Lab 10: The What the Hell Do I Do with All These Trees Lab

Introduction

We've generated a lot of trees in the last few weeks. Today we're going to explore different ways to view, compare, and manipulate those trees. First we're going to use *PAUP** to generate several types of consensus trees. Then we're going to use *FigTree* to look at these consensus trees. Finally we're going to generate a consensus tree from trees that have overlapping but not identical taxa using Matrix Representation with Parsimony.

We're going to use a couple different files for this lab. Find them online at the IB200 Syllabus and Handouts page.

The goals of the lab are to:

      I.   Generate consensus trees in PAUP*
     II.   Manipulate a consensus tree in FigTree
    III.   Matrix Representation with Parsimony

Exercise 1: Generate Consensus Trees in PAUP*

We're going to generate several different consensus trees from the tree file **MrBayesCephalopodTprobs.nex** using *PAUP** which contains thirteen trees generated using *MrBayes*. I want to emphasize that this is not the appropriate way to generate a consensus tree from *MrBayes*. It is much better to use the **sumt** command in *MrBayes*, because that will consider the trees based on their estimated posterior probabilities and will also calculate branch lengths. However, there are many other situations when you would want to use this method, such as if you generate several most parsimonious trees. I'm just using this tree file, because it is convenient.

1. Download **MrBayesCephalopoTprobs.nex** from the web site. It contains the first 13 trees from the tprobs file for the cephalopod dataset that we have used in a previous lab. These are the most probable 50% of trees that were found during stationarity.

2. Open PAUP*.

3. Execute **MrBayesCephalopoTprobs.nex** in PAUP.

4. Change your wording directory to your folder for the day.

5. First let's generate a strict consensus of all the trees in memory. Type:

   ```
   contree all /strict=yes;
   ```

   This will output a tree that only contains nodes present in all your input trees.

1. When generating consensus trees, PAUP* will not hold the consensus trees in memory. If you want to keep the consensus trees, you have to save them to a file. Let's practice this now:

```
contree all /strict=yes treefile=mytree.tre;
```

This will save a consensus tree file to your working directory.

2. Now let's generate a Majority Rule tree with a cut off at 50%. This will output a tree with all the nodes that appear in more than 50% of the tree. It will also tell you in what percentage of those trees the nodes occurred.

```
contree all /strict=no majrule=yes percent=50 treefile=majrule50.tre;
```

Does this have the same topology as the strict consensus? Are they compatible?

3. Generate another Majority Rule tree, only this time up the cut off, so that you eliminate some clades. The cut off point is always kind of arbitrary, but cannot be less than 50%. If it were less than 50%, then you couldn't be sure that all the clades are compatible.

Question #1: **How high would the cut off have to be to guarantee that you are going to get the same tree as strict consensus?**

Exercise 2: Tree-viewing with FigTree (or Dendroscope)

By now you will have noticed that most of the programs that generate trees either don't print trees at all or make really crappy ones. Luckily there are a lot of different ways to view trees. Today we'll be using *FigTree* (which you have already used a few times.) There is another tree-viewing program called *Dendroscope*, which you might also want to check out. It is a bit more complicated that *FigTree*, but may have some options that are not available in *FigTree* (http://ab.inf.uni-tuebingen.de/data/software/dendroscope3/download/welcome.html)

1. Open *FigTree* and open the consensus trees that you just generated in PAUP* from the previous exercise.
2. Pull down the **Tip Labels** menu and change the font size, so that you can easily read the names of the taxa.
3. Click on the **Layout** item from the list of tools on the left. You can change your tree topology to be represented in rectangular, polar, and radial coordinates.
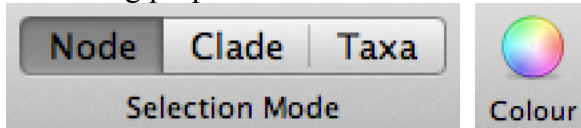

4. Change the Selection Mode to **Node** at the top of the window. Select a branch. See what cartoon, collapse, reroot, and rotate do.
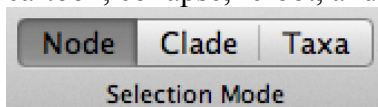
5. Change the Selection Mode to Clade at the top of the window. Select a clade. Click the Colour button and select a new color. This may be used to highlight a clade with

interesting properties or characters.

6. To generate a picture for use in your paper pull down the **File** menu and select **Export Graphic**. You can then select to save the tree with a Vector or Bitmap format. Notice the default place where it wants to save your file. You know enough about pathways at this point to change this to where you want your file to save. You can then import this type of file into Adobe Illustrator to spruce it up for final publication quality. If you save as a pdf, you can select Object → Clipping Mask → Release and then Object → Ungroup.

Did you know you can get Adobe Illustrator for free from the University?

Check it out: http://ist.berkeley.edu/software-central/

Question #2: **Send me a pdf file of your manipulated tree from FigTree.**

Exercise 3: Matrix Representation with Parsimony (MRP)

So it's easy to generate consensus trees if they all have exactly the same taxa, but what do you do if all the trees have different taxa? For example how would you put together a bunch of trees from different studies with overlapping but not identical taxa? Well, it is a matter of big debate. Maybe you shouldn't even do it at all. Maybe it is best to take the data matrices from all those studies and combine them into one supermatrix for analysis. If you do decide to combine trees it is not at all clear what the best method is. The mostly commonly used method is Matrix Representation with Parsimony (MRP). Here we're going to do a made up easy example of it.

We are going to do MRP on the three trees of rays that you will find on the next page. I just took a single data set of rays, randomly deleted two taxa from it three times, and used those reduced data sets to make trees by parsimony. This is a totally unrealistic situation for several reasons. The taxa have a lot of overlap. If theses were three trees picked from the literature they would have very little overlap. This would mean that the MRP matrix would have a lot of question marks. Also the trees are all generated from the same data set, so that you know that you won't have a contradictory signal from two different data sets, which you are likely to have in reality. However, I didn't have a time to find a more realistic set of trees, and these will make filling out the matrix easier.

1. Download the Ray matrix from the web site. Open it in Mesquite. This is just an empty matrix with the taxa names on it (and apparently lots of spelling mistakes), so that you don't have to bother filling them all in. You are just going to fill in the data.

2. Now code the three trees into the matrix. You do this by treating each interior branch from each tree as a separate character. Remember that every branch separates the taxa into two groups, one on each side of the branch. You can assign each of these partitions a separate character state, so that all the taxa on one side of a branch get a 0 and the other side a 1 for that character. All the taxa that don't appear in that tree should get a ?. Every tree should have 6 internal branches.

3. For example the branch that I marked as A in the first tree should be coded:

Raja polystigma   **1**
Raja montagui   **1**
Raja brachyura   **1**
Raja microocellata **1**
Raja asterias   **0**
Raja undulate   **0**
Raja radula   **?**
Raja clavata   **?**
Leucoraja meitensis 0
Leucoraja naevus   **0**
Leucoraja fullonica 0

4. When you're done with the matrix save it and close Mesquite.

5. Open PAUP* and open the MRP matrix in it.

6. Run an exhaustive search for the most parsimonious tree. Type:

   alltrees;

7. Did you get one tree? Was it fully resolved? Is there any homoplasy (which in this case would indicate a disagreement between the trees)?

Well that looks very pretty. It wouldn't be so pretty if I hadn't cheated.

Question #3: **Save a copy of this tree and send it to me.**

We have only touched on a few consensus methods that you will most commonly encounter in scientific papers and which you might employ for your final preojects. If you would like to know more about these types of consensus trees and many others, here is good resource:

http://www.mathnet.or.kr/mathnet/paper_file/McGill/Bryant/03ConsensusAMS.pdf

Leucoraja fullonica

Raja undulata

Raja asterias

Raja polystigma

Raja montagui

A

Raja brachyura

Raja microocellata

Leucoraja melitensis

Leucoraja naevus

Leucoraja fullonica

Raja undulata

Raja polystigma

Raja brachyura

Raja microocellata

Raja radula

Raja clavata

Leucoraja melitensis

Leucoraja naevus

Leucoraja melitensis

Leucoraja naevus

Raja undulata

Raja radula

Raja asterias

Raja clavata

Raja polystigma

Raja brachyura

Raja microocellata