

Bayesian Phylogenetic Analysis Using *BEAST*

Informal workshop – all welcome

**Friday 29 August
&
Friday 5 September**

Friday 29 August

Morning Session (Talks)

10.00 – 10.55 Introduction to Bayesian phylogenetic analysis

Introduction to phylogenetic methods
Introduction to Bayesian phylogenetic analysis
Introduction to Markov chain Monte Carlo sampling

Simon Ho, BoZo Seminar Room

11.05 – 12.00 Bayesian phylogenetic analysis in practice

Comparison of *MrBayes* and *BEAST*
Strict molecular clocks and relaxed molecular clocks
Calibrating estimates of rates and divergence times
Handling intraspecific and interspecific data

Simon Ho, RSBS Robertson Lecture Theatre

Afternoon Session (Workshop)

14.00 – 15.00 Phylogenetic analysis using *BEAST*

Creating and manipulating input files
Interpreting *BEAST* output
Basic troubleshooting techniques

Simon Ho, BoZo Meeting Room

15.00 – 17.00 Hands-on *BEAST* workshop*

Simon Ho and Matt Phillips, BoZo Meeting Room
Limited to 15 places; bring your own laptop and data set

Friday 5 September

14.00 – 17.00 (Repeat of Afternoon Session from Friday 29 August)

*Previous experience in phylogenetic analysis is *required*, and some amount of experience in Bayesian phylogenetic analysis is *desirable*. If you are unable to bring your own laptop: please contact Simon Ho (simon.ho@anu.edu) in advance to arrange for a desktop computer.

Description of the Morning Session

10.00 – 10.55 Introduction to Bayesian phylogenetic analysis

The basic theory behind Bayesian phylogenetic methods will be explained. This talk will begin with a brief introduction to available phylogenetic methods, with a comparison between model-free methods (maximum parsimony) and model-based methods (neighbour-joining, maximum-likelihood, and Bayesian inference).

Introduction to phylogenetic methods. The various phylogenetic methods will be briefly explained and compared.

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 very important differences between maximum likelihood and Bayesian methods. Priors play an important role in Bayesian methods, but they are frequently the subject of criticism.

Introduction to Markov chain Monte Carlo sampling. Markov chain Monte Carlo simulation, or MCMC for short, is an integral part in the implementation of Bayesian phylogenetic methods. This complex aspect of Bayesian analysis will be explained in some detail.

Assumed: Basic knowledge of phylogenetic methods.

11.05 – 12.00 Bayesian phylogenetic analysis in practice

Compared with maximum likelihood, Bayesian phylogenetic methods allow the use of rather complex evolutionary models. This talk will cover some of the capabilities of Bayesian methods.

Comparison of MrBayes and BEAST. *MrBayes* is the most widely used software for performing Bayesian phylogenetic analyses, while *BEAST* is rapidly growing in popularity. There are several key differences between these two programs. In addition, *BEAST* implements a considerable number of models that are not available in *MrBayes*, and vice versa.

Strict molecular clocks and relaxed molecular clocks. The introduction of Bayesian phylogenetic methods has allowed the development of sophisticated relaxed molecular-clock models. These Bayesian relaxed clocks will be explained.

Calibrating estimates of rates and divergence times. Estimating substitution rates and divergence times is an important component of many molecular evolutionary studies. Calibration techniques have been improving in realism and sophistication over the past few years, and the latest methods are only available in Bayesian phylogenetic software.

Handling intraspecific and interspecific data. There are some fundamental differences between population-level (intraspecific) and species-level (interspecific) sequence data, and they need to be analysed using quite different models. Both types of data can be analysed using Bayesian methods, either separately or, with some difficulty, in a combined manner.

Assumed: Basic to moderate knowledge of Bayesian phylogenetic methods.

Description of the Afternoon Session

Note: This session will be run in duplicate on 29 August and 5 September, for the benefit of those who are unable to attend one or the other.

14.00 – 15.00 Phylogenetic analysis using *BEAST*

This talk will cover the various aspects of a *BEAST* analysis, from creating an input file to interpreting the results.

Creating and manipulating input files. *BEAUti* is a program that provides a user-friendly, menu-based interface for creating input files for *BEAST*. Usage of *BEAUti* will be described in detail. Due to large range of evolutionary models available in *BEAST*, not all have been made available in the *BEAUti* menus. However, editing *BEAST* input files is not always straightforward. The components of a *BEAST* input file will be explained.

Interpreting BEAST output. *BEAST* produces a number of output files that need to be processed independently. This part of the talk will explain how to check whether the analysis has been run for a sufficiently amount of time (*i.e.*, ‘reached convergence’), and how to interpret the output produced in a typical *BEAST* analysis.

Basic troubleshooting techniques. Due to the complexity and flexibility of *BEAST*, it not uncommon for problems to be encountered. Users can run into problems when creating input files, when trying to start a *BEAST* analysis, or when attempting to process the results. Some basic troubleshooting techniques will be described.

Assumed: Moderate knowledge of Bayesian phylogenetic methods.

15.00 – 17.00 Hands-on *BEAST* workshop

This workshop is designed for researchers who have either already used *BEAST*, or intend to do so at some point in the near future. Participants should bring their own laptops, if possible. It is also recommended that participants bring their own data sets (although very large data sets should be avoided, for reasons of computational speed), although an example data set will be provided too. Participants will learn how to create and manipulate input files, run an analysis, and implement various models available in *BEAST*. All participants will be able to work at their own pace throughout the workshop. Whenever appropriate, demonstrations will be shown on the projector.

This is a partial list of *BEAST* analyses that can potentially be covered in this workshop:

- Inferring phylogenetic trees
- Divergence time estimation using fossil calibrations
- Divergence time estimation using a known mutation rate
- Bayesian skyline plot analysis (inferring demographic history)
- Comparing demographic models
- Estimating effective population sizes
- Model selection using Bayes Factors

Assumed: Basic practical knowledge of phylogenetic analysis.

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, and 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, and Bollback JP (2001) *Science*, 294: 2310-2314.

Potential applications and pitfalls of Bayesian inference of phylogeny

Huelsenbeck JP, Larget B, Miller RE, and 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 and 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, and 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.

Introduces the Bayesian skyline plot in *BEAST*:

Bayesian coalescent inference of past population dynamics from molecular sequences

Drummond AJ, Rambaut A, Shapiro B, and Pybus OG (2005) *Mol. Biol. Evol.*, 22: 1185-1192.

BEAST resources

A Rough Guide to *BEAST* 1.4

http://beast-mcmc.googlecode.com/files/BEAST14_Manual_6July2007.pdf

This is the closest thing to a *BEAST* manual.

BEAST Google Group

<http://groups.google.com/group/beast-users>

This site enables users to post questions about using *BEAST*. The questions are frequently answered by the authors of *BEAST* (Alexei Drummond and Andrew Rambaut).