

Output Organizer - a software to facilitate POTION results interpretation

Mariana Teixeira Dornelles Parise¹, Douglas Parise¹, Marcus Vinicius Canário Viana², Anne Cybelle Pinto Gomide², Vasco Ariston de Carvalho Azevedo¹

1 UFMG

Abstract

Positive selection studies have been used to identify genes involved in the emergence of new phenotypic traits, speciation and host-pathogen interaction. The automatic detection of positive selection can be performed using POTION software, which is an end-to-end pipeline to genome-scale analysis. This software allows users to configure and run the analysis easily and quickly, but gives the results in many files. Although the pipeline gives the most important information in some files, examining the other result files for more details required to write a manuscript is time-consuming and may be error prone due to human errors. In order to facilitate researcher's information retrieval, a software to organize and summarize the most relevant information in the POTION results was created. This solution was developed using Java in the NetBeans IDE. It receives the POTION log and positive.out files, the directory containing the intermediate files of the analysis, a file containing gene homology relations in the OrthoMCL 1.4 format, the folder which contains the nucleotide fasta files and the number of the organism genetic code. Using the given files, the software gives to the user an overview showing eight key features of the analysis and a table showing how many genes were pre-filtered and the reasons why. In addition, the program creates a file for each positive selected group showing the significant statistic model, the Bayes Empirical Bayes results table and a detailed table for each codon position under selection. The presented software facilitates the interpretation of the POTION results, giving to the final user an easy, organized and brief overview of the analysis.

Funding: CAPES, CNPq, FAPEMIG