

Table of Contents

Linux Basic	1
Mac Instillation Instruction	2
Windows Instillation instruction	2
Basic Linux Command	2
Download the GitHub Repository	3
Getting Setting UP	4
Mac User	4
Windows User	4
Alignment with MAFFT	5
Mac Instillation	5
Windows Instillation	5
Running a Basic Alignment	5
TNT - Parsimony	5
Mac Instillation	6
Windows Instillation	6
Introduction to TNT	6
TNT Analysis	6
Setting Up TNT Analysis	6
Basic TNT Analysis	7
Exporting and Quitting	8
Viewing your Phylogeny	8
Homework Questions!	8
Homework Question 1:	9
Homework Question 2:	9
Homework Question 3.	9
References	9

% Linux, Alignment and Parsimony % Jesus Martinez-Gomez % s
c(current_date) % This worksheet will walk you through some basic
in Linux, aligning phylogenetic sequences with the program MAFT then
inferring a phylogeny using the program "Tree analysis New Technology" (TNT).
It also contains Homework question at the end.

Linux Basic

Linux is a operating system that phylogenetic programs are created with. The Mac operating system is Linux based meaning that it runs Linux. The Windows operating system is now Linux based, therefore we need to install Linux. These

Mac Instillation Instruction

1. Linux is already installed we just need to open it.
2. Look for the Terminal.app using your finder. It should be in the following location Applications/Utilities/Terminal.app
3. Be moral support for your classmates with Windows because its complicated.

Windows Instillation instruction

For Window user we are going to install a Linux virtual machine. In essence this is a computer-in-computer. There are many benefits to installing Linux as many program are Unix basic. These instruction are for Windows 10 and they were taken from this site

1. Search for Windows PowerShell in Windows search bar, right click "Run as administrator"
2. Copy and paste the following into the Window:

```
Enable-WindowsOptionalFeature -Online -FeatureName Microsoft-Windows-Subsystem-Linux
```

3. You will be prompted to restart your computer by pressing "Y", do so.
4. Open Microsoft Store using the Windows Search Bar.
5. Search for the program "Ubuntu" and download it.
6. Open "Ubuntu", you will be required to set a Username and Password. This password is different from your Windows login, **Make sure you don't forget it!**
7. You are done!

Basic Linux Command

Now that we have Terminal open we will run through a few tips n' trick. Terminal essentially allows you to navigate your computer through a command line rather than a Folder system. In this section we will walk you through creating a new folder and navigating around which is the most basic function of Linux.

First of all it is important to figure out where you are. Type the following, which stands for 'Print Working Directory'

```
pwd
```

Ever

```
ls
```

This should list all the files and folder in your current location. Mac User will notice that you have a lot of files, this because Linux is built into your computer and has access to all your files. Windows user will have nothing. This is because, as we said before, you are running a virtual machine (computer-in-a-computer), so you don't have any files. Lets make a new directory. The function "mkdir" is short for make directory.

```
mkdir testFolder  
ls
```

You should see that you just made a new directory called "testFolder". Now let move into that new folder. We will use the function 'cd' which stands for "change directory"

```
cd testFolder
```

We are now in testFolder! If we type 'ls', we will notice that there is nothing. That is because we haven't made anything. Lets go back to our previous directory using 'cd'

```
cd ..
```

Two periods after cd moves you up one step in directories. Now lets delete the testFolder.

```
rmdir testFolder
```

These instructions are super basic but will get you started. This is a link to a Linux cheat sheet. If you want to learn more Linux google it.

Download the GitHub Repository

All the material for this course are found on a Github repository. This is essentially a cloud storage system that allows for version control (very important but we don't cover). You can view this repository using a browser by following the link.

You can download the data using your browser but I will have you download it using terminal Before you do this but first you need decide a place put the folder. For Mac user I recommend putting it on your Desktop, for Windows user I recommend putting it in your /home/ directory.

1. Using the skills you learn in the previous section navigate to the location you want to download the folder.
2. Use the following bit of code to download the data.

```
git clone https://github.com/Jesusthebotanist/PLBI06401_EcologyEvolution_Module_2019.git
```

Using a finder check to see if you have successfully download the folder. A folder called "PLBI06401_EcologyEvolution_Module_2019", should now exist on your computer.

Getting Setting UP

Installing and getting programs to work can be a challenge. Bioinformaticians spend a great deal of time simply getting a program to run. These are general instructions and with the help of the instructors we will get the program running on your machine. But an equally important aspect is keeping yourself organized. There is nothing worst than having 8 different files all named some variation of, "Final1", "Final2_forReal", "Final3_for_real". To avoid this we will create a new folder where you will do all your analysis.

In the following bit of code we will make a new folder called "My_first_Parsimony_Analysis" inside of the folder "Day1". We will then copy and paste a two programs and a data file into this folder. **When you do your Homework name folder something different.**

Mac User

```
cd PLBI06401_EcologyEvolution_Module_2019
mkdir Day1/My_first_Parsimony_Analysis
cp Day1/programs/mac/mafft-7.450-signed.pkg Day1/My_first_Parsimony_Analysis
cp Day1/programs/mac/tnt-mac.command Day1/My_first_Parsimony_Analysis
cp Ruhfel_unaligned_fasta/ruhlf_32_by_5000_unaligned.fas Day1/My_first_Parsimony_Analysis
cd Day1/My_first_Parsimony_Analysis
```

Windows User

```
cd PLBI06401_EcologyEvolution_Module_2019
mkdir Day1/My_first_Parsimony_Analysis
cp Day1/programs/windows/mafft_7.450-1_amd64.deb Day1/My_first_Parsimony_Analysis
cp Day1/programs/windows/tnt-linux
cp Ruhfel_unaligned_fasta/ruhlf_32_by_5000_unaligned.fas Day1/My_first_Parsimony_Analysis
cd Day1/My_first_Parsimony_Analysis
```

Next we are going to make copies of the programs and data that you will be using and placing them into this folder.

Alignment with MAFFT

We will be using the program MAFFT to run a basic alignment.

Mac Installation

Using finder simply double click on the file called, "mafft-7.450-signed.pkg". Then delete the file

```
rm mafft_7.450-1_amd64.deb
```

Windows Installation

```
sudo dpkg -i mafft_7.450-1_amd64.deb  
rm mafft_7.450-1_amd64.deb
```

Running a Basic Alignment

```
mafft ruh_f_32_by_5000_unaligned.fas > ruh_f_32_by_5000_aligned.fas
```

TNT - Parsimony

We will be inferring phylogenies using the program "Tree analysis New Technology" (TNT). This is a command line program, which means in order to run it you will use the "Terminal" app in Mac/ Linux or the "Command Prompt" app in Windows. It is developed in part by Kevin Nixon a Plant Biology faculty! The following is a worksheet that you can follow to perform a basic parsimony analysis in TNT.

Mac Instillation

```
sudo chmod +x tnt-linux  
./tnt-linux
```

Windows Instillation

```
sudo chmod +x tnt-mac  
./tnt-mac
```

Introduction to TNT

As mentioned, TNT is a command line program meaning all inputs will be through the command line (you'll feel like a hacker). An alternative is to write script. A script is a simple text file that has all the instructions (aka commands) you want a program to run. We will not go through how to write a script, instead we will enter the commands one by one into the command line. This section will introduce you to some TNT basic and tips.

First a brief note on syntax. Most TNT commands are a simple word or abbreviation, but in order to get the program to run them they must end with a semi-colon ";". So for example, try typing in the following command "help". You can simply copy and paste it into TNT.

```
help;
```

The "help" command above will give you all list of all the commands that TNT can run. Nifty. The "help" command can also be used to find information about other commands. For example we will be using the "mult" command later on in this workshop. To find out more information about "mult". Type the following.

```
help mult;
```

If you are ever unsure about what a particular command does always use "help"

TNT Analysis

Setting Up TNT Analysis

We will now perform a basic parsimony inference. The first step is to specify settings in TNT. The "log" commands will generate a text file, (myLog.txt) that will keep track of all the subsequent commands you type and the output TNT produces. It is a good coding practice to keep log files, as they are a written history of the analysis you've done (future you will thank you).

```
log myLog.txt;
```

Look at the folder, a new file called myLog.txt should be in there. You should be able to open this up by double clicking on it. It should be empty right now, but will fill up after we are done. The second step is to set the RAM memory usage. When TNT runs it takes up memory on your computer, the default setting for TNT memory usage is low. We'll set it to 200mb which is good for our tree size and no problem for a modern day computer. If you are analyzing a large tree, consider setting the memory to something larger.

```
mxram 200;
```

The next step is to read in our data. Two important things to note:

1. Make sure your input file is in the correct format. TNT takes a special file a .tnt. Consult the TNT website for more detail on how to specify this file.
2. Make sure your input file is in the same folder as the tnt.command (Mac) or tnt.exe (Windows) file. It should be if you followed "Download TNT" instructions above.

```
p ruhf_32_by_5000_aligned.tnt;
```

TNT searches through tree space to find the most parsimonious tree or set of most parsimonious trees. TNT searches through many many trees most of which are not good (aka have a good parsimony score). Therefore we don't want to save every single tree, instead we will use the "hold" command to save the top 1000 trees.

```
hold 1000;
```

Basic TNT Analysis

We are ready to perform a basic parsimony tree search. We don't have time to explain what these do, but Kevin Nixon recommend these steps for good search. You may notice that instead of a single command there are four in the following line of code. T

```
Mult; Rat; Drift; Sect;
```

This analysis may take a little while, and you will see some fun stuff pop up on your screen. Once its done you can view your most parsimonious phylogeny in your terminal window using the following code.

```
tp;
```

```
resample replications 200
```

Exporting and Quitting

Now that we have successfully completed an analysis we may want to export our best tree into a file for safe keeping. First we will use the command "taxname=", TNT changes our tip name during this analysis, "taxnames=" changes them back. Next we use the "export" command to generate a .nex file, which stands for Nexus. A Nexus file is a standard phylogenetic file.

```
taxname=;  
export - ruhf_32_parsimony.nex;
```

Double check to see if a file called myPhylogeny.nex was created. Finally we can exit out of TNT using the "quit" command

```
quit
```

Open the myLog.txt file again (if you already have it open, close it first), it should now be full of all the commands you inputted and the TNT output! Yay phylogenies!

A final note. If you rerun the above code it will overwrite any files that exist on your computer. If you are running a different analysis make sure you change the names of the .log file and the .nex file

Viewing your Phylogeny

You should now have a .nex file in your folder. This .nex file contains the phylogeny you just inferred and can be viewed in a tree viewer. There is a free Browser based tree viewer called (IcyTree)[<https://icytree.org/>]. Just drag and drop your .nex file into the browser

Homework Questions!

Homework Question 1:

Homework Question 2:

Homework Question 3.

References
