

Homework no. 4

Due Wednesday, March 2

Take the same data set you used for the previous homework (if possible) and do Bayesian inference of the phylogeny. Here are eight programs for this (some work only on DNA, only on RNA, or only on protein sequences): When this PDF is viewed on-line, the links below should be active and lead you to the programs' web sites.

MrBayes	<a href="http://www.mrbayes.net">http://www.mrbayes.net</a>
BEAST	<a href="http://beast.bio.ed.ac.uk">http://beast.bio.ed.ac.uk</a>
Bayes Phylogenies	<a href="http://www.evolution.rdg.ac.uk/BayesPhy.html">http://www.evolution.rdg.ac.uk/BayesPhy.html</a>
PHASE	<a href="http://www.bioinf.man.ac.uk/resources/phase/">http://www.bioinf.man.ac.uk/resources/phase/</a>
PhyloBayes	<a href="http://www.atgc-montpellier.fr/phylobayes/binaries.php">http://www.atgc-montpellier.fr/phylobayes/binaries.php</a> (a web server is near there too)
PAML	<a href="http://abacus.gene.ucl.ac.uk/software/paml.html">http://abacus.gene.ucl.ac.uk/software/paml.html</a>
BAMBE	<a href="http://www.mathcs.duq.edu/larget/bambe.html">http://www.mathcs.duq.edu/larget/bambe.html</a>
p4	<a href="http://p4.nhm.ac.uk/">http://p4.nhm.ac.uk/</a>

Alternatively, you can explore the other Bayesian programs in the Bayesian Inference section of the list of programs by Methods in my “Phylogeny Programs” web pages (use that phrase to find the pages).

- Make the models as similar as you can to the ones used in the likelihood analysis.
- Show a tree summarizing the results of Bayesian inference on your data, with clade posterior probabilities.
- Explain what prior on trees (and what priors on parameter values) were used. Don't just say “the default priors” but describe them.
- Compare the clade probabilities in the Bayesian analysis to the bootstrap  $P$  values that you found in homework #3.
- If there is a parameter in your model for transition/transversion ratio, have the program give it a prior, and allow it to vary during the run. What information did the program give you on the posterior on the value of this parameter?
- If you can, do more than one Bayesian run. Do they get similar results? How did you decide how long to run the Bayesian program?
- If you choose rather different priors, does this change the results?
- Comment on the biological reasonableness of the results, if possible.
- Comment on the programs you used, how well they worked and how easy they were to use.

- Report this to me in a short (3-5 pages or so) report. Show some results (including a tree with clade probabilities).

e-mail me (joe (at) gs.washington.edu) with a report in PDF or MS Word (.DOCX or .DOC) format. format on the results.