

## Homework Project

due the last day of classes

Do the following:

1. Choose a data set of interest to you. If you can't think of any, you can use one of three of the four data sets that are available using the links on the course web page (the fossil elephants discrete character data set does not lend itself to all the kinds of analyses which we ask you to do below). The data set should not be too small – not less than 100 bases or amino acids, and not less than 15 sequences or species.
2. Make sure you have access to phylogeny computer programs. At the bottom of the course web page you will find links for my own free package PHYLIP as well as my comprehensive Phylogeny Programs web pages which list 195 packages and where you can get them, often with links for downloads.
3. Analyze the data set using parsimony, distance matrix, and likelihood (or Bayesian) methods.
4. Write a report explaining clearly what you did and what the results were.
5. For at least one of these methods, carry out bootstrapping or jackknifing (or Bayesian posterior probabilities) to investigate the statistical support for parts of the tree. If you can do some other kind of analysis of statistical support too, that would help.
6. Discuss what the assumptions of these methods are, and the extent to which these assumptions are met in those data.
7. Discuss the biological reasonableness of the resulting tree(s).