BMC Bioinformatics

POSTER PRESENTATION

Open Access

Phylotree – a toolkit for computing experiments with distance-based methods for genome coevolution

Elissaveta Arnaoudova[1], Jerzy W Jaromczyk[1*], Neil Moore[1], Christopher L Schardl[2], Ruriko Yoshida[3]

Background

We have developed software called *Phylotree* as a toolkit for running experiments to study gene cophylogenies for genome evolution using distance-based methods. In particular, the toolkit has been instrumental in conducting processing-heavy experiments with the new "difference of means" statistical method. *Phylotree* was used to run experiments using simulated data as well as biological sequences of well known host and parasite species, and is distributed with data and configuration files allowing these experiments to be reproduced.

Materials and methods

Phylotree is organized as a collection of scripts. A single run takes as inputs two input alignments in Nexus format, and consists of a number of bootstrap iterations, comprising the following steps: 1) produce modified sequence alignments by bootstrapping the original alignments; 2) sample trees from the posterior distributions for each of the original and modified alignments using MCMC-based software (MrBayes or BEAST); 3) compute distances among the tree distributions, for example as an average of the pairwise distances between trees in the samples; and 4) check similarity criteria based on these distances. A full experiment consists of one or more runs with different parameters, tree distance methods, and input alignments.

Results and conclusion

Phylotree supports a number of distance methods for phylogeny trees, including path-difference, matrix tree distance, Robinson-Foulds distance and tree space geodesic distance. Additional distance methods may be used by supplying programs that take as input Newick-format tree files, and output distances in a simple text-based format. *Phylotree* can be run in two different modes: interactive (the user is prompted for experiment-specific parameters: number of bootstrap iterations, link files to be used, configuration files for the MCMC software, etc.), and command-line (parameters are specified on the command line or in a configuration file). Whichever mode is used, a configuration file is generated that allows the experiment to be easily reproduced.

The *Phylotree* package is open-source, and is distributed together with source code for the included third party software. It includes tools, documentation, and guidelines for easy installation and building of the executables on different platforms. The software is available from http://csurs.csr.uky.edu/phylotree/ .

Author details
[1]Department of Computer Science, University of Kentucky, Lexington, KY 40506, USA. [2]Department of Plant Pathology, University of Kentucky, Lexington, KY 40546, USA. [3]Department of Statistics, University of Kentucky, Lexington, KY 40506, USA.

* Correspondence: jurek@cs.uky.edu
[1]Department of Computer Science, University of Kentucky, Lexington, KY 40506, USA