

Phylogeny computer program exercise

Due 3/8

Get the program DNATREE by using the links on the course web page, which take you to the DNATREE web page which has links for downloading.

Use the program to simulate two trees with different rates of evolution. For each of these runs, **use the same random number seed**. This way both will have the same true tree. Simulate two data sets, with a tree of 10 species and sequences of length 200 sites (these are the default settings). For two different values of rate of evolution, 0.1, and 2.0, simulate a tree (*and make sure to write it down*), then simulate a data set, and use parsimony to reconstruct the best estimate. Compare the estimate to the true tree.

Summarize the results by counting, for each internal branch in the unrooted version of the true tree (there should be 7 such branches), whether it is also present in the estimated tree. Note that the root of the tree is not estimated by the parsimony method (it is put in an arbitrary place). So the two branches coming from the rootmost fork of the tree are to be considered to really be one branch. An interior branch is one that does not have a tip species at either end. Look at each interior branch of the true tree and ask whether you see the same branch in the parsimony tree. A branch is counted as the same if it separates the species into the same two sets. Thus if a branch of one tree has species A, C, D, E connected to one of its ends, and species B, F, G, H, I, J to the other, it is considered the same branch as one with has BFGHIJ and ACDE (order of these two sets does not matter).

You may or may not see a clear pattern from your run, but we will be summarizing the results in class and on the web site and when they are all put together we should be able to see whether the higher rate of evolution leads to greater or less accuracy of the estimate of the tree.

Report briefly (1-2 pages) on the results.