to their native states, deep entangled states can last for several minutes, which is considerable for an organism that has a generation time of 20 minutes such as *E. coli*.

This study was released just a few months after *Shen et al.*[3] investigated the *in vivo* impact of synonymous mutations in yeast. The authors demonstrated that synonymous mutations are strongly non-neutral and, similarly to non-synonymous mutations, can have a severe impact on an organism's fitness when introduced extensively throughout an organism's entire genome. While *Shen et al.* bring forward the destabilising effects of alternative codon on mRNA stability and translation, their results suggest an additional mechanism that could be linked to *Jiang et al*. discovery. These two studies combined show how codon usage can regulate protein expression and achieve optimal protein structure and function.

More work will be needed to understand the underlying mechanisms and broader significance. Clearly, codon refactoring does not abolish or severely limit the function of synthetic gene sequences, as demonstrated by thousands of investigations. However, once fully understood, this phenomenon could provide the potential to manipulate the activity of engineered enzymes to some degree. Further research utilizing both native and refactored DNA will provide the means to investigate the connection between codon-usage efficiency and protein-activity optimization, demonstrating the importance of synthetic biology approaches as a key field of research for sustainable bioenergy applications by metagenome mining and functional genomics.

Nicolas Grosjean and Ian K. Blaby
DOE Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
e-mail: ikblaby@lbl.gov
https://doi.org/10.1038/s41579-XXX-XXXX-X

1. Harvey, A., Edrada-Ebel, R, & Quinn, R. The re-emergence of natural products for drug discovery in the genomics era. *Nat. Rev. Drug. Discov.* **14**, 111–129 (2015).
2. Jiang, Y. *et al*. How synonymous mutations alter enzyme structure and function over long

timescales. *Nat. Chem.* (2022).
3. Shen, X. *et al*. Synonymous mutations in representative yeast genes are mostly strongly non-neutral. *Nature* **606**, 725–731 (2022).

Competing interests
The authors declare no competing interests.

Credit: Philip Patenall/Springer Nature Limited