AUTOMATIC PIPELINE FOR ENVIRONMENTAL DNA SEQUENCE ANALYSES

PISA, G.; RAITTZ, R. T.; FAORO, H.; KADOWAKI, M. A. S.; MAGNANI, G. S.; SOUZA, E. M.;
PEDROSA, F. O.; CRUZ, L. M.
Universidade Federal do Paraná – Departamento de Bioquímica e Biologia Molecular

The microbial diversity of different environments has been accessed by direct DNA extraction, PCR amplification, gene library construction (e.g., 16S rDNA) and high throughput DNA sequencing. These data can be used to determine the species composition and the diversity of different environments. Consequently, attention has also been directed to computer and statistical analyses of the large amount of data produced, resulting in the development of softwares for sequence alignment, phylogeny and statistical analyses. We have developed a PERL script that automates the analyses of environmental libraries, in an UNIX environment. This script integrates the execution of a set softwares in a pipe line, from sequence alignment to graphical comparisons of DNA libraries. DNA sequences are aligned using ClustalW; phylogenetic distances are calculated using DNAdist (PHYLIP package); statistical analyses are performed using LIBSHUFF and graphics are generated by gnuplot. All the software used are freely distributed. The script described can be obtained by contacting the corresponding author.