A Mesquite package for support value integration

Jeffrey C. Oliver, Jolanta Miadlikowska, A. Elizabeth Arnold, David R. Maddison, & François Lutzoni.

Introduction

These Mesquite modules allow users to draw grids on branches of trees to reflect support values from multiple analyses, as in Miadlikowska et al. (*Mycologia*, 2006). This package allows users to show multiple node support values on a tree (referred to as the **Display Tree**), either as the actual support value, or as a color based on a user-defined threshold value. These support values are harvested from **Support Value Trees** (**SVTs**), which, in the current implementation, can be non-parametric bootstrap values or Bayesian posterior probabilities.

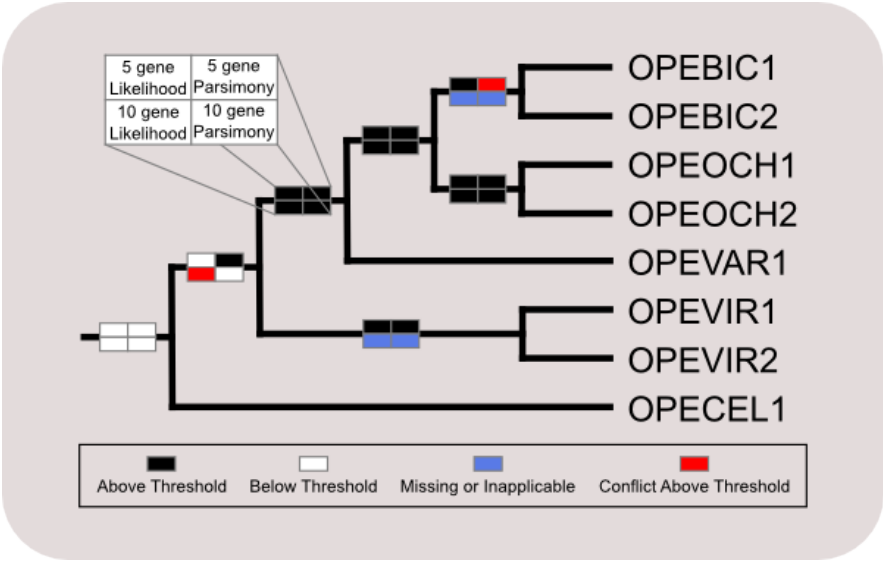

Figure 1. Example display of support values from four separate trees.

Support values can be reported from a SVT onto a display tree as long as all taxa in the SVT are also found in the display tree: it will be possible to show support values from a SVT with less taxa onto a more comprehensive display tree as long as the SVT is a subset of the taxa part of the display tree. When topological bi-partitions of the display tree are not found on the SVT, this will be reported on the display tree as inapplicable. If a bipartition in a SVT is not present on the display tree, because it is part of a polytomy of the display tree, all support values will not be reported. As long as all lineages part of a polytomy are represented by at least one descendent on one side of a topological bipartition of a SVT, this value will be reported on the display tree.

I. Installation

Hypha requires Mesquite, so start by installing Mesquite from:

http://mesquiteproject.org/mesquite/download/download.html

or, if Mesquite is already installed, check to make sure you are using a compatible version (version 2.75 or later). To install hypha, download the package from:

http://mesquiteproject.org/packages/hypha/manual/download.html

Download the hypha.zip file and extract the single directory 'hypha'. Place this directory (and all of its contents) in your Mesquite_Folder/mesquite/ directory (Figure 2). In a standard installation of Mesquite, the hypha directory will be among a number of other directories, including align, genesis, io, and trunk. Be sure not to place the hypha inside one of these other directories and to be sure the archive extraction process does not create an additional nesting in the hypha directory. If you look inside the hypha directory, there should be seven directories, including aHyphaIntro and BranchLabelFromOtherTree; there should not be a single directory called hypha. If there there is just a single hypha directory nested in hypha, move the first up one level (to Mesquite_Folder/mesquite/) effectively deleting the (now empty) second hypha directory.

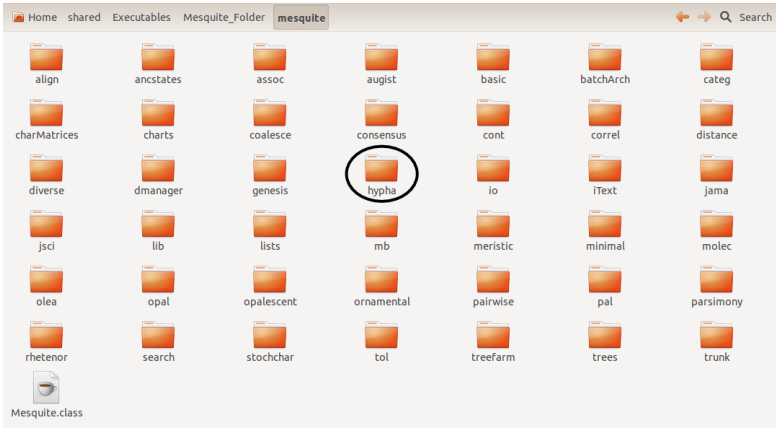

Figure 2. Proper installation of hypha within Mesquite_Folder/mesquite/

II. Using Hypha to display support values

A. Setting up the Mesquite file

To use the node grid coordinator, you will need a Mesquite file with a Taxa block with all taxa from all trees you are considering. For information in setting up a Mesquite file, refer to the Mesquite Manual's sections on files. You can find basic file information in Mesquite (Help > Mesquite Manual, Basics: Files) or in the online manual:

http://mesquiteproject.org/Mesquite_Folder/docs/mesquite/managingFiles.html

For information on taxa, see (Help > Mesquite Manual, Objects: Taxa) in Mesquite or go to the online manual at:

http://mesquiteproject.org/Mesquite_Folder/docs/mesquite/Taxa.html).

The names of the taxa in the Taxa block must match the taxon names in the trees files *exactly* for tree importation to occur smoothly. The file should also include the phylogenetic tree on which you would like to draw the support value grids, hereafter referred to as the display tree. The easiest way to accomplish this would be to open a nexus file from within mesquite, consisting of (a) a TAXA or DATA block containing all the taxa and (b) a TREES block containing the display tree.

Next, include the trees from which support values will be harvested (support value trees). To do this, select Taxa & Trees > Get File with Trees > Include Contents… and select the tree file(s) you would like to include. If you have multiple taxa blocks, you may be asked to instruct Mesquite which taxa block the trees correspond to. Additionally, if there are taxa in these tree files that are not included in the Taxa block of the Mesquite file, users may encounter problems incorporating those support value trees into the file. Differences in spellings of taxon names may lead to problems, but Mesquite reports the taxon names that are causing problems. If there is a problem including a tree file, users should correct the tree file in question, so taxon names match those in the Mesquite file.

When all the support value trees have been included, and all these trees correspond to the same taxa block as the phylogenetic tree, you are ready to set up the Grid Coordinator. If you only have a single taxa block in your file, then all included trees should refer to that taxa block. However, if you're not sure, you can check by opening Taxa & Trees > List of Tree Blocks and looking at the 'Taxa' column. The included support value trees should all refer to the same Taxa block as the display tree you wish to use for support value visualization.

B. Setting up the Grid Coordinator

Start by opening up the display tree in a tree window (Taxa & Trees > New Tree Window > Stored Trees). Currently, the Grid Coordinator is only set up to work on Square, right oriented trees, so make sure the display tree is drawn in this manner (Drawing > Tree Form > Square Tree and Drawing > Orientation > Right). Begin the Grid Coordinator by selecting Analysis > Grids for nodes. You will be prompted to enter the number of rows and number of columns. This will dictate the number of support values shown for each node. Currently, all cells in the grid must correspond to some support value, and the maximum number of columns and rows is three (for a total maximum of nine possible support values displayed in a three by three grid). The number of cells in the grid will likely correspond to the number of support value trees you have included in the file.

Mesquite will then prompt you for the following information for each cell in the grid:

1 – Number For Node of Tree

2 – Tree source

3 – Threshold value for cell

The prompts will occur cell by cell. In other words, you will be prompted for 1, 2, and 3 above for the first cell, then prompted for 1, 2, and 3 for the second cell, then for the third cell, etc.

1. For the first prompt (Number For Node of Tree), you will select either "Branch *Lengths* for Node from Other Tree" or "Branch *Labels* for Node from Other Tree". Select the former if the support value tree corresponding to that cell has support values stored as branch lengths (such as bootstrap scores from PAUP), and select the latter if the support value tree corresponding to that cell has support values stored as branch labels (such as posterior probabilities from MrBayes). If you're not sure how the support values are stored, you can open a tree window for the support value tree in question (before you set up the Grid Coordinator; you may cancel the grid coordinator setup and start over if necessary), via Taxa & Trees > New Tree Window. Click on the 'Text' tab and look at the newick format of the tree. Trees with support values stored as branch *lengths* will show support values immediately after a colon (:). For example:

(((A, B):80, C):95)

Is a newick format tree, showing support value of 80 for an A + B clade, and a support value of 95 for an A + B + C clade. Trees with support values stored as branch labels will show support values immediately after a clade (if the tree also has branch lengths stored, the support values will precede the colon). The same tree as above, with support values stored as branch labels will be represented as:

(((A, B)80, C)95)

If this tree also had branch lengths (showing, for example, the expected amount of evolutionary change), it would be written as:

(((A, B)80:0.50, C)95:1.2)

Indicating an A + B clade with support value of 80 and a branch length of 0.50, and an A + B + C clade with support value of 95 and a branch length of 1.2.

2. The second prompt, the Tree Source (for "Branch *Lengths* for Node from Other Tree" or "Branch *Labels* for Node from Other Tree", depending on choice for previous prompt), refers to the support value trees corresponding to the grid cell. Choose 'Stored Trees'; you will assign specific trees to specific cells in a later step.

3. The final prompt asks the user to enter a threshold value. This is the value used to determine the cutoff point at which different colors are displayed on the grid. For example, if the threshold value entered was 70, any branches with applicable support values equal to or greater than 70 would be filled with black, and any branches with applicable support values below 70 would be filled with white (inapplicable values are colored grey; these are the default colors - users can change the colors later. See Step III, below). The thresholds for determining conflict are the same as the threshold for support.

After this information for each grid cell is provided, Mesquite will draw grids on the display tree. However, Mesquite, by default, is using the display tree for support values, so you will have to change the tree source for each cell. To do so, in the Grid Coordinator Legend, click on a cell and select the support value tree you would like to correspond to that particular cell (if the legend is not visible, it may be at the lower left corner of the tree drawing panel, but not visible if the tree drawing is larger than the window; you can find the legend by using the tree window scroll bars). If the desired support value tree is stored in the same Tree Block as the display tree, you should select "Choose tree…" and choose the appropriate support value tree from the list. If the support value tree is stored in a different Tree Block (it will be if you followed the directions in the earlier part of these instructions, and haven't edited the Tree Blocks), select "Tree block (for Branch Lengths/Labels for Node from Other Tree)". Mesquite will use the first tree in the tree block by default, so if the desired support value tree is not the first tree, you will then need to use the "Choose tree…" option, as above. After you have assigned support value trees for all the cells in the grid, the cells should all be colored according to the threshold values, which are displayed in the Grid Coordinator Legend.

III. Options

After the grids are drawn on the tree, you may change the display options from the Grids menu (Figure 3).

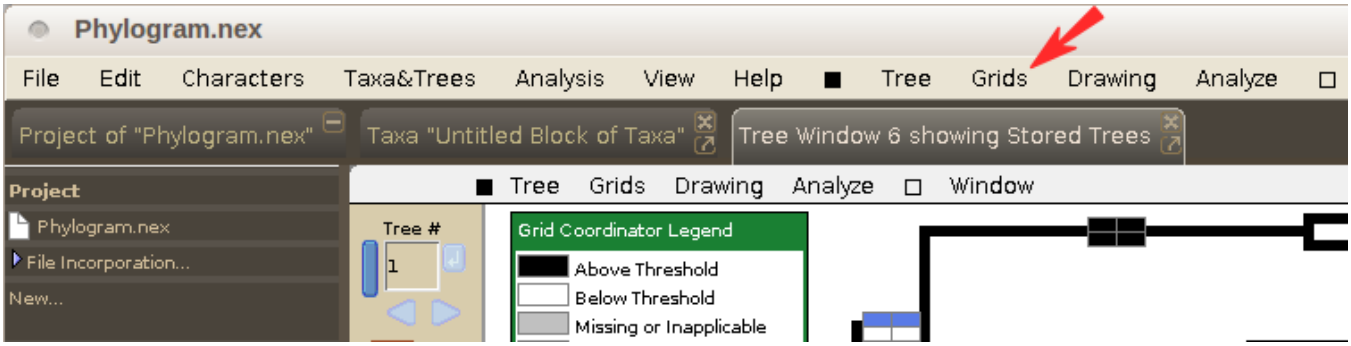

Figure 3. Display options are available in the Grids menu, indicated by the red arrow.

Options available from the Grids menu:

Option	Default Setting	Explanation
Set Cell Drawing Size...	Height: 10 Width: 20	Sets the height and width of the individual cells. All cells will be the same height & width, and the cells can have a height between 4-20 pixels, and a width between 8-40 pixels.
Above Threshold Cell Color	Black	These options allow you to dictate which colors to use. The color choices are applied to all cells in the grid.
Below Threshold Cell Color	White	
Missing or Inapplicable Cell Color	Grey	
Low Conflict Cell Color	Light Blue	
High Conflict Cell Color	Red	
Cell Display > Display Cell Color	On	Colors cells based on whether corresponding value is above or below the threshold or missing/inapplicable. If this option is turned off, all cells will be colored white (this may be desired if the values are shown, see next option).
Cell Display > Display Cell Value	Off	Includes text of values in cell. If this option is turned on, the values from the support value trees will be shown in the cells of the display tree. Missing or inapplicable values will be shown as dashes (-) if Include Missing Values (see below) is turned on.

Option	Default Setting	Explanation
Include Missing Values	Off	Indicates missing or inapplicable values by (-) when Display Cell Value option is turned on. If this is off, no text is drawn in cells with missing or inapplicable values. This option is only available if Display Cell Value is turned on.
Draw External Grid Border	On	Draws a border around the grids when turned on; no border will be drawn if option is turned off.
Format Grids for PDF Printing	On	Formats graphics so grids are displayed properly in exported PDF files. If this option is turned on, there may be slight imperfections in Mesquite's display of the grids, but users are encouraged to leave this option on if PDF-quality images are desired.
Font	Sans Serif	Only applicable if Display Cell Values is turned on.
Font Size	10	Only applicable if Display Cell Values is turned on.

Options available from the Grid Coordinator Legend, which act only on the settings for the corresponding cell, not the entire grid. To edit the settings for a cell, click on the black triangle in the corresponding grid in the Grid Coordinator Legend (Figure 4).

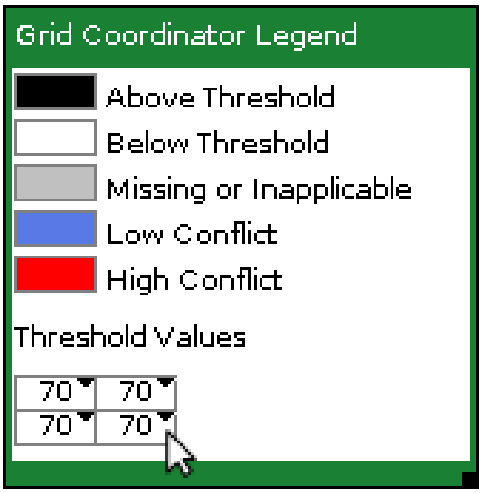

Figure 4. The Grid Coordinator Legend

Option	Explanation
Set Threshold...	Sets the threshold value to use for that cell.
Set Significant Digits...	Sets the number of significant digits (to the right of the decimal) to display if Cell Values are displayed.
Set Number For Node...	Use this option if the support values are stored in a different format than indicated during the initial setup (i.e. if you selected Branch *Lengths* for Node from Other Tree, but the support values are stored as Branch *Labels*, you can use this option to reflect the

Option	Explanation
	proper support value storage).
Tree Source (for Branch Labels/Lengths for Node from Other Tree)	Sets the source of trees to use for support value trees for that cell. You will not likely change this setting.
Choose Tree…	Chooses the tree to use for support value.
Tree Block (for Branch Labels/Lengths for Node from Other Tree)	Sets the tree block to use for support value trees for that cell. This will be necessary to use when initially setting up the Grid Coordinator (see Step II).

IV. How to cite

If you use Hypha for display purposes, cite this package:

Oliver, J.C., Miadlikowska, J., Arnold, A.E., Maddison, D.R., & Lutzoni, F. 2013. Hypha: a Mesquite package for support value integration. Version 1.0. http://mesquiteproject.org/packages/hypha.

as well as Mesquite:

Maddison, W. P. and D.R. Maddison. 2011. Mesquite: a modular system for evolutionary analysis. Version 2.75 http://mesquiteproject.org.