Vol. 107, No. 1, 1982 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
July 16, 1982 Pages 225-228

POSSIBLE EVOLUTIONARY STEPS IN THE GENETIC CODE

Thomas H. Jukes
Space Sciences Laboratory
University of California, Berkeley
Berkeley, CA 94720

Received May 24, 1982

In mammalian mitochondrial codes, fourfold codons wobble-pair with UNN anticodons so that U wobbles with U, C, A and G. Twofold pyrimidine-terminated codons pair with GNN and twofold purine-terminated codons pair with UNN. These properties enable a prediction to be made for evolution of the universal genetic code. It was postulated (1) that an archetypal code of 16 quartets coded for 15 amino acids. If this code used UNN anticodons, then duplication of tRNA genes, followed by mutations in the anticodons and aminoacylation sites, would give rise to the present universal code.

The genetic code in mitochondria differs in certain respects from the universa genetic code (2, 3). One of the differences is that fourfold groups ("quartets") of codons such as GUN, valine, are translated by a single tRNA with an anticodon starting with U, such as UAC for valine, in which pairing takes place between U in the anticodon, wobbling with U, C, A and G in four codons (2). Also, amino acids having two pyrimidine-terminated codons such as phenylalanine, UUY, pair with an anticodon starting with G, such as GAA for phenylalanine, and amino acids having two purine-terminated codons, such as glutamine, CAR, pair with anticodons starting with U, such as UUG for glutamine. Both these examples follow classical wobble pairing rules (3).

These findings make it possible to postulate a pathway for expansion of the code to its present universal form. I suggested (4) that the universal genetic code might have evolved from an earlier archetypal code in which each quartet of codons was assigned to a single amino acid (Table 1). The archetypal code would include only 15 amino acids rather than the present 20. Embodied in the proposal was the suggestion that the archetypal amino acids were those that currently have pyrimidine-terminated doublets of codons, and have, in eight cases, "lost" the other two codons (one codon in the case of isoleucine) in a

0006-291X/82/130225-04$01.00/0

quartet to another amino acid or to a chain-terminating function. Thus, histidine was formerly coded by the quartet of codons that now codes for histidine and the "new" amino acid, glutamine.

If the archetypal code embodied the pairing procedure that the mitochondrial code uses for fourfold codons, it is possible to trace the steps by which the archetypal code expanded to form the present universal code. The first postulation is that an archetypal tRNA, for example, the archetypal $tRNA_{Phe}$, with anticodon UAA, duplicated to form two identical tRNAs. In one of these, the anticodon mutated from UAA to GAA, pairing with UUU and UUC. The second duplicate underwent mutational change in its aminoacylation site, so that it became charged with leucine, and retained the anticodon UAA. Thus tRNA paired with UUA and UUG. A change of this nature in amino acid acceptance evidently has taken place recently in yeast mitochondria. These use CUN for threonine instead of leucine, while Neurospora has retained the use of CUN for leucine as in the universal code (5).

By these fairly simple procedures, the expansion of the code took place to form the present universal code (Table 1). Subsequent steps would include additional duplications and mutations to produce, for example, four $tRNAs_{Val}$, three with anticodons GAC, CAC and IAC, respectively, and the fourth retaining anticodon UAC. $tRNA_{Trp}$ was formed when the Trp anticodon mutated from UCA to CCA without duplication of the tRNA and the anticodon UCA was therefore lost, so that UGA became a stop codon. The present $tRNA_{Met}$ was formed when a duplication of its former tRNA took place, and one of the duplicate tRNAs underwent a mutation from UAU to CAU. The other, which retained anticodon UAU, was "recaptured" by isoleucine, the amino acid originally coded by AUN.

The archetypal tyrosine anticodon did not duplicate, but mutated from UUA to GUA, pairing with UAU and UAC. The two codons UAA and UAG were left without an anticodon, and became chain-terminating codons because they had lost cognate tRNA.

An analogous event seems recently to have occurred in the mammalian mitochondrial code in which AGA and AGG have become chain-terminating codons and no tRNA pairing with them is in the human, bovine or mouse mitochondrial genomes

Table 1. Proposed Archetypal Genetic Code and Subsequent Evolution (From Ref. 1)

Quartet	Archetypal Assignments	Present Assignments
UUN	Phe	UUY-Phe UUR-Leu
CUN	Leu	No change
AUN	Ile	AUY-Ile AUA-Ile AUG-Met
GUN	Val	No change
UCN	Ser	No change
CCN	Pro	No change
ACN	Thr	No change
GCN	Ala	No change
UAN	Tyr	UAY-Tyr UAR-CT
CAN	His	CAY-His CAR-Gln
AAN	Asn	AAY-Asn AAR-Lys
GAN	Asp	GAY-Asp GAR-Glu
UGN	Cys	UGY-Cys UGA-CT UGG-Trp
CGN	Arg	No change
AGN	Ser	AGY-Ser AGR-Arg
GGN	Gly	No change

N = A, G, C or U; Y = U or C; R = A or G;
CT = chain termination

(6, 7). These codons are still used for arginine in yeast mitochondria (5). It is also noteworthy that mitochondria use both UGA and UGG as codes for tryptophan and both AUA and AUG for methionine, for these two examples correspond to earlier evolutionary stages in the universal code as proposed above.

ACKNOWLEDGEMENTS

Support from NASA Grant NGR 05-003-460 and the assistance of Carol Fegté are acknowledged.

REFERENCES

1. Jukes, T.H. (1968) in Molecules and Evolution (Second printing), Columbia Univ Press, N.Y., p. 68.

2. Barrell, B.G., Anderson, S., Bankier, A.T., de Bruijn, M.H.L., Chen, E.,

Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., Schreier, P.H., Smith, A.J.H., Staden, R. and Young, I.G. (1980) Proc. Natl. Acad. Sci. USA 77, 3164-3166.

3. Jukes, T.H. (1981) J. Mol. Evol. 18, 15-17.

4. Crick, F.H.C. (1966) J. Mol. Biol. 19, 548-555.

5. Bonitz, S.B., Berlani, R., Coruzzi, G., Li, M., Macino, G., Nobrega, F.B., Nobrega, M.P., Thalenfeld, B.E. and Tzagoloff, A. (1980) Proc. Natl. Acad. Sci. USA 77, 3167-3170.

6. Anderson, S., Bankier, A.T., Barrell, B.G., de Bruijn, M.H.L., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., Schreier, P.H., Smith, A.J.H., Staden, R. and Young, I.G. (1981) Nature 290, 457-470.

7. Bibb, M.J., Van Etten, R.A., Wright, C.T., Walberg, M.W. and Clayton, D.A. (1981) Cell 26, 167-180.