VCS 0.2: Variable Coalescent Simulator

Alexei Drummond

October 23, 2008

Introduction

VCS is a simple command line program to simulate coalescent trees from an arbitrarily varying population size history specified as a piecewise-linear function.

Usage

The software is distributed as a Java Archive (jar file) and will run on any operating system that have Java 1.5+ installed.

The command line usage is

```
java -jar vcs.jar [options] <infile> <outfile>
```

The options are

- `-g <gentime>` sets the generation time to the given value. This is used to scale times in the input file into units of generations. The default value is 1.

- `-n <sampleSize>` sets the number of samples to generate the trees from. The default value is 50.

- `-p <popSizeScale>` sets the population size scale factor. This scale factor transforms the population sizes in the input file to effective population sizes. This is useful if the population size profiles are expressed, for example, as prevalences, so then scale factor would represent the effective numbers of hosts in the population of interest. The default value is 1.

- `-reps <reps>` the number of replicate simulations that will be performed. Each replicate will be separated in the output file by a comment line that labels the replicate, e.g. `#rep 0`.

- `-se <sampledEnd>` the time at which the last taxon is sampled, in the time scale provided by the input file. If this differs from the time specified be `-ss` option then the samples will be evenly spaced between the two times. The default value is the last time if `-f` is set and the first time otherwise.

- `-ss <sampleStart>` the time at which the first taxon is sampled, in the time scale provided in the input file. If this differs from the time specified be `-se` option then the samples will be evenly spaced between the two times. The default value is the last time if `-f` is set and the first time otherwise.

- `-f` specifies that the population size history proceeds forward in time. Otherwise the population size history is assumed to proceed into the past.

The file arguments are

- `<infile>` A whitespace-delimited plain text file. The first column contains the time and should be ascending from zero. Subsequent columns contain the population size histories, for which a tree will be simulated for each.

- `<outfile>` The file to which the trees will be written in NEWICK format.

Example infile

```
0.0000 0.0000143266 0.0008941020 0.0000143267 0.0008941025
0.0832 0.0000142609 0.0008927429 0.0000142609 0.0008927434
0.1665 0.0000142073 0.0008914258 0.0000142073 0.0008914263
0.2497 0.0000141426 0.0008896486 0.0000141426 0.0008896491
0.3329 0.0000140794 0.0008877067 0.0000140794 0.0008877073
0.4161 0.0000140179 0.0008856322 0.0000140179 0.0008856328
0.4994 0.0000139585 0.0008834640 0.0000139586 0.0008834646
```