

Comment

Topological, statistical, and dynamical origins of genetic code
Comment on “A colorful origin for the genetic code: Information
theory, statistical mechanics and the emergence of molecular codes”
by T. Tlusty

Hiroaki Takagi^a, Kunihiro Kaneko^{b,*}

^a Department of Physics, School of Medicine, Nara Medical University, 840 Shijo-cho, Kashihara, Nara, 634-8521, Japan

^b Department of Basic Science, University of Tokyo, Komaba, Meguro, Tokyo 153-8902, Japan

Received 3 August 2010; accepted 4 August 2010

Available online 5 August 2010

Communicated by E. Shakhnovich

Keywords: Genetic code; Evolution; Dynamical systems; Endosymbiosis

How has the genetic code evolved to link genetic sequences to corresponding amino acids? Why are only 20 amino acids used from 64 possible triplet codes when forming proteins? Since the discovery of codons, these two questions have been integral to evolutionary biology, and addressed by distinguished biologists. For example, a beautiful theory by Crick et al. [1] to explain the number 20 from the optimal code without stopping symbol (comma) was selected as the most elegant ‘wrong’ theory ever to be proposed in science by Maynard-Smith [2]. Later on, Crick himself described the ‘universal’ genetic code as luck with the ‘frozen accident’ in the common ancestor of living organisms [3].

In the present paper [4], the author presented a statistical-physics based explanation for the evolution of the genetic code by using a more sophisticated mathematical theory than that of Crick et al. Topological and information-theoretic approaches were adopted by imposing the optimization of a certain quantity that consists of diverse amino acids, error-tolerance and minimal energy cost. A second-order phase transition from random to non-random mapping between codons and amino acids was found to coincide with changes in this quantity. This resembles known behavior in magnetic spin systems, which is analogous to the error-catastrophe seen in chemical evolutions [5]. Furthermore, the author introduced the topology of the codon graph determining the lowest modes of the graph-Laplacian and related it to the mathematical map coloring problem. By removing the ambiguity in the third position of the codon and considering the optimization process under this topological constraint, the maximum number of amino acids to be coded is predicted to be 25. This is an elegant theory that can be experimentally verified.

DOI of original article: [10.1016/j.plrev.2010.06.002](https://doi.org/10.1016/j.plrev.2010.06.002).

* Corresponding author. Tel./fax: +81 3 5454 6746.

E-mail address: kaneko@complex.c.u-tokyo.ac.jp (K. Kaneko).

Now, are these topological and statistical-physical arguments sufficient for understanding the evolution of genetic code? In general, it is hard to determine what quantity should be optimized through the evolution. Furthermore, it is not easy to answer whether the coding system can reach and stay near its optimal point through the system's evolution. To address these two points, a dynamical-systems-study of evolving cells is necessary. In fact, we have already simulated a model consisting of reaction processes among tRNA-amino acid complexes, aaRSs, and other enzymes and metabolites with variable mapping between codons and amino acids [6], finding that a stable genetic code with fixed mapping evolves through transcription and translation dynamics and cell–cell interactions.

The emergence of the fitness function and a possible mechanism to keep the genetic code system in the upper vicinity of its phase transition point, as postulated in the present paper, can potentially be discussed from such dynamical-systems point of view. In this regard, co-evolution between the reading matrix rij and the genetic code, as the author mentioned, is an important issue to be pursued. Along this line, endosymbiosis between two biological systems should cause a drastic change in chemical states for both systems, which might alter the reading matrix and the genetic code of each. This appears true in mitochondria, where non-universal coding is known to exist. A preliminary discussion about the causes of this coding is given in [6]. By combining the beautiful statistical-physics approach the author proposes with the dynamical-systems approach mentioned above, it is hoped that the evolutionary dynamics of genetic code will be elucidated and later verified with in-vitro artificial experiments.

References

- [1] Crick FHC, Griffith JS, Orgel LE. Codes without commas. *Proceedings of the National Academy of Sciences of the United States of America* 1957;43:416–21.
- [2] Maynard-Smith J. Too good to be true. *Nature* 1999;400:223.
- [3] Crick FHC. The origin of the genetic code. *Journal of Molecular Biology* 1968;38:367–79.
- [4] Tlsty T. A colorful origin for the genetic code: Information theory, statistical mechanics and the emergence of molecular codes. *Physics of Life Reviews* 2010;7(3):362–76, this issue.
- [5] Eigen M. Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften* 1971;58:465–523.
- [6] Takagi H, Kaneko K, Yomo T. Evolution of genetic codes through isologous diversification of cellular states. *Artificial Life* 2000;6:283–305.