

10kTrees - Exercise #2

Viewing Trees Downloaded from 10kTrees: FigTree, R, and Mesquite

The goal of this worked exercise is to view trees downloaded from *10kTrees*, including tree blocks. You may wish to complete Exercise #1, so that the names of files and figures correspond to what you will see on your computer screen and you have experience with the site itself. However, this is not strictly necessary because we provide the files for you on the site.

Underlying text files

The trees from *10kTrees* are provided as text files. These text files can be read by a variety of programs to view the trees and run analyses on the trees. Before doing anything else, let's take a look at the files you just downloaded, which are in "Nexus" format (which typically end with .nex or .nxs).

Open `treeblock_example.nex` in a text editor. Notice that it is declared to be a Nexus file, and then provides a tree block, which first includes translation of the species names into numbers.

As you scroll down the text file, notice that the trees are identified as "tree 1 =" etc., down to 200 trees (the number of trees downloaded in Exercise 1). The trees are shown in bracketed form (Newick tree format), which indicates the hierarchical structuring of the nodes; species are indicated with the numbers from the translation table, and branch lengths are given after colons.

The Nexus file format contains blocks of information. Note also that this Nexus file does not yet have a Data Block, but it does have a Tree Block. A Data Block can be added in Mesquite, see below, but for *R* and many other programs, we will simply provide a text file with the data, and the data will be linked to the phylogeny using species names.

Now, open the consensus tree saved as `treeconsensus.nex`. Notice that it has the same structure, but it contains only a single tree.

When you are finished examining the Nexus files, close them, and let's view the trees graphically.

FigTree:

FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>) is a very useful program for viewing and printing trees, especially for publication quality graphics. Open the program, and then open the `treeblock_example.nex` (your tree block of 200 trees). Be aware that it

could take several seconds for your tree file to open, so please be patient... if it takes more than a minute or two, however, something is probably wrong.

Chances are that the tree will need formatting, and many options exist in FigTree for doing this.

First, using the options in the left panel, let's change the "Tip Labels" to be larger and italicized (choose the "font" option, select Arial with font size 12 and italicized). Now you can actually see the names, which is obviously useful!

Next, let's add branch lengths to the tree. Use the "Branch Labels" option to show the branch lengths with two significant digits. Notice that you could also show node ages by selecting that option in the pull-down menu.

Further options are available for adjusting the line weight (under appearance) or the scale bar (you could remove the scale bar, given that branch lengths are shown).

In terms of the "layout" options, it is possible to show the tree in different formats. Try the unrooted tree option (the button to the far right under "Layout"). A bit hard to see... we can fix that by adjusting the "zoom" dial. You may also want to remove the branch lengths, as they appear to be cluttering our picture somewhat.

Let's go back to the rooted tree layout by clicking the button to the far left under "layout." For kicks, make the branches more "curved" with the curvature option at the top-left.

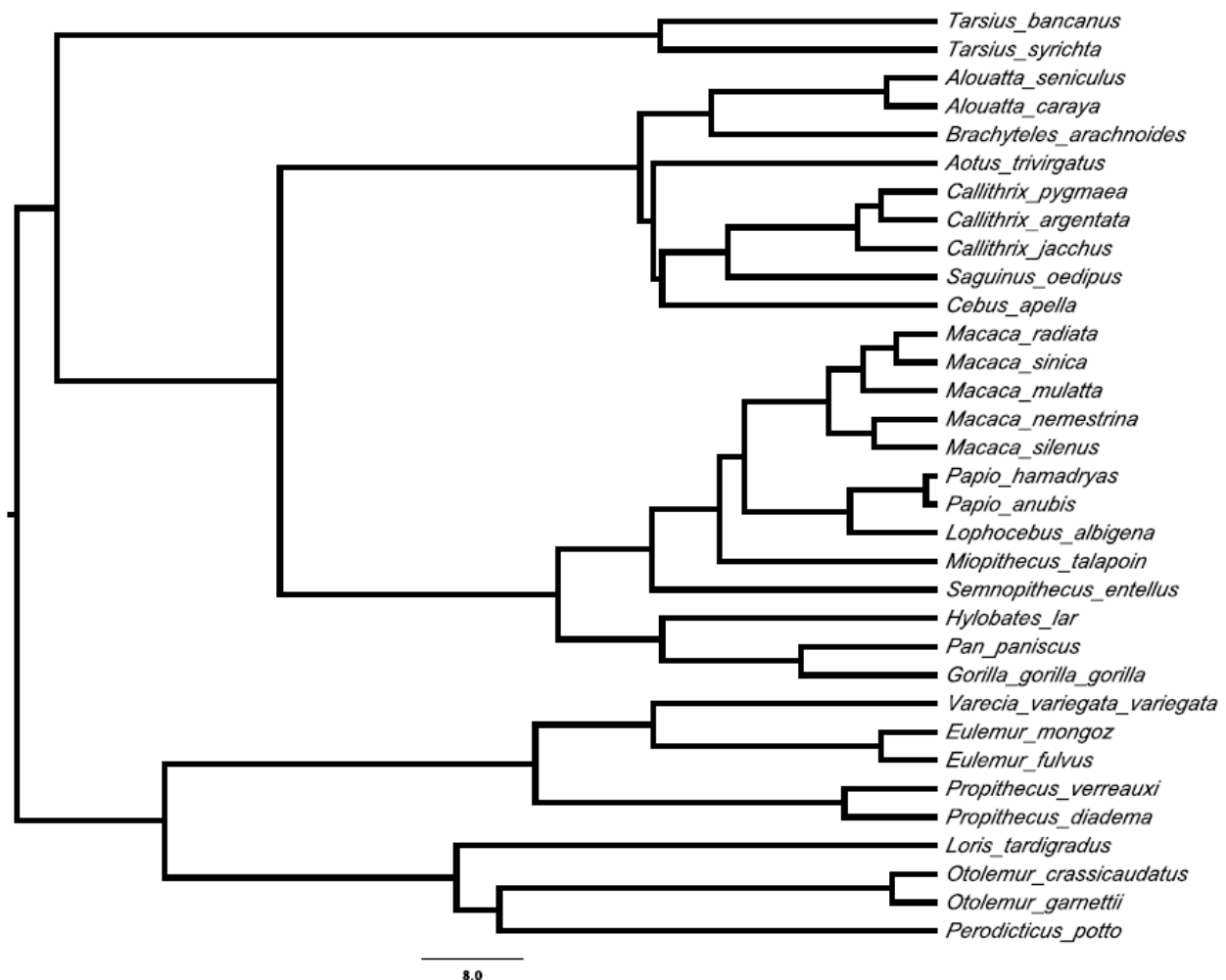
Obviously there are many other fun things to try with regard to formatting the tree, and you can figure out many of them through trial and error. Let's move on to view the different trees in your treeblock. Remember that there are 199 other trees to consider!

To compare the trees visually, use the "Prev/Next" arrows at the top; you can also select a particular tree under "Current Tree" on the left panel. Notice that the tree number is given in the panel. Do you see any major differences in the topologies or branch lengths? You may need to "rotate" the branches around to make the trees more comparable; see the button at the top. Or, under "Trees," select "Order Nodes" to give more consistent representation of trees across the tree block. In this case, the tree topologies are generally identical (I didn't check them all...); it is the mainly the branch lengths that differ.

Note that it is also possible to reroot the tree or even to collapse nodes (useful for representing large trees in Powerpoint or in publications). Give some of these options a try with your file.

Finally, let's print a high quality image of tree 21 (randomly chosen!) for publication by selecting that tree, and then going to the menu "File -> export graphic" and saving the tree as an eps file. (Actually, probably you'd want to do this for the consensus tree, but it works as an example.) Notice that all major options are provided for saving different

file types. Open your beautiful creation in the program of your choice. You should see something like this:



Now, repeat many of these steps with the consensus tree.

R's ape package: Next, let's take a quick look at the trees in R. For this, you will need to download and install R. Go to <http://www.r-project.org/> for more information.

Start up R. The next step is to download a "package" called ape, which provides functionality for importing Nexus files and viewing the trees graphically. To download ape, you can either use "Package Installer" under the "Packages and Data" menu. Or, type this command on the command line, `install.packages("ape")`. To load the package so that it can be used in R, type `library("ape")`.

Now, change the working directory to the folder where your Nexus files are stored. This can be done with the command under the Misc menu (on a Macintosh), or use the function `setwd("<path>")`. Import the tree using the “read.nexus” function in ape.

```
treefile = read.nexus("treeblock_example.nex")
```

Notice how quickly this works, relative to importing the trees to FigTree! (It is especially noticeable with larger files.)

All your trees are now stored in the variable “treefile.” Find information about particular trees by typing, e.g., `treefile[[8]]` for the 8th tree – *note the double brackets!* The “summary” function gives more information about the tree, e.g. `summary(treefile[[8]])`... give it a try to determine the range of branch lengths, and whether the tree contains any polytomies.

Let’s now view tree #8 graphically by typing:

```
plot.phylo(treefile[[8]])
```

Try making a red and blue tree by setting the branches of the tree to red with an additional argument (`edge.color="red"`), and the tip.color to blue (`tip.color="blue"`). You can also adjust the width of the branches (e.g. `edge.width=3`), and make the text bold italics (`font=4`):

```
plot.phylo(treefile[[8]], edge.color="red",  
tip.color="blue", edge.width=3, font=4)
```

Remember that you can recall previous commands (and edit them) by using the up and down arrows – *a very useful trick to use that will save you a lot of time.*

Many more options are available, including adding annotations, scale bars, showing the value of data at the tips, and placing multiple trees side-by-side. The book “Analysis of Phylogenetics and Evolution in R” by Emmanuel Paradis (2006) gives a nice overview of the options.

Lastly, save the tree for incorporation into a presentation or publication (e.g., copy and paste into powerpoint, or Save as... Rplot).

Mesquite: Finally, let’s view the trees in Mesquite and learn about adding trait data.

Open Mesquite, and then open the `treeblock_example.nex` file from within Mesquite. A warning will appear stating that a block of trees has been read that lacks data. Just click “OK”, and Mesquite will go on its merry way.

Click on “View trees,” and a tabbed window will appear that shows the first tree. Notice, however, that Mesquite is by default showing a simple topology without branch length

information. That is because the default view in Mesquite is to not show branch length information. You will normally want to view “Branches proportional to their lengths,” which can be done by selecting that option under the “Drawing” menu.

Scroll through the trees using the arrows at the upper left corner of the tree window. Experiment with other “drawing” options, and when you are satisfied with the way that Tree 8 looks, save it as a pdf under the “File” menu.

You might have noticed that with these trees, all the species line up at the tips. These are “ultrametric” trees because the branches have been adjusted to reflect time since two species last shared a common ancestor (see the *10kTrees* documentation to find out how this was done with fossil calibration points and the program r8s). In many publications, trees often have variable lengths from the root to the tips. While this might be expected if the tree has fossils, in many cases the trees do not have fossils... they are showing number of changes along branches in the data used to infer the trees rather than showing branches proportional to time.

Let’s go back to *10kTrees* to quickly view the trees before they were transformed in r8s. Simply open your web browser and go to the *10kTrees* website again. Under “Download Trees,” select the option “Load previously selected set of species” under “Which species should be included”. Choose your file that saved settings from the last time you were at *10kTrees* – it is typically a weird looking complex name, like “10kTrees_Primates_storedSession2011-01-06T21-12-33-05-00”. Then, select “Load Session”.

Read the directions carefully... you close the window, and going back to the “Download Trees” page, refresh the page. It should go back to your previous settings (well, it might already have done so, depending on your browser and the last values selected at *10kTrees*).

Now, go to “Branch Lengths” and change the setting to Phylogram. Download the trees and save the file as `treeblock_ex_phylogram.nex`.

When you open this tree in Mesquite and view “Branches proportional to lengths,” you will see that the root to tip distances vary (the tree is not ultrametric). There are cases when such trees are preferred to the ultrametric trees that we used earlier in this exercise.

Adding Trait Data in Mesquite. There are many options for adding trait data to your Nexus file. To do this in Mesquite, go to the “Characters” menu and select “New Empty Matrix.” Give this matrix the name “Data”, and use the default setting of 2 characters. Select “Continuous Data,” and click “OK.”

A data matrix window will appear with a list of all the species. Give the data two names (they are actually random values!).

You can type in the data manually, or, if you import the text file into Excel (or some other spreadsheet program), you can copy and paste the data into Mesquite (e.g., try this with the data from Exercise 1). ***Before copying and pasting, just be sure that the species are in the same order in the two files.*** Even a minor re-arrangement can wreck all subsequent analyses. Note that it is possible to sort the data in ascending or descending alphabetical order in Mesquite with the little “beehive” icons to the left of the list of species. Give this a try, and double check that your data in the data file match what is in the text file.

There are actually other ways to link up various datasets within your Mesquite project. The developers of Mesquite provide a number of tutorials showing how to do this. Just be sure that the final data set matches up the species as in original files.

Next: Remember from Exercise 1 that we were missing two taxa from the trees in *10kTrees*. Go to Exercise 3 to learn how to add lineages to a block of trees using an R script.

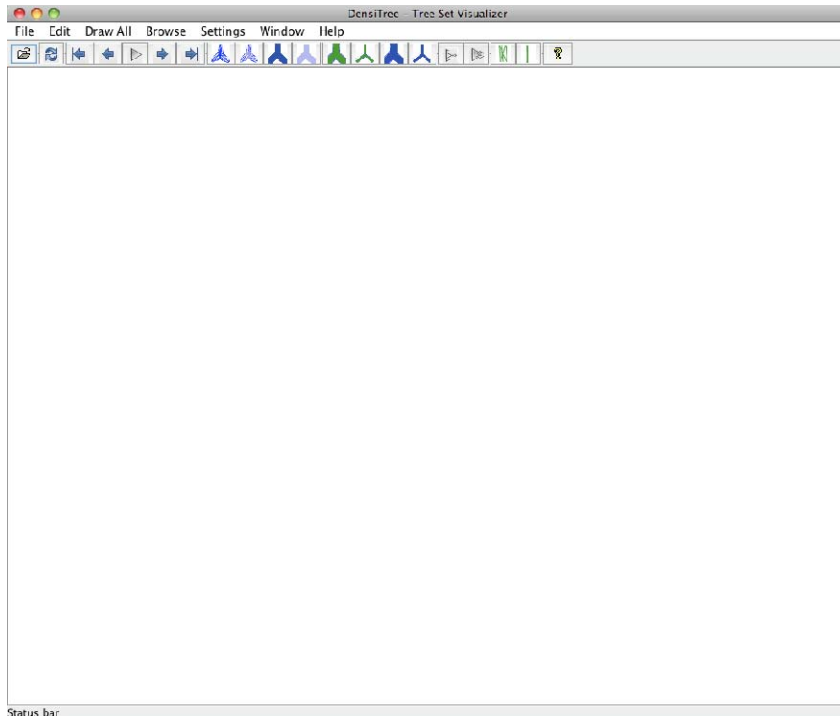
DensiTree

Another cool program to use is DensiTree (Bouckaert 2010). It provides a way to visualize variation in the tree. It is a bit quirky in terms of start-up procedures, but actually easy to use; directions here are for a Mac.

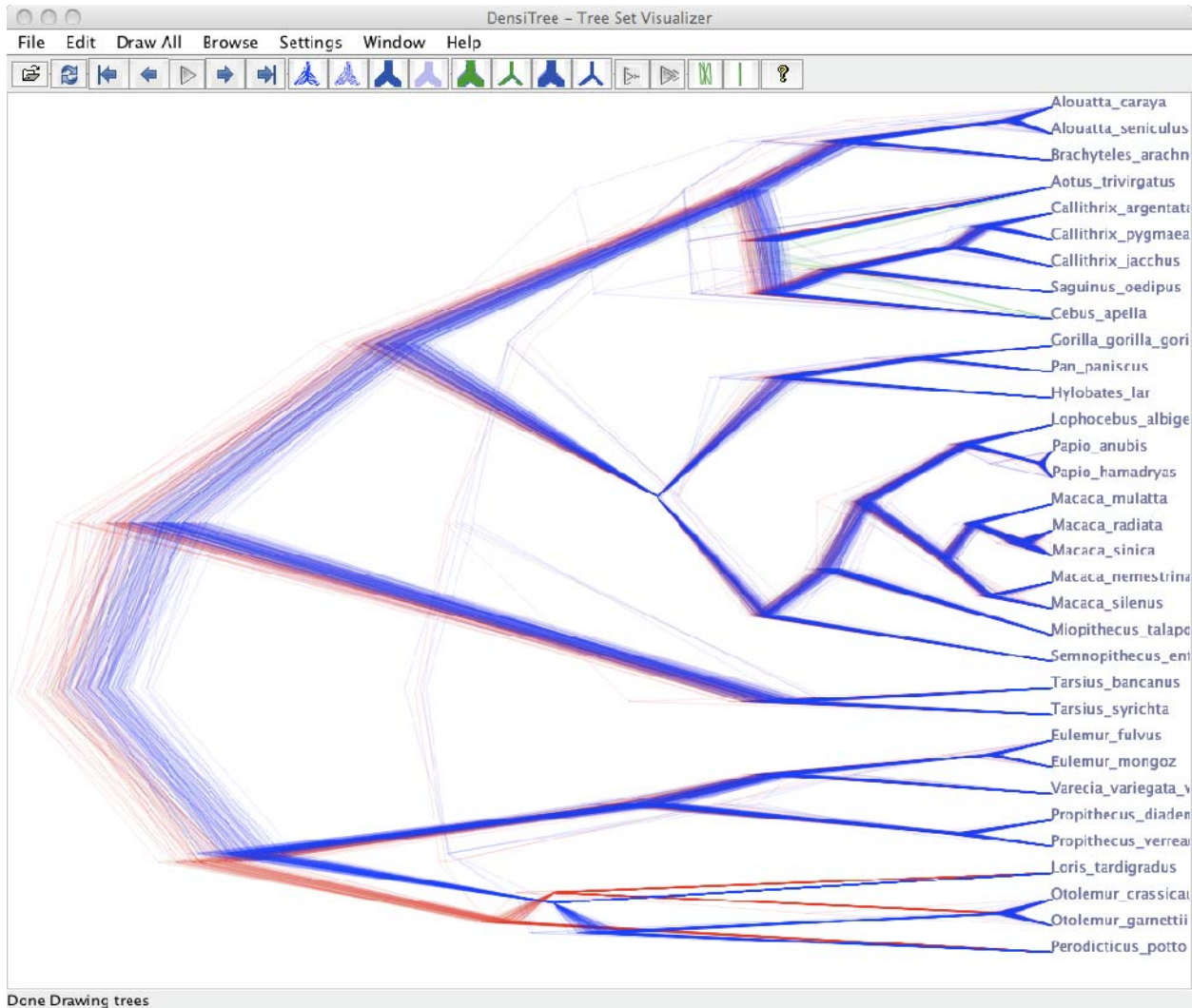
Simply go to the program [website](#) and download a copy of the “jar” file, ensuring first that you have a recent version of Java installed (at least version 1.6. is required).

Then, open a Terminal window (the program should be in the “Utilities” folder of your “Applications” folder). Navigate to the folder where the jar file was downloaded by typing “cd”, space, and then dragging the folder to the command line (a cool trick to enter a path on a Mac command line...). Press return.

Then, simply type “java -jar DensiTree.jar”. The program will start up, and you will see a blank window, like this:



Under the “File” menu, select “Load,” and locate your “treeblock_example.nex” file. Open it, and you should see this:



DensiTree has plotted each of the 200 trees by overlaying them on top of each other. It has also used different colors for incongruent topologies. The graphic shows, for example, topological incongruence for the *Loris-Otolemur-Perodicticus* clade. Congruent topologies are of the same color, but appear more fuzzy and spread out where there is more variation in branch lengths between the trees. Explore the settings and resources available on the DensiTree website.

References

Paradis, E. 2006, Analysis of Phylogenetics and Evolution with R, Springer Verlag.

Bouckaert, R. 2010. DensiTree: making sense of sets of phylogenetic trees. *Bioinformatics* 26:1372.