Student Guide for Mesquite

This guide describes how to: 1. create a Mesquite project file, 2. construct phylogenetic trees, and 3. map trait evolution on branches (e.g., morphological traits).

Morphological Data Module (discreet or categorical data)
This section describes how to use Mesquite to analyze morphological sequence data.

Creating the Project and Data File
1. From top toolbar select **File, New**. A pop-up window will appear.
2. Name the project (e.g. Morphtree.nex). You must add the **.nex** extension and choose where to save the file in the line below the file name. Click **Save**.
3. The **New File Options** window will open.
   a. Click on the **Name** box and name of the matrix (e.g. Morphology.)
   b. Click on **Number of Taxa** and enter the number of taxa , (e.g. 5) in your study. Taxa can be added or removed later.
   c. Check the **Make Character Matrix** box. The box for **Make Taxa Block** should also be checked. Click OK.
4. The New Character Matrix window will open.
   a. Enter a name for the character matrix in the text box.
   b. Enter in the number of characters (e.g. 10)
   c. Select **Standard Categorical Data** (default). Click OK.
   d. The project window will now open, with the project file name (e.g. Morphtree.nex) appearing at the top of the window.
5. The project window will show the name of your data matrix (e.g. Morphology) and the main project window. The Project and Character Matrix tabs are always displayed by default when creating a new project (Figure 1). Subfiles for the project are shown on the left. The Character Matrix toolbar (Figure 2) is next to the data matrix.
6. You can exit a tab by clicking the **X** in the top right of the tab. It will close the tab, but it is not deleted. To restore a tab, click on **Show Matrix** in the subfiles.
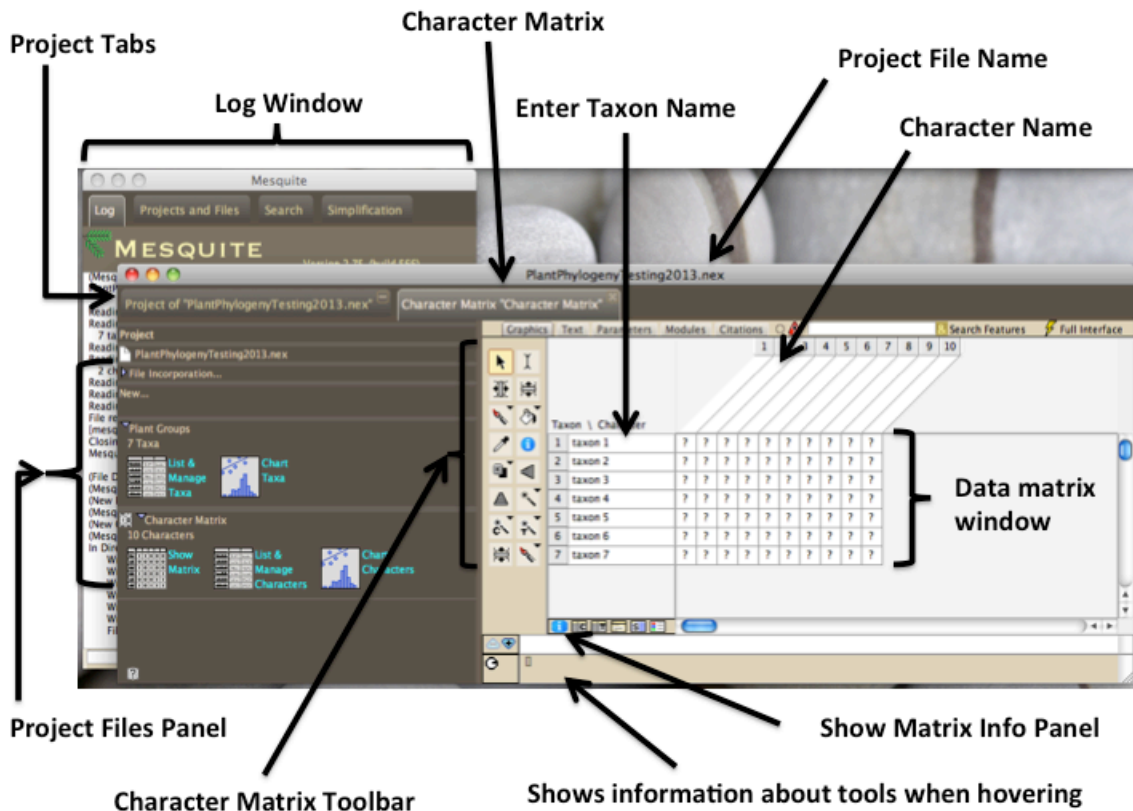
Figure 1. View of sample project window showing project tab, Character Matrix tab, Character Matrix toolbar, and Data Matrix.

Entering Taxon and Character names in "Character Matrix" Tab
1. Double-click the box named **taxon 1** in the first row of Character Matrix. Enter the taxon name. Press ENTER, and the cursor will move to the next cell. Repeat for all taxa.
2. Double-click or click the first column in the Character Matrix and name the character. Repeat to name the other characters.
3. To add additional characters or taxa , select the **Add Characters** or **Add Taxa** tools to the Character Matrix toolbar (Figure 2a). To delete characters or taxa, choose the **Select Tool**, click on the column or row, and press Delete.
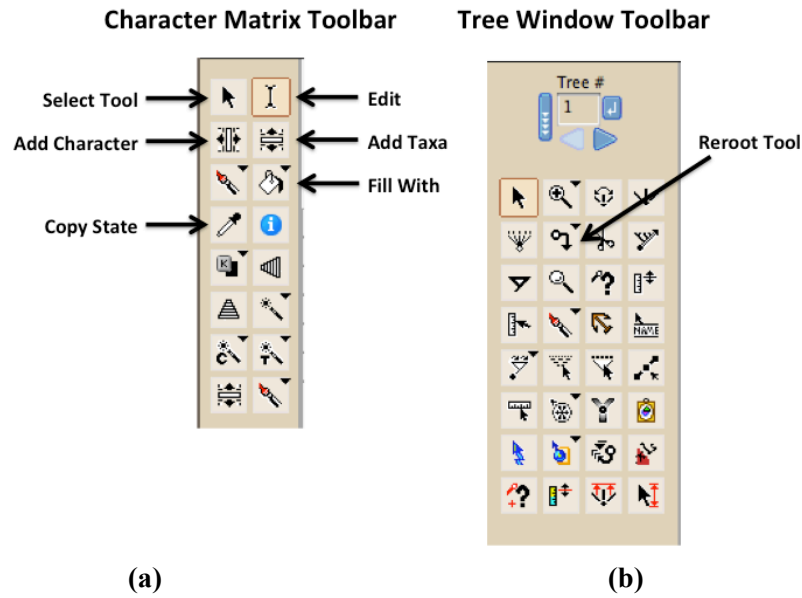
(a) **(b)**

Figure 2. Important Character Matrix Toolbar (a) and Tree Window Toolbar (b) buttons.

Entering Data
1. Cells in the data matrix initially contain a "**?**" indicating a missing character.
2. Manually enter values in the matrix by choosing the **Edit Tool** (Figure 2a) or double-clicking on the cell
3. Enter 0 for ancestral character states or 1 for derived.
4. If a state is not entered as a numeric value, a "?" will remain in the cell and Mesquite will read it as a missing or not applicable character. This is also the best option if a taxon lacks a character, or if the character state is unknown.
5. To examine the matrix with character state names displayed, choose **Show Matrix Info Panel** button in the lower left corner of the Character Matrix Window (Figure 1).
    a. A window will open on the right with the character name displayed above the numbers 0, 1. Click in the box next to a number and enter a character state. Again, zero should indicate the ancestral condition. Press Enter to move to the next character state.
    b. *Note:* Entering text is strictly a visual aide for the user. The program still reads the matrix as numeric values.
6. Proof the matrix to ensure data have been entered correctly.
7. Save the file by clicking **File** on the Mesquite menu bar, and choose **Save File**.

Constructing the Tree
With the data matrix completed, the next step is to use Mesquite's tree searching algorithm to produce the most parsimonious trees.
1. On the Mesquite menu bar, click **Taxa&Trees**. Highlight **Tree Inference**, then highlight **Tree Search**, and click **Mesquite Heuristic Search (Add & Rearrange)**.
2. A new window will open. Select Treelength and Click OK.
3. If the program asks, **Source of characters (for Treelength)**, select **Stored Matrices** and click OK.
4. The Select window will open. Choose **SPR Rearranger** and Click OK.
5. Set MAXTREES to 100 (the default setting). Click OK, and in the **Separate Thread?** window choose **NO**.

6. The commands will execute and a new window will open with a representation of your tree from the data you entered. This first tree is unrooted, meaning the Outgroup must be designated. (An Outgroup is the ancestral taxon that you used to assign character states.)
7. Select the **Reroot Tool** from the **Tree Window Toolbar** (Figure 2b). Choose the **Select tool** and move the cursor over the branch of the outgroup taxon. Click the branch and the tree will root automatically with the outgroup you have chosen.
8. To save this tree, select **Tree** in the Mesquite menu bar, and a drop-down list will appear. Choose **Store Tree**. If this is not done, Mesquite will erase your changes.
9. A new section will appear in the Project File Panel called "**Trees from Mesquite's heuristic search**". These are all of the trees produced by the analysis. You can view these trees by clicking the blue arrow in the Tree Window Toolbar.
10. *Note:* To modify the Tree Shape, choose **Drawing** on the Mesquite menu bar , highlight Tree Form, and select desired tree shape. This will demonstrate that different tree forms show the exact same relationships.

Constructing a Consensus Tree
A consensus tree combines the equally possible tree arrangements into a single representative tree showing the most likely relationship among the taxa.
1. Choose **Taxa&Trees** on the Mesquite menu bar, highlight **Make New Trees Block from**, and Select **Consensus Tree**. A new window will open.
2. Select **Stored Trees** and click OK.
3. In the Consensus Calculator window, select **Majority Rule Consensus**, and click OK.
4. The Majority Rule Consensus Options window will open. The default settings of **Consider Tree Weights** and **Write Group Frequencies** should be checked. Root as specified in the first tree should be selected in this window. Click OK.
5. Click OK to accept the defaults for Majority Rule consensus options.
6. Click NO in the **Separate Thread?** window.
7. If a new window opens and asks which **Tree Block** do you want to use, choose Trees **from Mesquite's heuristic search**.
8. The Trees Ready will appear. Click Yes to view them.
9. The tree will be unrooted. Root as described above.

Tracing Characters on Tree branches (Parsimony and Ancestral States)
Having constructed a Majority Rule Consensus Tree from your dataset, you can evaluate how informative the chosen characters are and their pattern of evolution.

1. Open a tree window. Select **Analysis** from the Mesquite menu. Choose **Trace Character History**.
2. The Select window will open. Choose **Stored Characters** and click OK.
3. Select **Parsimony of Ancestral States** and click OK.
4. A Trace Character box. Click and drag to move this box. Clicking the left and right navigation buttons on the Trace Character Box will scroll through all characters in the data matrix.
5. To remove Trace Character from the Tree Window, choose **Trace** from the Mesquite menu bar and click **Close Trace**.

Saving Tress and Saving Work
1. To save work, select File on the Mesquite menu bar, and choose **Save File**.
2. To save a PDF of the phylogenetic trees you have built, select which tree or trees you want to save, and from the **File** drop-down, select **Save Tree as PDF**. You must add the .pdf extension to the file name.

Molecular DNA Sequence Data Module
This section describes how to use Mesquite to analyze molecular sequence data.

Loading Sequence Data files
1. Start Mesquite and open the **File** dropdown menu.
2. Select Open File and choose you molecular data file (*e.g.* PlantRBCL2012.fasta). Click **OPEN**.
3. In the Translate File Window, select **FASTA (DNA/RNA)** and click OK.
4. Save the Imported File as a NEXUS file (.nex). You must add the .nex extension or the file will not save correctly. Click **SAVE**.
5. A new project tab will open.
6. Click **Show Matrix** in the project subfiles, and a matrix containing colored, alphabetic characters representing the four DNA bases (A, T, C, G) will open.

Sequence Alignment
The DNA sequences must next be aligned, otherwise, the analysis will be meaningless. Alignment searched to match base locations and identify where mutations such as deletions or insertions have occurred.

Sequence Alignment
1. NOTE: If you are conducting the analysis on a personal computer, you will need to download the MUSCLE alignment program from http://www.drive5.com/muscle/). Pay attention to where you install it as that information is required to run the alignment.
2. Select **Matrix** from Mesquite menu bar.
3. Highlight **Align Multiple Sequences** and select **Muscle Align**.
4. A new window will open. Choose **NO** to running on a separate thread.
5. A new window will open. In the box, indicate the path to the MUSCLE aligner program (i.e. its location on the computer). After defining the path for MUSCLE, choose OK.
6. The terminal window will open while the program executing. Once finished, exit the terminal window and return to Mesquite.
7. The sequence will now be aligned and show where bases in the sequence are identical (same color), different (different color), or absent (empty box)

Constructing Trees Using DNA Sequence Data
1. Select **Taxa&Trees** from the Mesquite menu bar.
2. On the Mesquite menu bar, click **Taxa&Trees**. Highlight **Tree Inference**, then highlight **Tree Search**, and click **Mesquite Heuristic Search (Add & Rearrange)**.
3. The **Criteria For Tree Search** window will open. Select **Treelength** and click OK.
4. If the program asks, **Source of characters (for Treelength)**, select **Stored Matrices** and click OK.
5. The Select window will open. Choose **SPR Rearranger** and Click OK.
6. Set MAXTREES to 10 (do not use the default setting or it can take hours to complete). Click OK, and in the **Separate Thread?** window choose NO.
7. The command will execute and a two New Windows will appear with the title "Tree Search" and "Command is executing". Do not close the windows! This step may take as long as 20 minutes depending on your computer speed.
8. Clicking the down arrow within the Tree Search progress bar will report how the analysis is running.
9. To quit the job, click **Emergency Cancel** and the job will terminate.
10. Once the search completes, an unrooted tree will be produced. Root the tree as described above with the outgroup taxon.

Construct a Consensus Tree and Save Files as described above.

Optional Activity for Combined Analysis of Morphological and Molecular Data (*Advanced)*
With the completion of both morphological and molecular analyses, students can investigate how sequence data give insights on morphological trait evolution by combining data. This can show how morphological traits can be mapped on top of the phylogeny produced from molecular data.

Adding Morphological Traits to the Molecular Phylogeny
1. Either continue from the molecular study or open the molecular study that was saved previously.
2. From the Mesquite menu bar, choose **Characters**.
3. Click **New Empty Matrix**.
4. Name the new data matrix as described above .
5. Choose characters from the morphological analysis (e.g. vascular tissue, seeds, flowers)
6. Enter the number of morphological characters to be added to the molecular data matrix into the box and choose **Standard Categorical Data**. Click OK.
7. A character matrix window will open. The names of the taxa from the molecular data will be present. Enter the characters you chose and score them as ancestral or derived in the columns.
8. Open the consensus tree tab and click **View Trees**.
9. Click **Analysis** from the Mesquite menu bar and choose **Trace Character History** as described above.
10. In the **Source of characters to reconstruct** window, select **Stored Characters**, and click OK.
11. In the next window, select **Parsimony Ancestral States**, and click OK
12. In the Select Data Matrix window, select the new data matrix you produced (in step 4 above) that added morphological traits to the molecular data set. Click OK.
13. Mesquite will now show the morphological characters you added on the previously constructed molecular tree. The trace character box can be moved as described above. As you click on the blue arrows in the **Trace Character** box, you can examine whether a trait has evolved once or multiple times. This shows patterns of morphological trait evolution relative to the molecular phylogeny.
14. Save the tree or project as described above