p-Genetic Code and Ultrametric Bioinformation

Branko Dragovich[1,2]

[1]Institute of Physics, University of Belgrade, Belgrade, Serbia
[2]Mathematical Institute, Serbian Academy of Sciences
and Arts, Belgrade, Serbia

Abstract

The genetic code is a mapping from the set of 64 codons onto the set of 20 amino acids and one stop signal. The codons are ordered triplets composed of the nucleotides cytosine (C), adenine (A), uracil (U) (or thymine (T)), guanine (G) and they are contained in the genes. The amino acids are building blocks of the proteins. The vertebrate mitochondrial (VM) code is rather simple and the other genetic codes can be considered as its slight modifications. In the VM code, an amino acid is coded by one, two or three codon doublets. When two codons code the same amino acid, they are similar (close) in the informational sense. We argue that the p-adic (ultrametric) distance is a simple, natural and adequate mathematical instrument to measure the bioinformation similarity (closeness). In particular, we show that the set of codons and the set of amino acids are p-adic ultrametric spaces and that the VM code is an ultrametric network. We also demonstrate that the (p-adic) ultrametrics is useful in investigation of informational similarity between sequences of nucleotides, codons and amino acids. This is a review talk with some new results.