

Bayesian Phylogenetic Analysis Using *BEAST*

Centre for Macroevolution and Macroecology

Tuesday 30 and Wednesday 31 March 2010

**Gould Wing Seminar Room
Division of Evolution, Ecology, and Genetics
Research School of Biology**

Tuesday 30 March: Introductory topics

9.00 – 9.45 **Introduction to Bayesian phylogenetic analysis**

Introduction to phylogenetic analysis
Introduction to Bayesian phylogenetic analysis

10.00 – 10.45 **Bayesian phylogenetic analysis in practice**

Markov chain Monte Carlo sampling
Interpreting the output of a Bayesian analysis
Problems with Bayesian methods

11.00 – 11.45 **Phylogenetic analysis using BEAST**

Relaxed molecular clocks and calibration
Handling intraspecific and interspecific data
Creating input files using *BEAUti*
Processing the output of *BEAST*

Wednesday 31 March: Advanced topics

14.00 – 14.45 **Reconstructing demographic history**

Reconstructing demographic history
Demographic models in *BEAST*
Skyline-plot methods
Sampling design for successful demographic inference

15.00 – 15.30 **Bayesian phylogeography**

Bayesian phylogeographic analysis
Applications of the model

16.30 **Happy Hour, Banks Wing Tea Room**

Abstracts

Introduction to Bayesian phylogenetic analysis

Rob Lanfear

This talk will provide a brief introduction to Bayesian phylogenetic analysis in the context of other available methods.

- **Introduction to phylogenetic analysis.** A brief outline of the various phylogenetic methods will be given, with a comparison between model-free methods (maximum parsimony) and model-based methods (neighbour-joining, maximum-likelihood, and Bayesian inference). The basic theory behind Bayesian phylogenetic methods will be introduced.
- **Introduction to Bayesian phylogenetic analysis.** Bayesian phylogenetic methods have two key components: the likelihood function and the prior. Despite sharing a basis in the likelihood function, there are important differences between maximum-likelihood and Bayesian methods.

Assumed: Basic knowledge of phylogenetic analysis.

Bayesian phylogenetic analysis in practice

Simon Ho

This talk will explain one of the key components of Bayesian phylogenetic analysis: estimating posteriors using Markov chain Monte Carlo sampling. Some of the problems with Bayesian methods will be described.

- **Markov chain Monte Carlo sampling.** MCMC simulation is an integral part of Bayesian phylogenetic analysis. This complex aspect of Bayesian analysis will be explained in detail.
- **Interpreting the output of a Bayesian analysis.** This part of the talk will describe how to interpret the output of a standard Bayesian phylogenetic analysis, which includes files containing estimates of parameters and trees.
- **Problems with Bayesian methods.** Some of the major criticisms of Bayesian phylogenetic methods will be described. One of the major concerns has been the influence of the priors, the choice of which can be somewhat subjective.

Assumed: Basic knowledge of Bayesian phylogenetic analysis.

Phylogenetic analysis using *BEAST*

Simon Ho

This talk will present some of the models available in *BEAST*, and will describe the processes of creating an input file and interpreting results.

- **Relaxed molecular clocks and calibration.** Estimating substitution rates and divergence times is an important component of many molecular evolutionary studies. Relaxed molecular-clock models in *BEAST* can allow for rate variation among lineages. In addition, a number of calibration techniques have been implemented in *BEAST*.
- **Handling intraspecific and interspecific data.** Population-level and species-level data need to be analysed using different approaches. For example, a coalescent-based approach needs

to be taken for intraspecific data, whereas we would expect species-level data to be described better by a speciation process.

- **Creating input files using BEAUti.** *BEAUti* provides a user-friendly, menu-based interface for creating input files for *BEAST*. The options available in the latest version of *BEAUti* can cater for the majority of standard *BEAST* analyses.
- **Processing the output of BEAST.** *BEAST* produces a number of output files that need to be processed independently using further software. This part of the talk will describe how to interpret this output.

Assumed: Moderate knowledge of Bayesian phylogenetic analysis.

Reconstructing demographic history

Simon Ho

The demographic history of a population leaves a signature in the DNA of its modern representatives. There are several Bayesian methods for estimating demographic history from a sequence alignment. The 'skyline plot' family of methods will be described here.

- **Demographic models in BEAST.** Various demographic models, such as constant size and exponential growth, are available in *BEAST* and can be tested using Bayes factors.
- **Skyline-plot methods.** The development of skyline-plot methods will be traced, beginning with the classic skyline, and continuing with the generalised skyline, Bayesian spline, Bayesian skyline, Bayesian skyride, and extended Bayesian skyline.
- **Sampling design for successful demographic inference.** Skyline-plot methods can only be used successfully with data sets that are relatively 'informative'. This part of the talk will cover the issues associated with selection of loci, selection of individuals, and data quality.

Assumed: Basic knowledge of Bayesian phylogenetic analysis.

Bayesian phylogeography

Simon Ho

Ancestral geographic reconstruction can be performed in a Bayesian framework, allowing the simultaneous estimation of phylogenetic relationships, node times, ancestral states, and demographic history. This has recently been implemented in *BEAST*, with the introduction of the geospatial model.

- **Bayesian phylogeographic analysis.** The theory and motivation behind this method, which models the diffusion process between discrete geographic regions, will be described.
- **Applications of the model.** The geospatial model has been used for several data sets. Two examples, involving African rabies virus and muskox, will be shown in order to demonstrate the capabilities of the Bayesian phylogeographic approach in *BEAST*.

Assumed: Moderate knowledge of Bayesian phylogenetic methods and *BEAST*.

Useful references

Introductory book chapters

- **Inferring Phylogenies**
Book by Joseph Felsenstein.
Chapter 18 “Bayesian Inference of Phylogenies” contains an introduction to Bayesian phylogenetic methods. Most of the chapter is devoted to criticising Bayesian methods, particularly with respect to the role of priors.

Introductory journal articles

- **Bayesian inference of phylogeny: a non-technical primer**
Archibald JK, Mort ME, & Crawford DJ (2003) *New Trends Plant Syst*, 2003: 187-191.
- **Bayesian inference of phylogeny and its impact on evolutionary biology**
Huelsenbeck JP, Ronquist F, Nielsen R, & Bollback JP (2001) *Science*, 294: 2310-2314.
- **Potential applications and pitfalls of Bayesian inference of phylogeny**
Huelsenbeck JP, Larget B, Miller RE, & Ronquist F (2002) *Syst Biol*, 51: 673-688.

Technical journal articles

Describes *BEAST* 1.4.6:

- **BEAST: Bayesian evolutionary analysis by sampling trees**
Drummond AJ, & Rambaut A (2007) *BMC Evol Biol*, 7: 214.

Introduces the relaxed molecular-clock models in *BEAST*:

- **Relaxed phylogenetics and dating with confidence**
Drummond AJ, Ho SYW, Phillips MJ, & Rambaut A (2006) *PLoS Biol*, 4: e88.

Describes the different calibration priors available in *BEAST*:

- **Calibrating molecular estimates of substitution rates and divergence times in birds**
Ho SYW (2007) *J Avian Biol*, 38: 409-414.
- **Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times.**
Ho SYW, & Phillips MJ (2009) *Syst Biol*, 58: 367-380.

Describes the skyline-plot methods in *BEAST*:

- **Skyline-plot methods for estimating demographic history from DNA sequences**
Ho SYW, & Shapiro B (in prep.) – available soon
- **Bayesian coalescent inference of past population dynamics from molecular sequences**
Drummond AJ, Rambaut A, Shapiro B, & Pybus OG (2005) *Mol Biol Evol*, 22: 1185-1192.
- **Bayesian inference of population size history from multiple loci**
Heled J, & Drummond AJ (2008) *BMC Evol Biol*, 8: 289.

Describes the geospatial phylogeographic method in *BEAST*:

- **Bayesian phylogeography finds its roots**
Lemey P, Rambaut A, Drummond AJ, & Suchard MA (2009) *PLoS Comp Biol*, 5: e1000520.
- **Phylogeography takes a relaxed random walk in continuous space and time**
Lemey P, Rambaut A, Welch JJ, & Suchard MA (2010) *Mol Biol Evol*, in press.

BEAST resources

- **BEAST Google Group (<http://groups.google.com/group/beast-users>)**
This site provides a forum allowing users to post questions about using BEAST.