Can a Minimal Gene Set for Cellular Life be Deduced from Comparisons of Completely Sequenced Bacterial Genomes?

Dagan, T.*, Mintz, S. and Graur, D.

Department of Zoology, George S. Wise Faculty of Life Sciences, Tel Aviv University

The 'minimal genome' approach attempts to estimate the smallest number of genetic elements sufficient to build a modern-type free-living cellular organism. There are two main approaches for inferring the minimal gene set for cellular life: the experimental and the analytical. The experimentalists use mutagenesis techniques in order to find which genes are essential to sustain cellular life, while the theoreticians infer the minimal-genome-set from known genomes of cellular organisms, i.e., lists of genes that are sufficient to sustain the life of a cell.

In this study, we follow the analytical approach. We estimate the minimum number of genes sufficient to sustain cellular life from the genomes of 30 bacteria species using bioinformatics tools. In our analysis, we make a distinction between free-living bacteria and parasite bacteria since parasitism invariably entails loss of genetic functions in the parasite and a consequent reduction in genome size. We also test the influence of the number of genomes used for the estimation on the final minimal-genome-set size.

Our results show that the minimal-genome-set of free-living bacteria is two times bigger than that of the parasite bacteria. However, the size of the minimal-genome-set is strongly dependant on the number of genomes in the analysis, the more genomes we use the smaller minimal-genome-set we obtain. This notion brings us to the conclusion that the assessment of the analytical approach is strongly biased by the number of genomes used in the analysis and therefore should be treated with caution.

*e-mail: tali@kimura.tau.ac.il