Emergence of genetic coding through self-guided self-organisation

Peter R Wills

Department of Physics, Te Ao Marama Centre for Fundamental Inquiry, Centre for Computational Evolution, University of Auckland, Tamaki Makaurau, Aotearoa New Zealand

The origin of life depends on the establishment of a fundamental connection between information and matter whereby causation within emergent systems of genetic coding transcends the laws of physics and chemistry (quantum mechanics) and is taken over by the exigencies of local information processing (computation). Genetic coding provides a mapping from information stored as strings of nucleotides (DNA or RNA) onto individual three-dimensional molecular structures (folded proteins) that collectively uphold the tightly integrated networks of chemical functionality operating in living cells. The origin of this genetic coding is the quintessential "chicken-egg" problem in biology ("Which came first?") because the process of translation from the nucleotide alphabet of DNA/RNA genes into the amino acid alphabet of protein chains cannot operate without the products of its own operation. This problem of apparently circular causation exists not only in relation to the molecular components that operate the system of genetic coding but also in relation to the organisation of the genetic information without which any such system could or would exist. How was the capacity for chemical self-organisation, innate in the nucleic acid-dependent formation of peptide bonds between amino acids, guided autonomously to produce the highly specific, computationally self-organised map that we call "the genetic code"? The laws of physics contain no hint of their own representation, let alone the creative utilisation of such a representation. However, the representation of amino acid properties in terms of nucleotide triplets (codons) was established through the faculties offered by that representation itself. The "cheap trick" of entwining coding self-organisation and genetic selection in an event of dynamic symmetry-breaking provided the computational kernel from which coding was able to guide itself to the limit of resolution of its chosen means of operation: the use of strings of amino acids from a diverse alphabet as a means of controlling the chemical processes that differentiate between amino acids according to their digitally represented molecular properties. I will discuss collaborative investigations of this problem that combine theoretical, computational and experimental perspectives.