

Explanation of dating analyses

Dominik Schrempf

June 28, 2022

Abstract

The most important files are `name.params.monitor` (open with *Tracer*) and `name.timetree.tree.combined` (open with *FigTree*), where `name` is the analysis name.

Let `name` be the analysis name.

1 Monitor files

`name.params.monitor` Monitor file of parameters; open with *Tracer*. The number of burn in iterations varies but is usually around 5000. Find the exact number of burn in iterations using:

```
1 grep "Burn in for " name.mcmc.log
```

`name.prior.monitor` Monitor file of prior; have a look at the separate contributions of the prior components.

`name.timetree.monitor` Monitor file of the absolute or relative time tree. Each branch stores the absolute or relative time between speciations. Times are absolute if at least one node age calibration is available.

`name.ratetree.monitor` Monitor file of the relative rate tree. Each branch stores the relative rate used on that branch.

2 Time tree files

The following files provide specific as well as summary statistics of the time tree.

`name.timetree.ages.full` Table of node ages. Each row shows the ages of a specific node across iteration. The nodes are integer indexed; leaf names are shown when available. See the description of file `name.meantree.index` below.

`name.timetree.ages.summary` Table of node ages. Each row shows the mean age and the variance of the age of a specific node. Further, the minimum ages, the maximum ages, and the 95% confidence intervals are shown.

`name.timetree.monitor.nex` Nexus file of the tree monitor; can be directly opened with *Densitree*.

`name.timetree.monitor.thinned` Thinned monitor file. Only every fifth iterations is printed. Used to create the Nexus file for *Densitree*.

`name.timetree.tree.mean` Tree with node ages denoting the means of the node age across iterations (open with *FigTree*).

`name.timetree.tree.variance` Tree with node ages denoting the variances of the node ages across iterations (open with *Figtree*).

`name.timetree.tree.combined` Tree with node ages denoting the means of the node ages across iterations, and with 95 percent confident intervals (open with *FigTree*). Calculation of the confidence intervals assumes normal distributed node heights.

The same files (excluding the summary statistics about node ages) are also available for the rate tree.

3 Other files of interest

`name.meantree` Tree with branch lengths storing the average number of substitutions per site inferred by Phylobayes.

`name.meantree.index` Same as `name.meantree` but the node names are the node indices. I assign indices in pre-order (sometimes called depth-first order). That is, I count the root first, then the sub trees from left to right.

4 Log files

`name.prepare.log` Log file of preparation script.

`name.run.log` Log file of run startup script.

`name.mcmc.log` Log file of MCMC sampler. Have a look at the proposals, the acceptance rates, and more.

5 Internal files related to the MCMC sampler

`name.data` Internal data (mean branch lengths measured in average number of substitutions per site, covariance matrix, and so on).

`name.mcmc.mhg` Stored parameters of the MHG (Metropolis-Hastings-Green) algorithm; used when continuing or restarting analyses with the MHG algorithm.

`name.mcmc.setting` Stored settings such as the analysis name, the burn in configuration and more.

6 Notes

The calibration and constraint names include the boundaries, so they can be nicely viewed in Tracer at the horizontal axis label. For example, when opening the file name.params.monitor, the age distribution of a node calibrated by a calibration may look like this:

