

BAYESIAN DIVERGENCE TIME ESTIMATION IN BEAST

Tracy Heath

Department of Integrative Biology, University of California, Berkeley

2013 Workshop on Molecular Evolution
Český Krumlov, Czech Republic

DIVERGENCE TIME ESTIMATION SOFTWARE

Program	Models/Method
r8s	Strict clock, local clocks, NPRS, PL
ape (R)	NPRS, PL
multidivtime	log-n autocorrelated (plus some others)
PhyBayes	OU, log-n autocorrelated (plus some others)
PhyloBayes	CIR, white noise (uncorrelated) (plus some others)
BEAST	Uncorrelated (log-n & gamma), local clocks
TreeTime	Dirichlet model, CPP, uncorrelated
MrBayes 3.2	CPP, strict clock, autocorrelated, uncorrelated
DPPDiv	DPP, strict clock, uncorrelated

Bayesian **E**volutionary **A**nalysis **S**ampling **T**rees

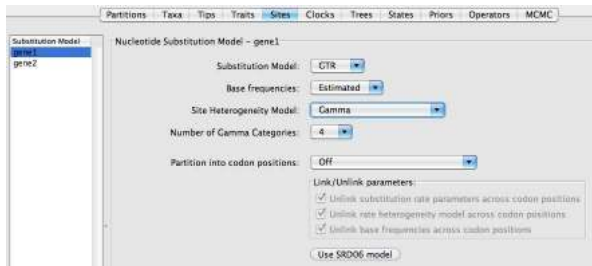
- population size
- growth/decline in population
- bottlenecks/transition points
- gene trees/species trees
- virus transmission dynamics
- recombination
- migration
- founder effects
- epidemiological tracking
- phylogeography
- trait evolution
- dates of MRCAs
- lineage rates
- ancestral character state reconstruction
- times of bottlenecks/transitions

Bayesian **E**volutionary **A**nalysis **S**ampling **T**rees

- free, open-source, cross-platform software package for estimating evolutionary parameters on rooted trees
- includes several utility programs for creating input files and summarizing output
- relies on a verbose XML syntax for executing analysis

BEAUti

- a GUI, utility program for generating properly formatted BEAST XML input files



BEAST XML: The eXtensible Markup Language

- specifies the sequences, node calibrations, models, priors, output file names, etc.
- dataset-specific issues can arise and some understanding of the BEAST-specific XML format is essential for troubleshooting
- there are a number of interesting models and analyses available in BEAST that cannot be specified using the BEAUti utility
- XML syntax help: http://beast.bio.ed.ac.uk/XML_format

```
<!-- site model
<siteModel id="gene2.siteModel">
  <substitutionModel>
    <gtrModel idref="gene2.gtr"/>
  </substitutionModel>
  <gammaShape gammaCategories="4">
    <parameter id="gene2.alpha" value="0.5" lower="0.0"/>
  </gammaShape>
</siteModel>
```

BEAST (analysis program)

- reads the commands in the xml input file
- performs MCMC and generates output files
 - `<file_stem>.trees` (contains the trees and branch rates for every n generations)
 - `<file_stem>.log` (contains the parameter samples for every n generation)
- the main BEAST binary only runs a single chain, for MCMCMC use BEASTMC3

LogCombiner

- combines the log or trees files from multiple, independent MCMC runs

File type: Tree Files

☐ Convert numbers from scientific to decimal notation

☐ Renumber output states

☐ Resample states at lower frequency:

Select input files:

File	Burnin
whme_100k.1.trees	5000
whme_100k.2.trees	5000

TreeAnnotator

- reads in the trees file and summarizes the topology, branch times, and rates
- places annotations on the tree that can easily be viewed in FigTree
- SumTrees in the Dendropy package (Sukumaran & Holder, 2010) is an alternative to TreeAnnotator. SumTrees is a richer program and offers more options for summarizing topology and branch parameters

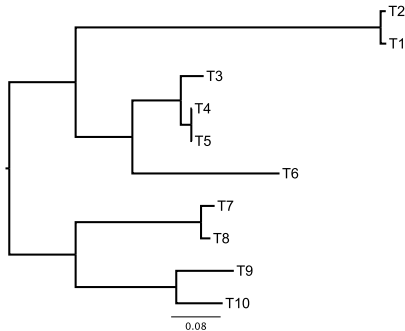
BEAST TUTORIAL



- Walkthrough using BEAUTi and executing the analysis in BEAST
- Independently complete the tutorial and summarize the BEAST output

BEAST TUTORIAL

This tutorial uses a simulated dataset of 10 taxa with 4 calibration points



ML phylogeny

