On the first law of genomics

Ryszard Rudnicki

Institute of Mathematics,
Polish Academy of Sciences,
Bankowa 14, 40-007 Katowice, Poland.
email: rudnicki@us.edu.pl

In 1998 on the conference *Microbial Genomes* Slonimski [1] formulated a low concerning the distribution of genes in a genome. Having several complete sequences of microbial genomes he divided genes into classes. He put into class n genes which appear exactly n-times in a genome. He observed that if x_n is a number of types of genes in a class n then

$$x_n \approx \frac{c}{2^n n}.$$

At the same time Huynen and Nimwegen [2] have compared the distribution of genes in eleven genomes and claim that x_n show power distribution. Recently, Koonin's group has published a few papers (e.g. [3], [4]) where they developed a simple model of this phenomena. In my talk I am going to present some general model in which we do not assume that the number of genes of a given type is bounded. The model is described by a system of infinite number of ordinary differential equations. We analyse asymptotic properties of the distribution of numbers of classes. We also compare theoretical results with the previous empirical studies. Our model rather justifies the Slominski's conjecture than that of Huynen and Nimwegen.

References

[1] P.P. Slonimski, M.O.Mosse, P. Golik, A. Henaut, Y. Diaz, J.L. Risler, J.P. Comet, J.C. Aude, A. Wozniak, E. Glemet, J. Codani, The first laws of Genomics, *C*onference: Microbial genomics II, Hilton Head, January 1998.

[2] M.A. Huynen, E. van Nimwegen, The frequency distribution of gene family sizes in complete genomes, *M*ol. Biol. Evol., **15** (1998), 583–589.

[3] E.V. Koonin, Y.I. Wolf, G.P. Karev, The structure of the protein universe and genome evolution, Nature, **420** (2002), 218–223.

[4] G.P. Karev, Y.I. Wolf, E.V. Koonin, Simple stochastic birth and death models of genome evolution: was there enough time for us to evolve? *B*ioinformatics, **19** (2003), 1889–1900.