Name \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

This assignment will predominately cover ML and Bayesian modules, though you will need a parsimony tree made from a single gene in the Ruhfel et al. data set to complete all the questions. Many of the answers can just be typed below; in some cases you will need to provide a tree or a plot made from one of the programs. You can either paste them below (if the resolution is good enough) or zip everything up into one file and email them to me (jbl256@cornell.edu). The due date for all materials will be Friday, Nov 22nd by 5 pm.

1. What gene did you select for the analyses? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
2. From jModeltest, what was the best model based on AIC criteria? What was the second-best model? What was the difference in AIC values between the two?
3. Infer a tree using your gene in MrBayes. make sure you run the analyses long enough so that the tree fully converges (split frequency of 0.01). How many generations did this take? Attach a Tracer plot using both chains (run1.p and run2.p).
4. Run your gene of choice with the appropriate model of molecular evolution (GTR+Gamma is fine for RAxML) in both RAxML and MrBayes. Using the R script, do the following comparisons:

- parsimony vs ML

- parsimony vs Bayesian

- ML vs Bayesian

For each comparison, how many clades are shared? How many clades are unique in each data set for each comparison? Provide the PDF plot of each comparison.

1. Compare your RAxML tree from a single gene to the tree published in Ruhfel et al. using the entire data set (provided in the data folder for Day3; Ruhful\_et\_al\_full\_data\_MLwBS.tre). How many clades are shared between the two? Are the differences more towards the tips or towards the ancestral nodes? For the nodes that differ between the two trees, are the nodes in your single gene tree well supported (BS > 70)?
2. Use any of your single gene trees and perform an ancestral state reconstruction (simulating either discrete or continuous data). Based on your simulation, is your trait changing rapidly across the tree? Provide the plot from R.

**BEAST questions**

1. After running the tree dating in Beast for the 32-tip tree, how old are the angiosperms in your tree? How old is the most recent common ancestor of all land plants?
2. From the 32 tip subset data, does the root selected in your Beast analysis match the root location from the Ruhfel et al. 2014 paper? Does the tree match the accepted evolutionary history of plants?
3. In the tutorial we used a strict molecular clock, even though this may not be appropriate for the range of taxa in the 32-tip subset. Rerun the BEAST analysis with a different molecular clock that allows for variation of evolution rate across different branches. How does this tree differ from the tree produced with the strict molecular clock?