

Proposed Workshop in Phylogenetic Methods for Estimating Species Trees for OSNA meeting, San Diego, Feb. 2010

Scott Edwards and Liang Liu, Museum of Comparative Zoology, Harvard University

September 2009

I propose to lead a computer workshop focused on new methods for estimating phylogenetic trees from multilocus DNA sequence data. A new class of phylogenetic methods is emerging for the estimation of so-called 'species trees', trees of species, populations and lineages that can be estimated by combining the information from multiple, independently segregating genes or DNA sequences. A number of computer software packages, several developed for my group, have emerged in recent years and are increasing in their sophistication and popularity. The workshop will focus on implementing and running these software packages and will allow participants to analyze their own data sets if they so choose or we can provide example data sets.

Specifically, we propose to demonstrate running of two species tree methods: Bayesian Estimation of Species Trees (or BEST)¹, and Species Trees Estimated by Average Ranks of Coalescence Times (or STAR)². The former is a stand-alone package integrated into the popular phylogenetics software mrBayes, whereas the latter is a suite of modules in R collected in the online package phybase³. No familiarity with R is required, but some basic familiarity with mrBayes and Bayesian analysis will be assumed. The STAR module will also incorporate a discussion/demonstration of the multilocus bootstrap, a technique whereby both genes (partitions) and sites within those partitions are bootstrapped to obtain confidence in multilocus estimates of phylogeny.

The workshop will begin with a 1-hour lecture of concepts behind species trees by Edwards and Liu, followed by questions. Then we will move into demonstration of data formats and analysis of multilocus data sets. Students will be encouraged to bring their own data. We will spend most of the time getting data sets to run in these two packages and then follow up with discussion.

References:

¹Liu, L., and D. K. Pearl. 2007. Species trees from gene trees: reconstructing bayesian posterior distributions of a species phylogeny using estimated gene tree distributions. *Systematic Biology* 56:504-14.

²Liu, L., L. Yu, D. K. Pearl, and S. V. Edwards. 2009. Estimating species phylogenies using coalescence times among sequences. *Systematic Biology* Online at 10.1093/sysbio/syp031

³<http://www.stat.osu.edu/~liuliang/research/phybase.html>