

COMPUTER PREDICTION OF BIOLOGICAL FUNCTIONS OF THE BACTERIAL GENES

Mikhail A. Golyshev¹, Eugene V. Korotkov^{1,2}

¹Bioinformatics laboratory, Centre of Bioengineering, Russian Academy of Sciences,
Prospect 60-letiya Oktyabrya 7/1, Moscow, 117312 Russian Federation

²Cybernetics department, National Research Nuclear University "MEPhI",
Kashirskoe shosse 31, Moscow, 115409 Russian Federation
Email:bioinf@yandex.ru

Over the last years a great number of bacterial genomes were sequenced. Now one of the most important challenges of computational genomics is the functional annotation of nucleic acid sequences. In this study we presented the computational method and the annotation system for predicting biological functions using phylogenetic profiles. The phylogenetic profile of a gene was created by way of searching for similarities between the nucleotide sequence of the gene and 1204 reference genomes, with further estimation of the statistical significance of found similarities. The profiles of the genes with known functions were used for prediction of possible functions and functional groups for the new genes. We conducted the functional annotation for genes from 104 bacterial genomes and compared the functions predicted by our system with the already known functions. For the genes that have already been annotated, the known function matched the function we predicted in 63% of the time, and in 86% of the time the known function was found within the top five predicted functions. Besides, our system increased the share of annotated genes by 19%. The number of false positives is less than 0.01%. The developed system may be used as an alternative or complementary system to the current annotation systems. The address of the system is: <http://genefunction.ru>.