The distribution of coding and non-coding segments along genomic sequences

Nestor Norio Oiwa, Carla Goldman
Instituto de Física/USP

We extend the statistical analysis reported elsewhere (Oiwa and Goldman, 2000) to characterize the distribution of coding and non-coding segments along genomic sequences using both linear and nonlinear methods. The present study includes new data now available at GenBank. The results confirm early indications of the presence of long-range order and self-similar structures in sequences of high eukaryotes. In addition, they show that, within statistical errors, the power decay of two-point correlation functions in these cases are either alike to each other or substantially different indicating a criteria for group classification. This suggests that the presence of non-coding regions among coding parts of high eukaryotic, might have an important role for organizing the genome in clusters with well-defined scaling laws.

Apoio: FAPESP