Heredity (2009) 103, 189

www.nature.com/hdy

NEWS AND COMMENTARY

Molecular genetics

The genetic code has a 'shift' key

SC Le Comber

Heredity (2009) **103,** 189; doi:10.1038/hdy.2009.40; published online 22 April 2009

The genetic code has attracted the attention of scientists, including biologists, chemists, physicists and mathematicians for more than 40 years (Crick, 1968). Some of this work is among the most beautiful in all of biology—for example, Marshall Nirenberg's early experiments using poly-U RNA in cell-free systems to show that the triplet UUU in an RNA transcript codes for the amino acid phenylalanine (Nirenberg and Matthaei, 1961). This is true even of ideas subsequently shown to be incorrect, such as the comma-free code proposed by Francis Crick, Leslie Orgel and John Griffith—'the most elegant biological theory ever to be proposed and proved wrong' (Judson, 1979). Now a recent paper by Anton Turanov and Vadim Gladyshev at the University of Nebraska suggests that the code may have more surprises in store for us. The authors describe how, in the ciliate *Euplotes crassus*, a single codon, UGA, can code for either cysteine (Cys) or selenocysteine (Sec), even within the same protein, depending on its proximity to a 3′ untranslated region known as the selenocysteine insertion sequence (SECIS) (Turanov *et al*., 2009).

Selenocysteine insertion sequence is already well known from studies of bacteria. The SECIS element itself forms a stem-loop structure that is bound by the translation factor, SELB, forming a quaternary complex with the Sec tRNA and GTP (Walczak *et al*., 1996). However, the findings of Turanov *et al*. differ markedly from this earlier work, which showed that a single tRNA simply fails to distinguish between different amino acids (or between one amino acid and a stop codon), inserting one more or less randomly; for a review, see Namy *et al.* (2004). In contrast, *E. crassus* has two different tRNAs that recognize the same codon, one of which inserts a cysteine and one that inserts a selenocysteine. Selenocysteine is inserted when the UGA codon occurs close to the SECIS element; further upstream, cysteine is inserted instead. Crucially, the choice of which element is inserted is non-random; the science writer Ed Wong has likened the SECIS element to the shift key on a typewriter. Turanov *et al*.'s analysis of the *E. crassus* genome showed three tRNAs that can recognize the UGA codon: Sec tRNA, mitochondrial Trp tRNA and a novel Cyst tRNA; an additional Cyst tRNA that recognizes the codons UGU and UGC was also detected.

Despite the undoubted interest of these findings, some of the claims made for the paper's importance are difficult to support. The authors state in the paper's first sentence that 'strict one-to-one correspondence between codons and amino acids is thought to be an essential feature of the genetic code', a claim surely belied by earlier studies showing that codons can have dual meanings—studies that Turanov *et al*. themselves cite. The authors go on to claim that 'insertion of different amino acids into separate positions within nascent polypeptides by the same codeword is believed to be inconsistent with ribosome-based protein synthesis'. Again, earlier work rather undermines this claim. In any case, believing something is ubiquitous is not the same as believing it is required—and it is certainly not the case that any of these studies are 'inconsistent with ribosome-based protein synthesis'.

This slight tendency towards hyperbole is surprising, as the study is surely interesting enough without it. Anything that clarifies our knowledge of the genetic code is likely to be of interest to scientists in a wide range of disciplines, as it is difficult to overestimate the importance of the genetic code to modern biology. Perhaps the strongest evidence that all life on Earth shares a common origin is the fact that—with minor variants—all living organisms use the same genetic code. Studies looking for evidence of positive or purifying selection, or testing neutral theories of evolution, would be impossible if our knowledge of the code did not allow us to distinguish synonymous and non-synonymous mutations. Scanning genomic data for protein structural motifs to identify genic regions would similarly be extremely difficult.

The work described here raises immediately the question of how widespread this mechanism is—or whether other, similar mechanisms exist. The authors correctly point out that it will be important to establish whether it occurs in organisms with completely sequenced genomes. One of the more startling elements of this study is that the structural arrangements of *E. crassus* mRNA preserve this SECIS-dependent dual function in mammalian cells, suggesting that similar variant forms of the genetic code may actually be quite widespread in nature.

Famously, Orgel's Second Rule states that evolution is cleverer than you are. According to the philosopher Daniel Dennett, this rule was intended as a warning to biologists who were too quick to dismiss as design flaws biological features that they could not currently explain (Dennett, 1996). In fact, it can just as easily serve as a reminder that evolution tends to come up with good ideas before we do. Anderson *et al*. (2004) showed how artificial tRNAs that recognized quadruplet codons could be used to insert unnatural amino acids into proteins, and suggested that limitations on the number of possible triplets might not be a barrier to further expansion of the genetic code. The results reported by Turanov *et al*. are a beautiful demonstration of Orgel's Second Rule, and show how nature has already found a different way to expand the code—using a 'shift' key.

Dr SC Le Comber is at the School of Biological and Chemical Sciences, Queen Mary University of London, Mile End Road, London, E1 4NS, UK.

e-mail: s.c.lecomber@qmul.ac.uk

Anderson JC, Wu N, Santoro SW, Lakshman V, King DS, Schultz PG (2004). An expanded genetic code with a functional quadruplet codon. *Proc Natl Acad Sci USA* **101**: 7566–7571.

Crick FHC (1968). The origin of the genetic code. *J Mol Biol* **38**: 367–379.

Dennett DC (1996). *Darwin's Dangerous Idea*. Touchstone: New York, NY, 74p.

Judson HF (1979). *The Eighth Day of Creation*. Jonathan Cape: London, 315p.

Namy O, Rousset JP, Napthine S, Brierley I (2004). Reprogrammed genetic decoding in cellular gene expression. *Mol Cell* **13**: 157–168.

Nirenberg MW, Matthaei JH (1961). The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides. *Proc Natl Acad Sci USA* **47**: 1588–1602.

Turanov AA, Lobanov AV, Fomenko DE, Morrison HG, Sogin ML, Klobutcher LA *et al.* (2009). Genetic code supports targeted insertion oftwo amino acids by one codon. *Science* **323**: 259–261.

Walczak R, Westhof E, Carbon P, Krol A (1996). A novel RNA structural motif in the selenocysteine insertion element of eukaryotic selenoprotein mRNAs. *RNA* **2**: 367–379.

Editor's suggested reading

Carter CW (2008). Whence the genetic code? Thawing the 'Frozen Accident'. *Heredity* **100**: 339–340.

Rodin SN, Rodin AS (2008). On the origin of the genetic code: signatures of its primordial complementarity in tRNAs and aminoacyl-tRNA synthetases. *Heredity* **100**: 341–355.