# DensiTree Manual:
# Making sense of sets of trees
# Version 2.1

Remco Bouckaert
`remco@cs.auckland.ac.nz`
`remco@cs.waikato.ac.nz`
`rrb@xm.co.nz`

June 3, 2013

# Contents

# 1 Introduction

Bayesian hierarchical clustering methods provide a powerful tool for phylogenetic analysis, linguistic research and hierarchical clustering in general such as applied in marketing, political science, customer preference grouping etc. One of the benefits of the Bayesian approach over classical hierarchical clustering methods such as single-link, complete link and Ward clustering is that it provides a measure of the lengths of the branches in the hierarchy. A further benefit over maximum likelihood clustering is that the Bayesian method provides insight in the distribution of possible hierarchies. Bayesian methods use MCMC sampling which results in a large number of trees representing the distribution over all possible hierarchies. Unfortunately, interpreting this distribution is not straight forward since the set of trees produced by an MCMC analysis can run in the thousands and examining them individually would be too laborious.

A popular method for analyzing tree sets are to find a single representative hierarchy and label the branches with uncertainty (for instance using the TreeLogAnnotator in BEAST [1]). The benefit of this method is that it is easy to interpret the single hierarchy by visualizing it in a tree drawing program (such as FigTree [2]) and use error bars to indicate uncertainty in branch lengths. Unfortunately, it takes some skill to interpret situations where there is uncertainty in the hierarchy. Such cases show in the tree as short branches with relatively large error bars. However, this is indistinguishable from the case where a single tree topology dominates but where there is large uncertainty due to model and/or data.

Another method for interpreting tree sets is to find subtrees (aka clades) that occur with high frequency (for example by using the TreeLogAnalyser in BEAST [1]). The number of relevant clades may become very large, especially with large datasets since the number of possible trees grows exponentially in the number of labels. Furthermore, interpreting uncertainty within high frequency clade may become cumbersome due to the large number of them.

Tree networks (as in SplitsTree [3]) are graphs containing edges wherever such edges appear (possibly at some threshold frequency) in the tree set. Tree networks do not allow easy representation of uncertainty and can become unwieldy when large numbers of distinct topologies are present in the tree set.

Here, we provide an alternative method for tree set analysis implemented in an open source tool DensiTree freely available under GPL license. DensiTree is a program for drawing sets of trees stored in Nexus format. The main idea is to draw all trees in the set, but instead of using opaque lines, we use transparency. As a result, areas where a lot of the trees agree on the topology and branch length, there will be many lines drawn and the screen will show a densely colored area. Areas where there are a few competing topologies will be highlighted by a web of lines. Uncertainty in node heights and their distribution can be shown by smears around the mean node height. Where summary trees and clade sets are quantitative approaches to tree set analysis, DensiTree provided a qualitative approach.

We start with some important concepts for understanding DensiTree in Section 1.1. The main features and analysis methods that can be performed with DensiTree are shown in the gallery (Section 3). The user interface is explained in Section 4, including the menu items and table of key short cuts. Section 5 is called FAQ and lists some common things one may want to do. Finally, command line interface is described in Section 6, which is useful for starting DensiTree with your favorite standard settings.

## 1.1 Important concepts

The complete set of trees represented in the NEXUS file will be referred to as the *set of all trees*.

The set of all trees has a limited number of topologies. For every topology, a so called *consensus tree* is calculated. The branch length of a consensus tree is calculated as the average of the branch length for all trees with the same topology. So, there are two sets of trees in DensiTree, a set of all trees and a set of consensus trees. Both can be drawn or either of them can be drawn.

There are two ways of viewing; the default is *draw all* which draws the set of trees and/or consensus trees. Alternatively, one can *browse* through the tree topologies.

## 1.2 What is new

From Version 2.01 to 2.1.10, the following changes were made:
o layout improved
o burn-in by percentage, default to 10%
o ability to hide labels
o tip tools
o user specified root canal tree

From Version 2.0 to 2.01, the following changes were made:
o Star tree, centralised tree and angle correction added.
o Button bar with tree type/style.
o List of clades, select clade by clicking in the list.
o Interactive clade movement.
o Numerics: better scales, height tracking of cursor, cumulative tree intensity reported when browsing, view clade support percentage

From Version 1.45 to 2.0, the following changes were made:
o Edit tree to manipulate rotation and height of internal nodes.
o Support of phylogeographical images.
o A choice of branch drawing methods (lines, arcs, steep arcs).
o Editing of trees be deleting taxa and saving resulting tree set.
o Pasting trees from clipboard.
o Label rotatable when the root is at top + more sensible label placement.

## 2 Getting started

To run DensiTree, you need a Java 1.6 runtime installed (for instance available from `http://java.sun.com`) and the DensiTree binary, which can be downloaded from `http://www.cs.auckland.ac.nz/~remco/DensiTree.jar`.

You can run DensiTree by calling `java -jar DensiTree.jar` from the command line. Make sure that the file DensiTree.jar is in the directory you start at.

Sorry, it's not more complex than this.

# 3 Gallery

This section shows the main features of DensiTree and highlights some of the methods useful for interpreting tree sets.

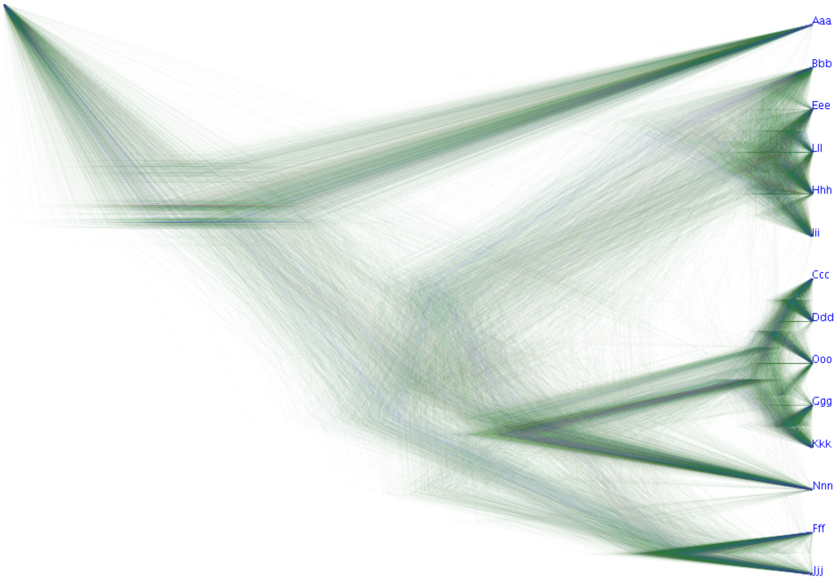


Figure 1: Default setting when opening a file. Show both consensus trees and set of all trees in triangular shape. In this tree set, there are five clearly distinguishable clades, with large uncertainty of the topologies within the two 5-leaf clades.

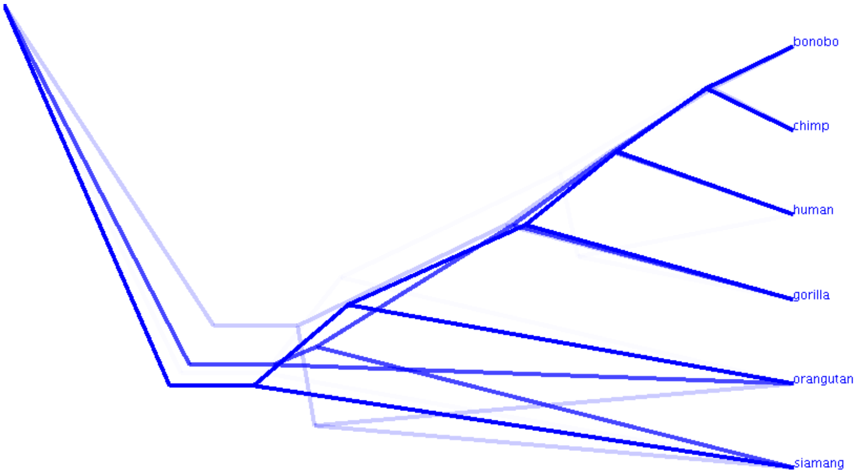


Figure 2: Show only consensus trees. This set shows that there is very little uncertainty in the topology of most of the tree, except for the few splits near the root.

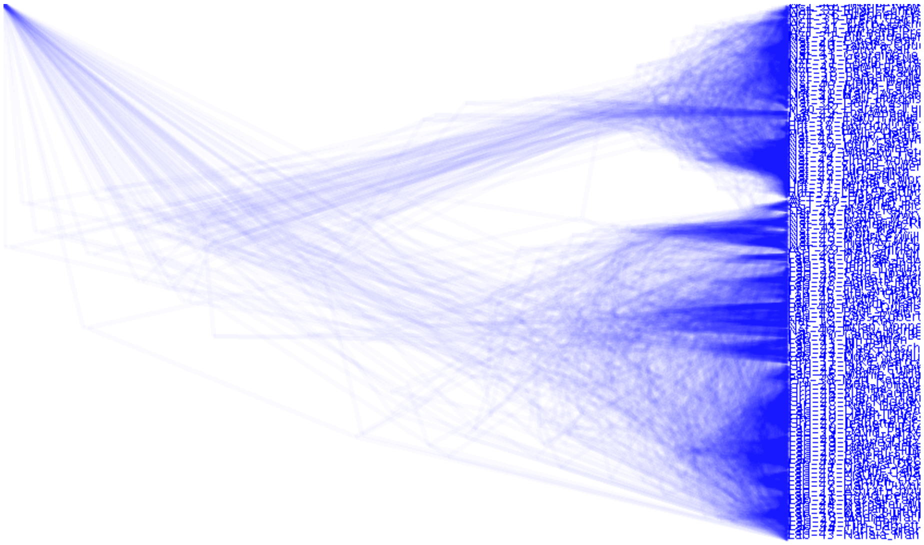


Figure 3: Show only consensus trees. This highlights the uncertainty inside the clades, but shows that the split at the root into two groups is very certain (split into progressive and conservative politicians).

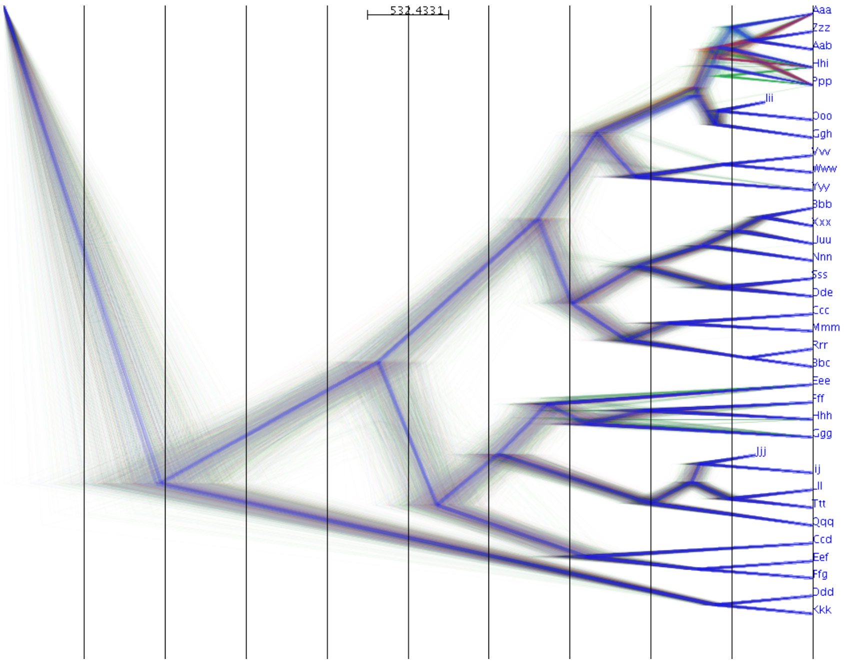


Figure 4: Show tree height by height grid and height bar. This tree set nicely demonstrates the increase in uncertainty of the node heights going from the leafs to the root.

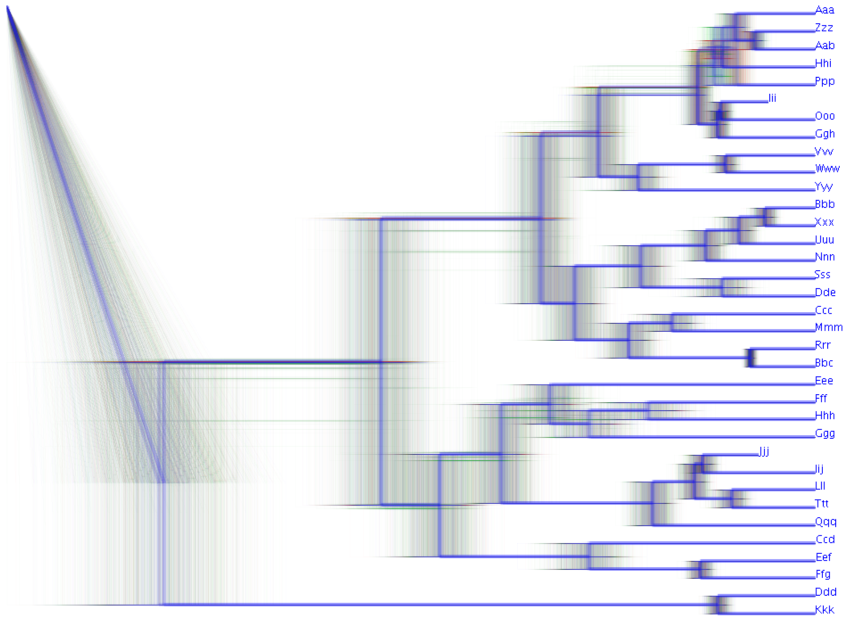


Figure 5: As Figure 4 but in block trees. This tree set was generated with calibration points, which show up as dense node heights, for example, the parent of Rrr and Bbb.

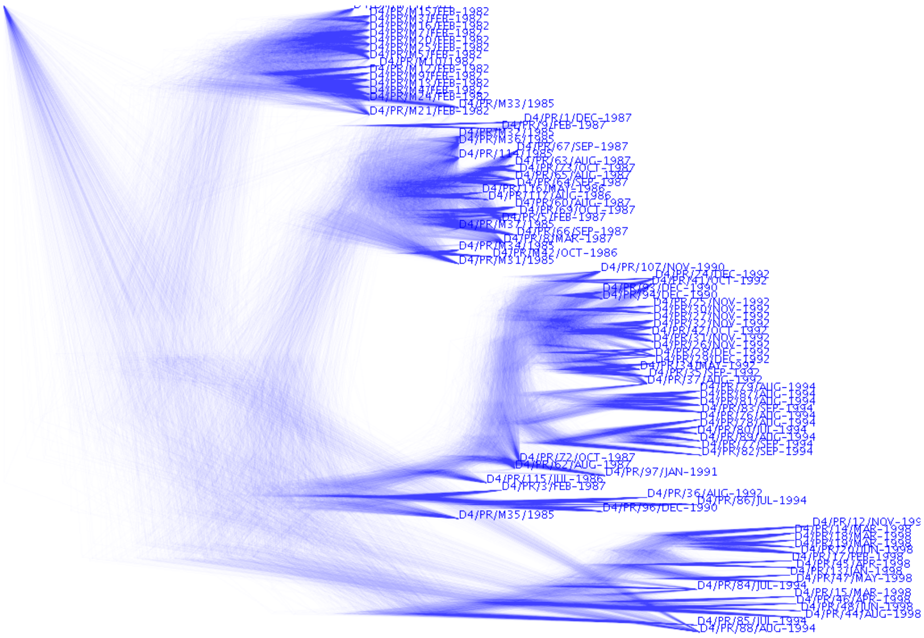


Figure 6: Decreased width of consensus trees, only consensus trees drawn. Intensity of consensus trees needed to be increased considerably. This is useful when there is large uncertainty in the topology and hence many consensus trees (over 900 in this example) with little overlap. Without intensity increase they would not show up and only a white image would be shown.

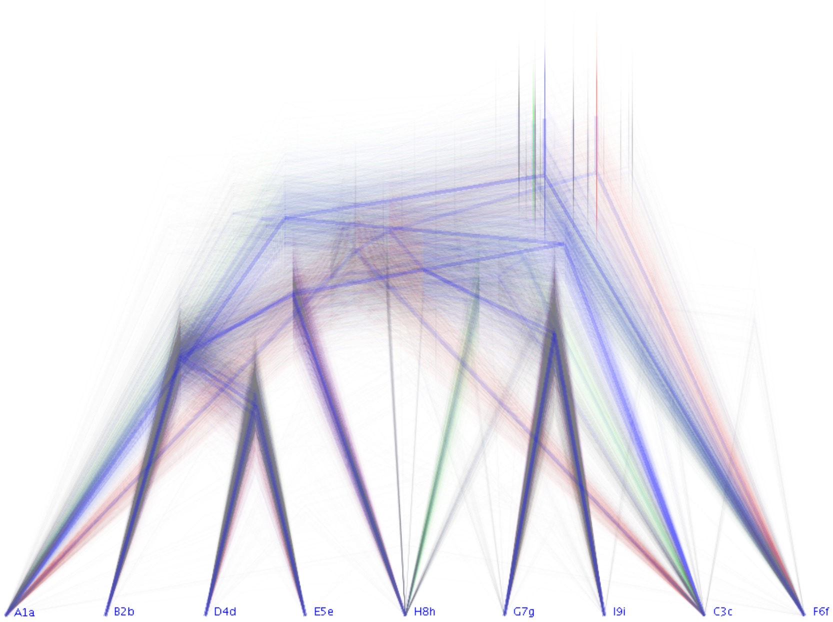


Figure 7: Triangle tree with root at top. This can be an attractive option when there are not a lot of leafs so that the labels do not overlap.

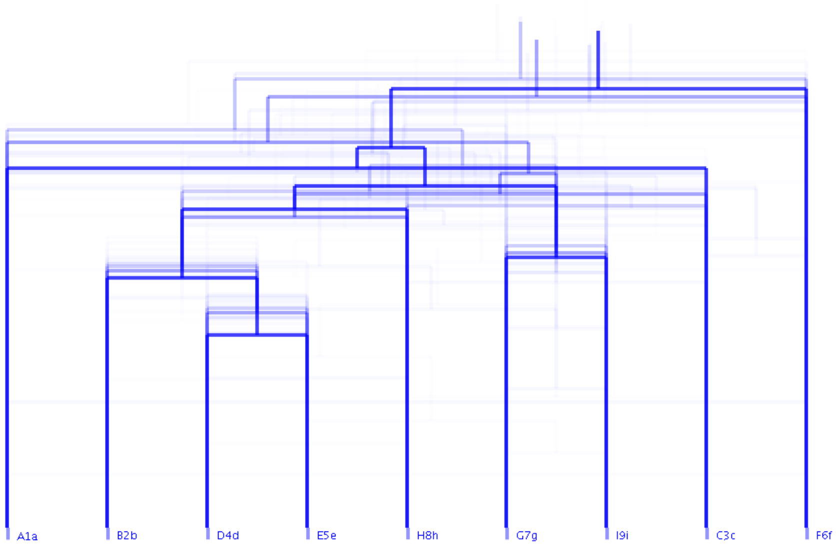


Figure 8: Increase intensity consensus trees. Consensus trees are drawn with intensity proportional to the frequency of occurrence of the topology in the set. As a result, low frequency topologies will be drawn very faintly and may become invisible. By increasing the global intensity, these faint trees can be made more visible.

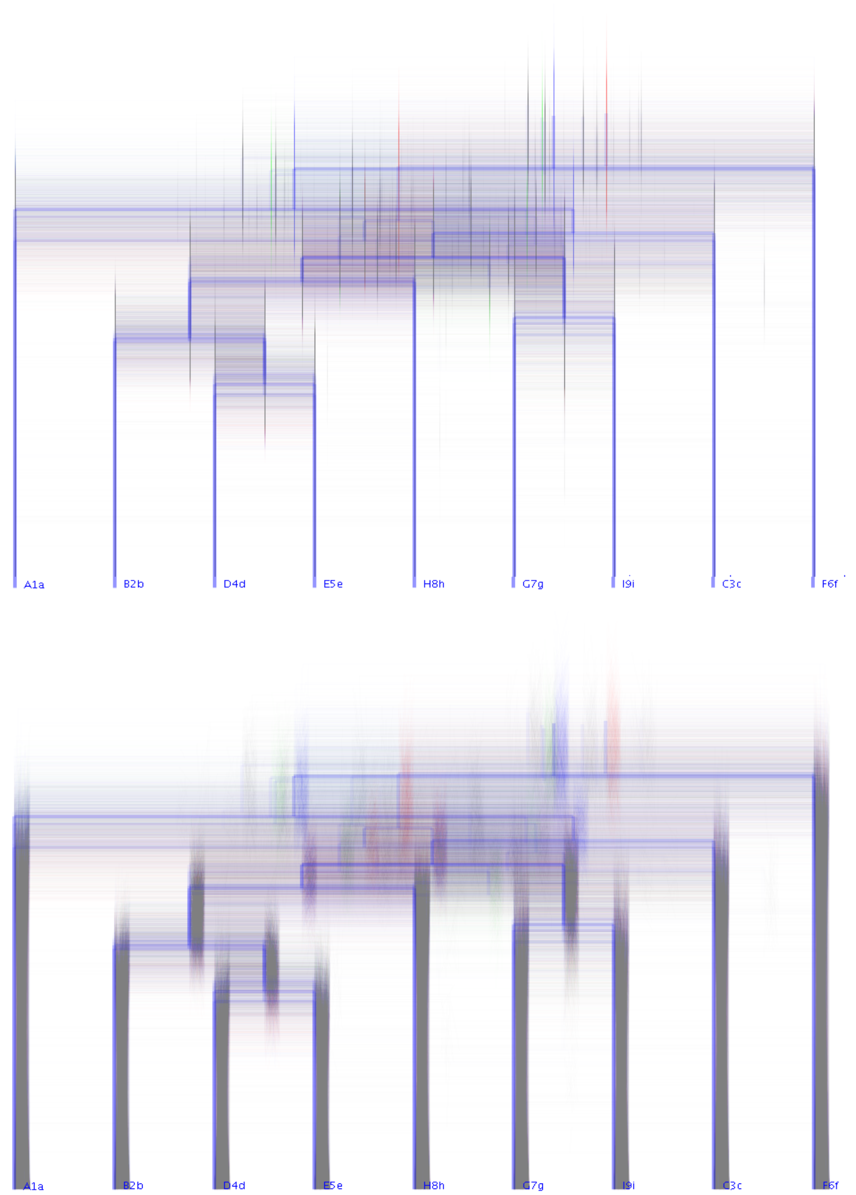


Figure 9: Jitter x-axis. Especially the block-layout of a tree results in many overlapping lines in the direction from the leafs to the root. These show up as dense stripes in the image (see top image). By adding jitter to the begin and end points, the lines get randomly perturbed with a small amount and the image gives a better sense of the densities (bottom image).

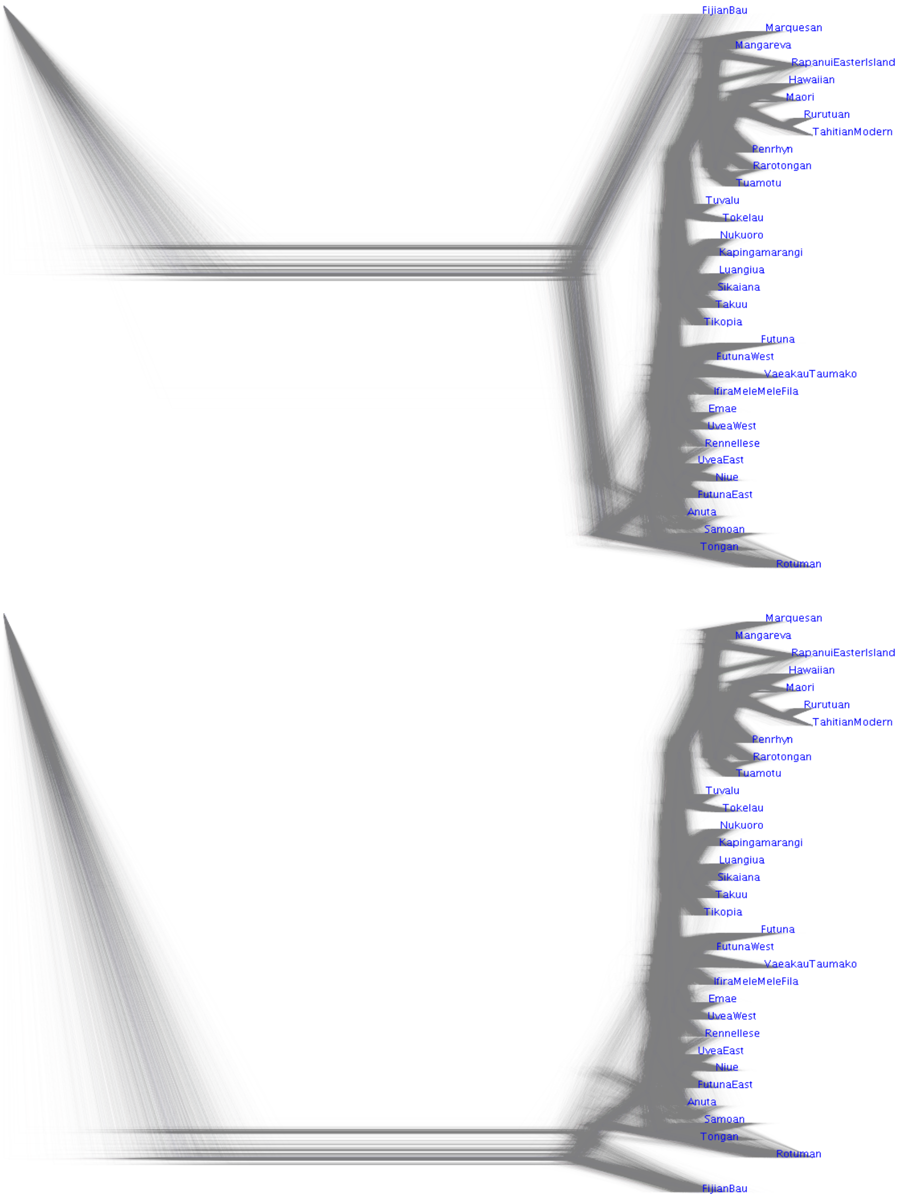


Figure 10: Move leaf for better layout on Polynesian language data [4]. In this case, the leaf at the top is moved to the bottom, resulting in the long branch to that leaf being shortened considerably. In situations where there is a lot of uncertainty in the topology near the leafs it is often possible to move a few leafs a bit to minimize the number of crossing lines and get a clearer view.

Figure 11: Zoom in (Data from [4]).

Figure 12: Draw trees up from selected leafs. Only tree branches with a direct path from the root to a selected leaf will be drawn. Data from [4].

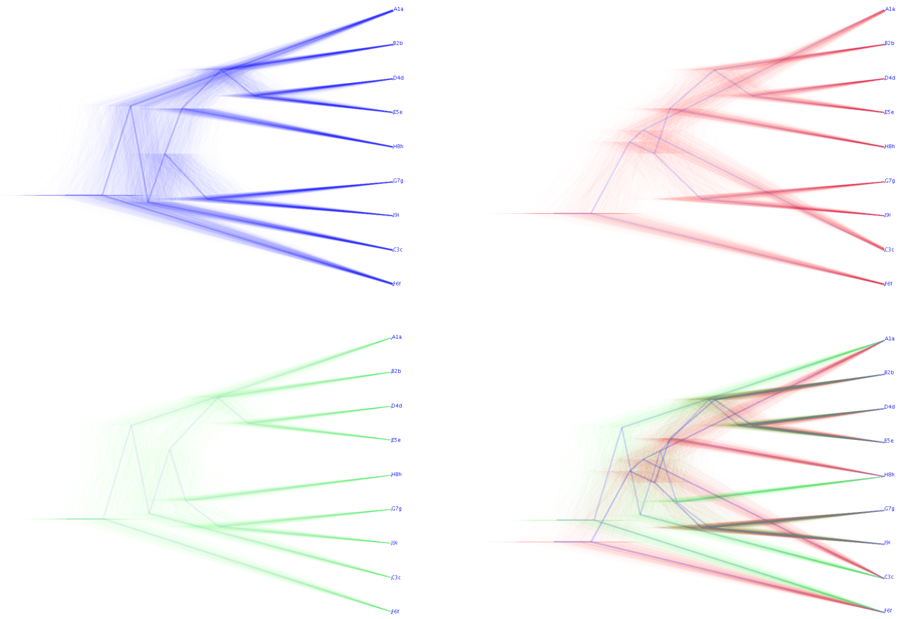


Figure 13: Show selection of tree topologies. Top left shows the most likely tree topology, and the consensus tree with that topology drawn in. Top right shows the second most popular topology. Bottom left the third most likely. Bottom right shows the second and third most likely tree topologies, leaving out the most likely topology.

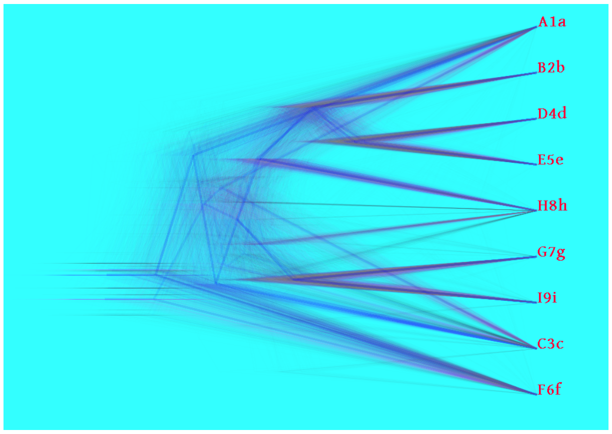


Figure 14: Not recommended, but possible: an example of color and font customization. Here, font size is increased to 48 point, color of all trees set to red, color of consensus trees to green and background color to yellowish green.

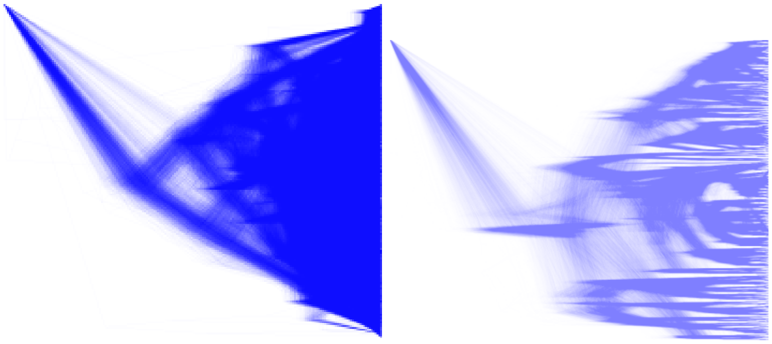


Figure 15: When there are many topologies, each occurring with low frequency, the result may be a mess (left image) because the order of leafs is chosen based on the most frequently occurring consensus tree. To untangle the mess, various shuffle methods are available under the Edit/Shuffle submenu (result in right image).

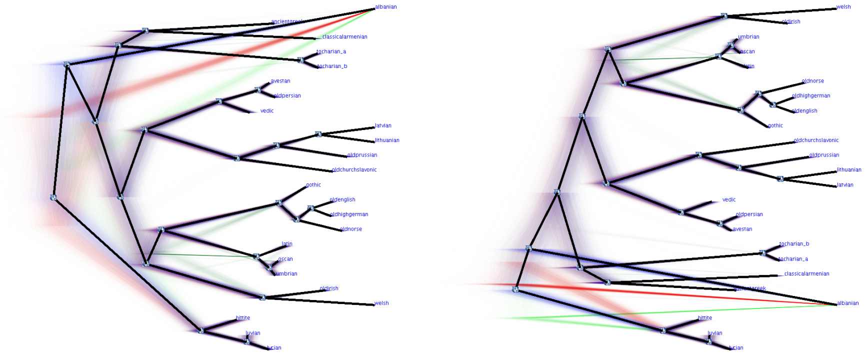


Figure 16: When edit-tree is shown (through menu Edit/Show edit tree) the order of taxa can be manipulated by clicking on the rotation icons drawn on the internal nodes of the black tree.

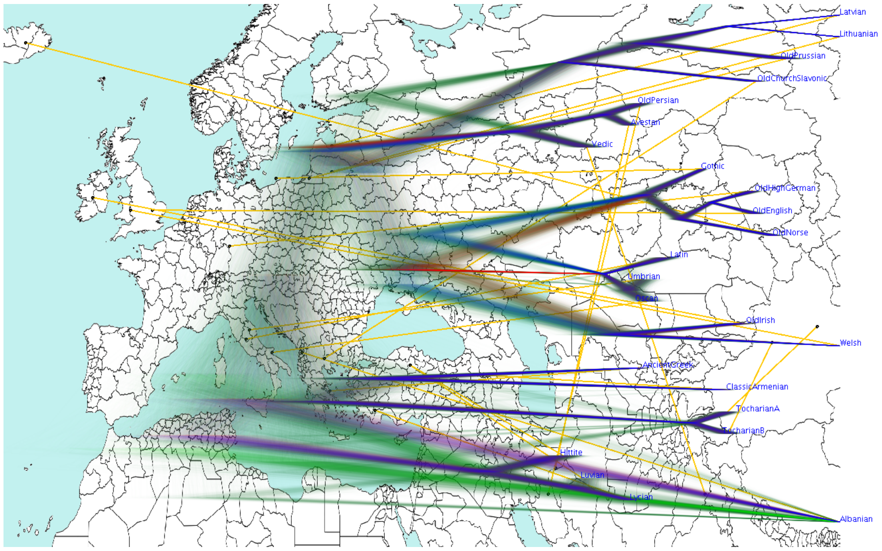


Figure 17: Phylogeographical DensTree where the taxa are connected with their geographical locations and a background image showing a relevant map.

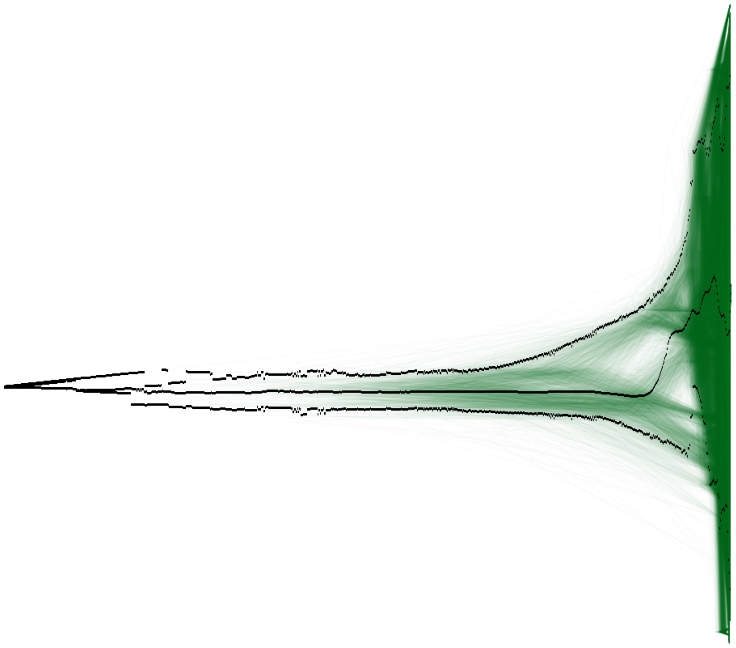


Figure 18: Drawing metadata embedded in the tree set to determine the y-position of the internal nodes. Black dots at the 90% highest probability density interval and the median.

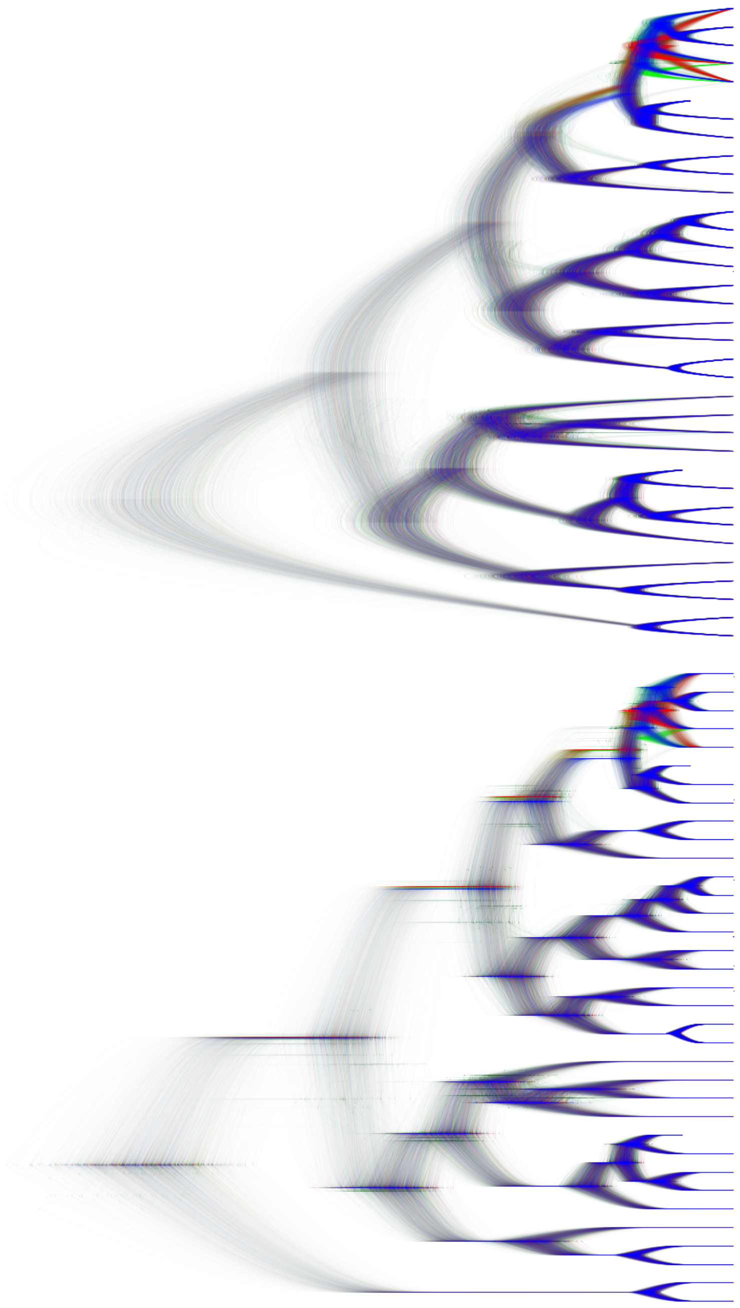
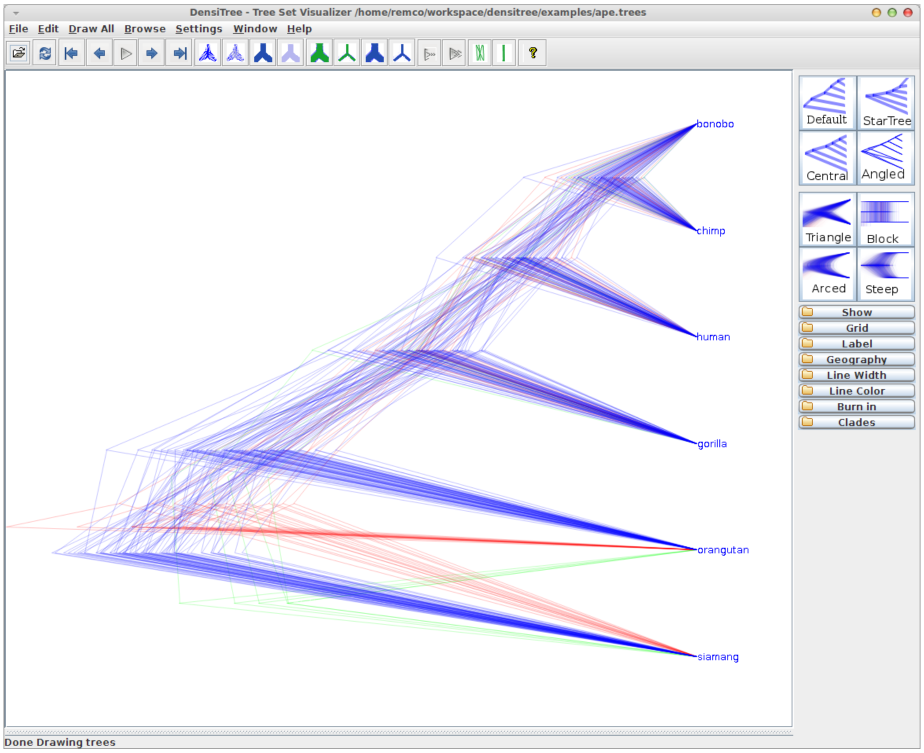


Figure 19: Different methods of drawing branches. Arcs at the top, steep arcs at the bottom.

# 4 GUI

When opening DensiTree the following screen appears with a main window containing a menu, tool bar, status bar.



Done Drawing trees

Figure 20: Main window

## 4.1 Menu items

DensiTree has the following menu items and short description of its function.

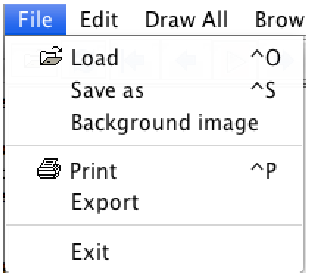


Figure 21: File menu

`File/Load`: open Nexus trees file containing set of trees
`File/Save as`: Save current tree set as Nexus tree file. Only useful after editing taxa in the tree
`File/Background image`: open bitmap file to be shown in the background. If a KML file is loaded, this is considered to be a world map, and only part of the image relevant to the locations of the taxa is shown. If the KML file has a file name of the form ”XYZ(<lat1>,<long1>)x(<lat2>,<long2>).png” for example ”NewZealand(-60,140)x(-10,180).png” then the image is considered to cover only the rectangle with corners (<lat1>,<long1>) and (<lat2>,<long2>). Not lat1 < lat2 and long1 < long2.
`File/Print`: print currently shown view of the tree set (untested)
`File/Export`: export currently shown view of the tree set in bitmap in BMP, JPG, PNG and SVG format. Note that the zoom factor (menu `Window/Zoom in` and `.. out`) has impact on the resolution of the image. For hight resolution images, zoom in more, then redraw, then export.
`File/Exit`: quit DensiTree.

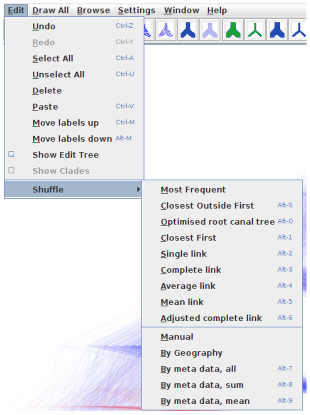


Figure 22: Edit menu

`Edit/Undo`: undo latest reordering of taxa. Note that the undo action only applies to reorderings, not to font changes, intensity settings, layout options, etc.
`Edit/Redo`: redo latest reordering of taxa.
`Edit/Select All`: select all leafs in the tree. This is useful when after manipulating the order of the leafs a redraw is required with branches to all leafs shown.
`Edit/Unselect All`: remove all leafs from selection. This is useful before moving a single node.
`Edit/Delete taxa`: remove selected taxa from tree. Branches between a taxon and its parent will be removed one by one and the internal parent node removed.
`Edit/Paste trees`: paste one or more trees in Nexus format from the clipboard.

`Edit/Move labels up`: move selected labels one higher in the ordering. It is recommended to use the short cut key `Ctrl-M` if a large number of moves need to be made
`Edit/Move labels down`: move selected labels one lower in the ordering. It is recommended to use the short cut key `M` if a large number of moves need to be made
`Edit/Show Edit Tree`: Show a tree to allow reording of nodes by selecting the rotate icons drawn on the internal nodes (see Figure 16.
`Edit/Show Clades`: Show a clades by drawing a circle with a radius proportional to its support. This allows clades to be moved when drawing a centralised tree.
`Edit/Shuffle`: Submenu with various methods for ordering the nodes.
`Edit/Shuffle/Most Frequent`: Use order that displays most frequently occurring tree nicely. This is the default used when opening a file.
`Edit/Shuffle/Closest First`: Orders leafs by starting with the closest two leafs, then adding the closest node to the left most or right most node. The distance measure used is based on the length of the edges averaged over all trees.
`Edit/Shuffle/Single link`: Use single link hierarchical clustering with the distance method as for 'closest first' and use an order that displays the obtained hierarchy pleasingly.
`Edit/Shuffle/Complete link`: As single link, but using complete link.
`Edit/Shuffle/Average link`: As single link, but using average link.
`Edit/Shuffle/Mean link`: As single link, but using mean link.
`Edit/Shuffle/Adjusted complete link`: As single link, but using adjusted complete link.
`Edit/Shuffle/Manual`: Key in order of nodes by hand.
`Edit/Shuffle/By Geography`: Only useful when geographic locations are loaded for the taxa. Orders nodes by longitude if root at top, or latitude otherwise.
`Edit/Shuffle/By meta data, all`: if meta data is available matching a pattern (can be provided with the -pattern command line option) the tree is draw with internal nodes at the level of the meta data value.
`Edit/Shuffle/By meta data, sum`: as above but the sum of values at each height in the tree is used as level.
`Edit/Shuffle/By meta data, mean`: as above but mean value at each height is used.

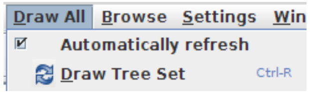


Figure 23: Draw all menu

`Draw all/Automatic refresh`: when selected, the tree set will be redrawn whenever a setting changes. If not selected, an explicit redraw needs to be done (using short cut key `R` or menu `Draw all/Draw Tree Set`) to draw the tree set after manipulating some settings.

`Draw all/Draw Tree Set`: draw tree set when automatic refresh is off, and switches to drawing mode when in browsing or animation mode.

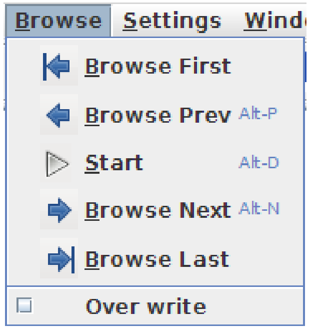


Figure 24: Browse menu

`Browse/Browse First`: Shows first consensus tree and/or accompanying trees from set of all trees. If in drawing mode, switch to browse mode. If in animation mode, animation is stopped.
`Browse/Browse Previous`: Show previous consensus tree. If it is the first tree, or in over write mode (see menu `Browse/Overwrite`), the screen is cleared before drawing the tree set. Otherwise, the trees are drawn over already drawn trees.
`Browse/Start`: Start/stop animation.
`Browse/Browse Next`: Show next consensus tree. If in over write mode (see menu `Browse/Overwrite`), the screen is cleared before drawing the tree set. Otherwise, the trees are drawn over already drawn trees.
`Browse/Browse Last`: Shows last consensus tree, clearing the screen if required.
`Browse/Overwrite`: When not set (default) the trees are drawn on top of each other when browsing the trees. This way, a subset of tree topologies can be drawn. If set, the screen is cleared and only a single topology is drawn when browsing or animating.

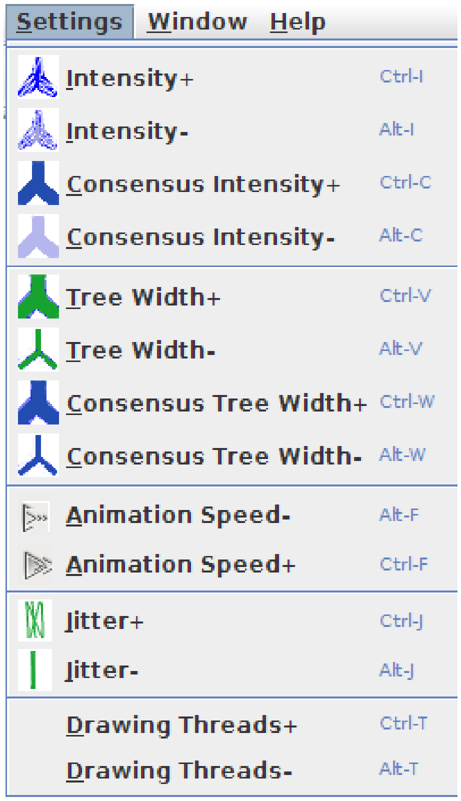


Figure 25: Settings menu

`Settings/Intensity- and +`: decrease/increase tree intensity.
`Settings/Consensus Intensity- and +`: decrease/increase consensus tree intensity.
`Settings/Tree width- and +`: decrease/increase tree line width.
`Settings/Consensus Tree width- and +`: decrease/increase consensus tree line width.
`Settings/Animation speed- and +`: decrease/increase animation time delay - shorter delay = faster animation.
`Settings/Jitter- and +`: decrease/increase jitter on trees (not consensus trees).
`Settings/Drawing threads- and +`: decrease/increase number of drawing threads for drawing tree set.

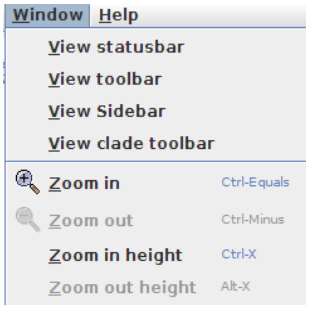


Figure 26: Window menu

`Window/View Statusbar`: Toggle visibility of status bar
`Window/View Toolbar`: Toggle visibility of tool bar at the top
`Window/View Sidebar`: Toggle visibility of tool bar at the side
`Window/View Clade Toolbar`: Toggle visibility of tool bar at bottom containing clade information
`Window/Zoom in`: Enlarge drawing. Note that this has an impact on the resolution when exporting bitmap images.
`Window/Zoom out`: Draw smaller image.
`Window/Zoom in height`: Enlarge drawing, but only in the direction of the tree branch heights. Again, this has an impact on the resolution when exporting bitmap images.
`Window/Zoom out height`: Draw smaller branch heights.

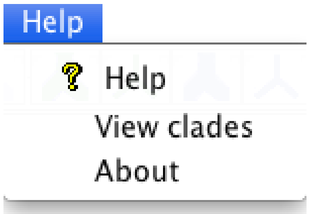
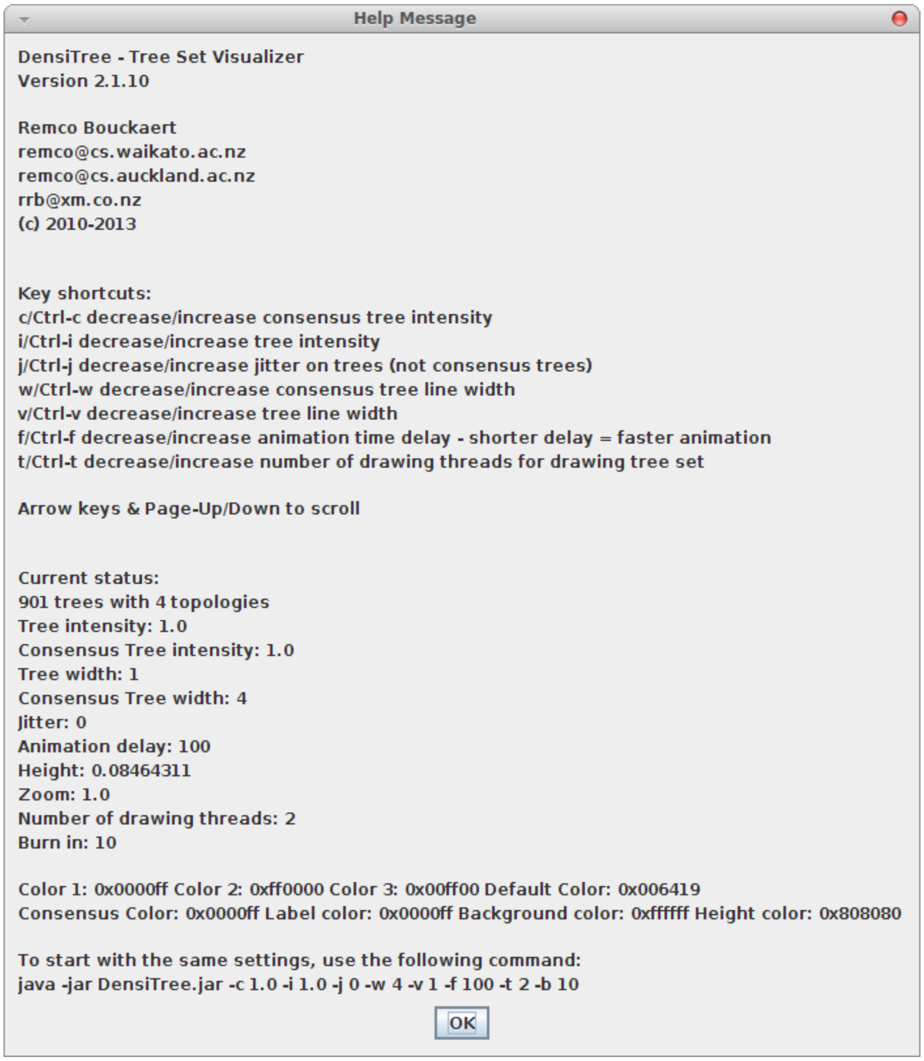


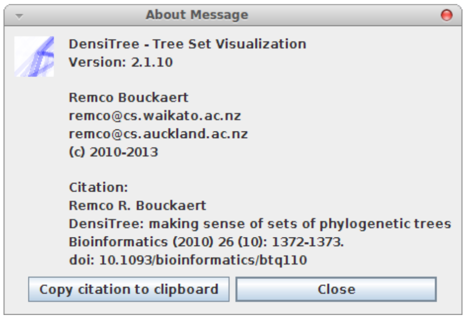
Figure 27: Help menu

`Help/Help`: Show short description as shown below.
`Help/View Clades`: Show clades in text entry, so the contents can be copied to clipboard.
`Help/About`: Show version and citation info.



```
DensiTree - Tree Set Visualizer
Version 2.1.10

Remco Bouckaert
remco@cs.waikato.ac.nz
remco@cs.auckland.ac.nz
rrb@xm.co.nz
(c) 2010-2013


Key shortcuts:
c/Ctrl-c decrease/increase consensus tree intensity
i/Ctrl-i decrease/increase tree intensity
j/Ctrl-j decrease/increase jitter on trees (not consensus trees)
w/Ctrl-w decrease/increase consensus tree line width
v/Ctrl-v decrease/increase tree line width
f/Ctrl-f decrease/increase animation time delay - shorter delay = faster animation
t/Ctrl-t decrease/increase number of drawing threads for drawing tree set

Arrow keys & Page-Up/Down to scroll


Current status:
901 trees with 4 topologies
Tree intensity: 1.0
Consensus Tree intensity: 1.0
Tree width: 1
Consensus Tree width: 4
Jitter: 0
Animation delay: 100
Height: 0.08464311
Zoom: 1.0
Number of drawing threads: 2
Burn in: 10

Color 1: 0x0000ff Color 2: 0xff0000 Color 3: 0x00ff00 Default Color: 0x006419
Consensus Color: 0x0000ff Label color: 0x0000ff Background color: 0xffffff Height color: 0x808080

To start with the same settings, use the following command:
java -jar DensiTree.jar -c 1.0 -i 1.0 -j 0 -w 4 -v 1 -f 100 -t 2 -b 10
```

Figure 28: Show help and other useful information



```
DensiTree - Tree Set Visualization
Version: 2.1.10

Remco Bouckaert
remco@cs.waikato.ac.nz
remco@cs.auckland.ac.nz
(c) 2010-2013

Citation:
Remco R. Bouckaert
DensiTree: making sense of sets of phylogenetic trees
Bioinformatics (2010) 26 (10): 1372-1373.
doi: 10.1093/bioinformatics/btq110
```

Figure 29: Show version and citation information

## 4.2 Toolbar

The toolbar () has the following functions:

open file,

draw all,

browse to first consensus tree,

browse to previous consensus tree,

start animation,

browse to next consensus tree,

browse to last consensus tree,

increase/decrease tree intensity.

increase/decrease consensus tree intensity.

increase/decrease tree line width.

increase/decrease consensus tree line width.

decrease/increase animation time delay - shorter delay = faster animation.

decrease/increase jitter on trees (not consensus trees).

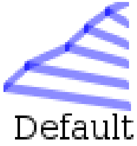
and

show short help and status.

The tool bar can be hidden or made visible again using the `Window/View Toolbar` menu.

## 4.3 Sidebar

The sidebar is by default shown at the right side of the window, but can be dragged to another place. It contains the following functions:

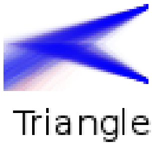


Place nodes in the middle of the top of its two child clades.

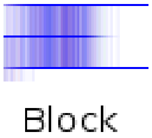
Position nodes as star tree.

Position nodes halfway the utmost left and utmost right taxon in the clade.

As above, but put clades at same position as its parent thus forming nodes with more than two child clades.



draw as triangle tree.



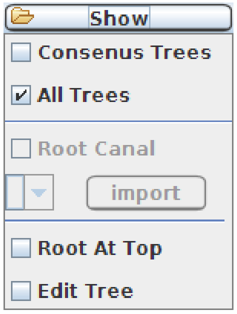
draw as block tree.

draw with arcs.

draw with steep arcs.

The side bar can be hidden or made visible again using the `Window/View Sidebar` menu.

### 4.3.1 Show Panel



Consensus Trees: Display consensus trees. There is one consensus tree for every topology in the tree set. The height of the nodes are the average of the heights for that topology.

All Trees: Show all trees in the tree set.

Root Canal: Show root canal tree. This is a single summary tree representing the complete tree set. There are many ways to construct a summary tree.

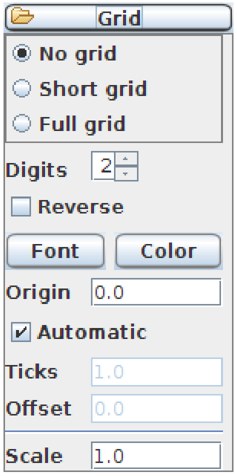
Root Canal Number: Select root canal tree to display.

Import: Import root canal tree from Newick or from the summary_tree program.

Root At Top: Display the root at the top of the display instead of on the left hand side.

Edit Tree: Display edit tree for manipulating order of tree and position of internal nodes. Works only with default drawing style.

### 4.3.2 Grid Panel



● No grid
○ Short grid
○ Full grid

Digits 2

☐ Reverse

Origin 0.0

☑ Automatic

Ticks 1.0

Offset 0.0

Scale 1.0

Grid: Show lines indicating timescale. Options are to show none, short lines at the side of the panel, or full lines over the complete tree set.

Digits: Set number of significant digits for the grid labels.

Reverse: By setting reverse, the time scale will be drawn forward in time. By default, time scale is drawn backward in time, so that the height of a tree is a positive number. Also, set 'origin' to the date of the youngest tip.

Font: Set font of the grid labels.

Color: Set colour of the grid labels.

Origin: Set date of the youngest tip.

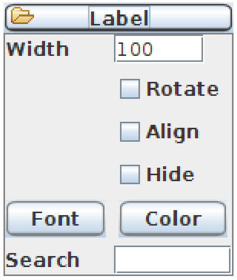
Automatic: Automatically determine the number of ticks.

Ticks: Interval between two ticks.

Offset: Time added to the ticks. This can be usefull when the youngest tip of the tree is on a number that is not quite a round number, for example 2003.4. Setting the offset to -3.4 ensures the grid lines will be drawn on 2000 instead of through 2003.4.

Scale: Scale time, which can be handy when the tree is in substitutions and a clock rate is available from the literature. A negative scale has the same effect as selecting 'reverse' with a positive scale.

### 4.3.3 Label Panel



Width 100
☐ Rotate
☐ Align
☐ Hide

Label Width: Width of the label.

Rotate: Rotate label – is only effective when root at top.

Align: Align labels with label for youngest tip. This is only useful when tips are not all from the same date.

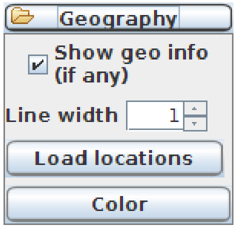
Hide: Hide labels.

Font: Font used for labels.

Color: Color used for labels.

Search: Search for labels. Labels matching the search string will be selected/highlighted.

### 4.3.4 Geo Panel



☑ Show geo info (if any)
Line width 1

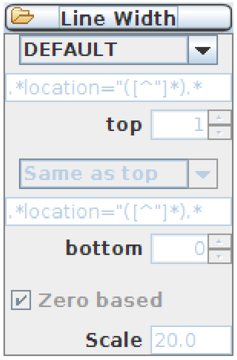
Show Geo Info: Show lines linking tip labels with geographic location of tip sample. Line only show up if goegraphic locations are specified using the 'load locations' function.

Line Width: Width of the line used to link tips with geo locations.

Load Locations: Load locations from KML file. The locations can be specified in google earth and saved in a KML file.

Color: Color of the line used to link tips with geo locations.

### 4.3.5 LineWidth Panel



Line Width

DEFAULT

.*location="([^"]*).*

top 1

Same as top

.*location="([^"]*).*

bottom 0

☑ Zero based

Scale 20.0

Line Width: Determine line width of trees of both tree set and consensus trees.For consensus trees, the average value for the topology is used.

DEFAULT: all lines are same width.

BY_META_DATA_PATTERN: use value of pattern specified below.

BY_META_DATA_NUMBER: use the N-th attribute value in the meta data.

meta data attribute: only available if any meta data attribute is specified. Use value of the attribute for line width of branches.

Pattern: Regular expression used for width of branches when BY_META_DATA_PATTERN is chosen. The string of the pattern between brackets is selected as value.;

Top: Specifies N-th meta data attribute for top of branch when BY_META_DATA_NUMBER is seleced.

Line Width Bottom: Line width at bottom of branch.

same as top: use same specification as for top of branch.

Fit to bottom: adjust bottom widths so they fit to top of branch below.

BY_META_DATA_PATTERN: use value of pattern specified below.

BY_META_DATA_NUMBER: use the N-th attribute value in the meta data.

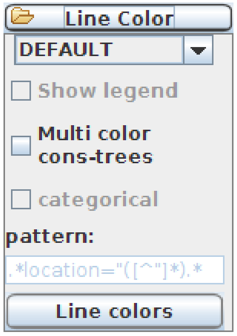
meta data attribute: only available if any meta data attribute is specified. Use value of the attribute for line width of branches.

Bottom: Specifies N-th meta data attribute for top of branch when BY_META_DATA_NUMBER is seleced.

Zero Based: If selected, the minimum value is zero, otherwise the minimum value of the range of whatever value is used.

Scale: Scale width with this number.

### 4.3.6 Color Panel



Line Color

DEFAULT

☐ Show legend

☐ Multi color cons-trees

☐ categorical

pattern: .*location="([^"]*).*

Line colors

Line Color: Determines line color for the complete tree set. So, this does not affect the consensus trees or root canal tree.

DEFAULT: color 1 for most frequently occurring topology, color 2 for the second most popular, color 3 for the third, and color 4 for the remaining trees. Colors can be changed using the line colors button.

COLOR_BY_CLADE: draw clades in one color.

COLOR_BY_META_DATA_PATTERN: draw trees matching the regular expression specified in the pattern entry below.

meta data attribute: only available if any meta data attribute is specified. Use value of the attribute to color branches.

Show Legend: Show legend mapping colors to attribute values in the DensiTree. This only works when a discrete attribute is selected for line coloring.

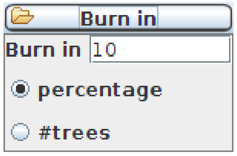
Multi Color Consensus Trees: Use different colours for consensus trees instead of the standard color.

Categorical: Interpret value of attribute as categorical data.

Pattern: Regular expression used for coloring trees when COLOR_BY_META_DATA_PATTERN is chosen. The string of the pattern between brackets is selected as value.

Line Colors: Specify custom colors.

### 4.3.7 Burnin Panel



Burn in

Burn in 10

● percentage

○ #trees

Burnin: Specifies the set of trees at the beginning of the set that are removed from the tree set. When the tree set represents a sample from an MCMC run, typically about 10% of the trees are sampled while the chain is in burn-in, and are not representative for the tree distribution.

Focus and press enter to reload file with adjusted burn-in settings.

Percentage: If selected, the burn-in is interpreted as a percentage, hence should be in between 0 and 100. If the burn-in falls outside that range, burn-in is reset to 10.

Number Of Trees: If selected, the burn-in is interpreted as the number of trees at the start of the set that should be removed. If burn-in is larger than the number of trees in the set, burn-in is reset to 0.

### 4.3.8 Clade Panel

☐ Show clades

☐ Selected only

Mean ☐ ☐

95%HPD ☐ ☐

Support ☑ ☐

Sign. Digits 3

Show Clades: Show information for individual clades. This is only activated when the drawing style is not default. The default style is not clade based, so there is no information to position clade information.

Selected Only: Show information only for selected clades. Clades can be selected by clicking them in the DensiTree, or selecting them from the clade-bar (at botom of the screen).

Mean: Show mean height of clades as line or and/as text.

95hpd: Show 95% highest probability density interval of the height of clades as bar and/or as text.

Support: Show support of clade as cricle and/or as text. The support is the fraction of trees in the tree set that contain the clade.

Digits: Number of significant digits to show clade information as text.

Font: Font used to show clade information as text.

Color: Color used to show clade information as text.

## 4.4 Clade toolbar

The clade toolbar at the bottom of the screen lists clades sorted by support in the tree set. When 'view clades' is set to true, clades can be selected in the list, and the selection changes. Also, when a clade is selected in the tree set by dragging it in a rectangle, the selected clades are highlighted in the list.

The clade toolbar can be hidden or made visible again using the `Window/View Clade Toolbar` menu.

## 4.5 Statusbar

The main function of the status bar at the bottom of the screen is to show the progress when drawing all trees and to show the number of the topology when browsing through the consensus trees.

When the mouse is moved, the height at the mouse position is displayed in the status bar.

When all trees are drawn, the numbers in the status bar count down to zero. When also consensus trees are drawn, the count goes up. When drawing is complete, the message 'Done Drawing Trees' appears in the status bar.

The status bar can be hidden or made visible again using the `Window/View Statusbar` menu.

# 5 FAQ

## 5.1 How do I move leafs?

This depends on the tree type. For default and centralised trees, first you select the leafs that you want to move. You can select leafs by clicking on them individually, or by dragging a rectangle over the set of leafs you want to select. By keeping the shift key pressed, the selection can be added to. By keeping the control key pressed, any leaf that is selected will be toggled from the selection. Once the appropriate set of leafs is selected, press M to move down and control-M to move in the other direction.

Furthermore, for star trees you can enable show clades. Circles appear on the taxa, which you can select by clicking on it, or drag a rectangle over it. The selection can be dragged. This allows you to position the taxa at irregular intervals.

## 5.2 I'm loading a large file, but nothing seem to happen?!?

If nothing happens after a few minutes, most likely it is java that has run out of memory. To start DensiTree with access to more memory, start it from the command line with

```
java -Xmx3g -jar DensiTree.jar
```

Here '3g' means 3 gigabyte. If you need even more, change the 3 to a higher number.

## 5.3 How do I start DensiTree with my favorite settings?

First, find out the values of the current settings. This can be done using the Help/Help menu, which makes a dialog pop up that shows the current status. At the bottom, it also shows how to start up DensiTree with these settings from the command line.

## 5.4 How do I specify burn in?

Start DensiTree from the command line with the -b option, e.g. using

```
java -jar DensiTree.jar -b 100
```

This will skip the first 100 trees every time a nexus tree is read. To check current burn in, select the Help/Help menu. The dialog shows the current state.

## 5.5 What does the color coding mean?

The most popular tree is blue, the next most popular red, the third most popular green and the rest is dark green.

Consensus trees are blue by default as are labels. Height bar and grid are black by default.

## 5.6 How do I change these colors?

Colors can be changed using the Settings/Set Color submenu. To start DensiTree with specific colors, colors can be specified from the command line or via a script as follows:

```
java -Dcolor.1=0x000000 -jar DensiTree.jar
```

sets the color of the most popular tree to black. The part 'color.1' specifies that it is the most popular color, and 0X000000 specifies the RGB value in hexadecimal.

The current color settings can be found using the Help/Help menu, where they are shown near the bottom of the dialog that pops up then.

The following colors can be specified that way:

`color.1` for most popular topology,
`color.2` for next most popular topology,
`color.3` for third most popular topology,
`color.default` for remainder of topologies,
`color.cons` for consensus trees,
`color.label` for color of labels,
`color.height` for color of height grid and height bar,
`color.bg` for back ground color.
`color.rootcanal` for root canal color.

So, to make consensus trees red and labels black, you start DensiTree like this:

```
java -Dcolor.cons=0xFF0000 -Dcolor.label=0X000000 -jar DensiTree.jar
```

## 5.7 How do I change the label font?

Use the Settings/Set Font menu.

## 5.8 Can I batch process tree set files?

Use the command line options to specify your settings, and use the

```
-o <filename>
```

option to specify a file. The DensiTree is exported as png file to the specified file.

## 5.9 My tree is a mess, what do I do?

Check that you have a clock like tree. DensiTree may work for non-clock like trees, but was not designed to do so. Otherwise, try one of the tricks in the next question.

## 5.10 My tree has many crossing lines, what do I do?

o Step 1: Try the reshuffle using the Edit/Shuffle menu and try the various methods. Most methods work on some tree sets, but none work for all.
o Step 2: If the tree looks half decent, use the tree editor to rotate parts of the tree. You can show the tree editor using the Edit/Show Tree Edit menu. To rotate part of the tree, click on an icon at an internal node.
o Step 3: The location of internal nodes can be changed by dragging them with the right mouse button. This can help getting rid of some crossing branches.

## 5.11 How can I visualise meta data?

When there is meta data on the branches, such as populations sizes, branch rates or migration rates, DensiTree can be used in two ways to visualise this meta data.

1. By positioning all internal nodes at the height that the meta data value indicates. Press 7, 8 or 9 to get such a tree. The difference between the methods is that 7 shows the meta data values for each individual branch, 8 shows the mean meta data value at a particular height in the tree and 9 shows the sum of the meta data values.
2. By drawing lines with a width proportional to the meta data value.

## 5.12 How can I specify which meta data is used?

The simplest way is when you know in which order the metadata is inserted in the tree set file. Say, meta data is of the form

```
[&rate=0.1234,popsize={13.31, 11.01}]
```

then there are three numbers: one for rate, and two fro popsize. To make the bottom of a branch equal to the rate, we need the first number. With menu settings/meta data/Bottom pattern number, a dialog pops up where you can enter '1' for the rate, '2' for the first popsize, and '3' for the last pop size.

For experts, under menu settings/meta data/pattern, you can specify a regular expression that will be matched with the meta data in the tree file. This pattern will be used by both methods of meta data visualisation.

## 5.13 How can I specify which meta data is used for the bottom and top of a line?

With menu settings/meta data/Top pattern number, you can specify the number in the meta data. Note that the top of a branch need to be before the bottom of a branch.

Also, for experts, as in the previous question, but define two groups in the regular expression. The first group is used at the bottom of the branch and the second group at the top.

## 5.14 How do I create a phylogeographical DensiTree?

o Step 1: load a tree set file (Menu File/load)
o Step 2: load geographical locations (Menu File/Load geographic) encoded in a KML or KMZ file as produced in Google earth.
o Step 3: load a bitmap image (Menu File/Background image) with a world map. Alternatively, load a file of a map in Mercator projection which has the corners of the rectangle covered by the image encoded in its file name (see description of File/Background image for details).

## 5.15 The labels get all mixed up when the root is at the top. What do I do?

Rotate labels, using the Settings/Rotate labels menu. This only has an effect when the root is drawn at the top.

# 6 Command line options

Starting DensiTree from the command line or from a script allows one to customize default settings. DensiTree is started using

```
java [vmoptions] -jar DensiTree [options] [file]
```

The following options can be used:

```
 -c <float nr>
```

intensity used for drawing consensus trees, default 1

```
-i <float nr>
```

intensity used for drawing trees, default 1

```
-j <int nr>
```

jitter used for drawing x-position of nodes, default 0

```
-w <int nr>
```

width of consensus tree lines, default 4

```
-v <int nr>
```

width of tree lines, default 1

```
-f <int nr>
```

delay between drawing two frames in animation in milliseconds, default 100

```
-t <int nr>
```

number of threads used for drawing tree set, default 2

```
-b <int nr>
```

burn in, i.e., the number of trees at the start of the tree set that are ignored, default 0. NB this is not the number of samples in the MCMC run that are ignored, but the actual number of trees.

```
-scalemode [none|short|full]
```

set the grid. Should be one of 'none' for no grid (default), 'short' to show a scale or 'full' to show lines throughout image

```
-li <float nr>
```

sets up label indent, default 0

```
-o <file>
```

sets output file. The tree set specified through the command line is drawn and exported in png format to the specified file name. This can be useful for batch image processing.

```
-KML <File name>
```

KML file with geographical locations

```
-geowidth <int nr>
```

widht of the link between taxon and its geographical location

```
-geocolor <color>
```

color of the link between taxon and its geographical location

```
-bg <file>
```

sets the background image

```
-bd <BranchDrawer class>
```

sets the method for drawing branches. Available methods are "viz.graphics.BranchDrawer" for lines (default), "viz.graphics.ArcBranchDrawer" for arcs, and "viz.graphics.SteepArcBranchDrawer" for steep arcs.

```
-pattern <pattern>
```

specifies meta data pattern. Useful for drawing trees with one of the shuffle options using meta data. A regular expression pattern should be specified following the java conventions (see for example `http://download.oracle.com/javase/1.4.2/docs/api/java/util/regex/Pattern.html`) which matches the complete meta data annotation in the tree.

```
-colorpattern <pattern>
```

Colors taxon labels differently where the color is selected randomly. A group needs to be defined in the regular expression. Every taxon matching the same part in this group gets the same color. For example

```
 -colopattern ^(..).*$
```

makes all taxa with the same two starting letters the same color.

```
-rotatetext
```

to rotate labels when the root is drawn at the top.

```
-singlechild
```

allow parsing files with nodes that have single child instead of two children.

If a file is specified, it should be after the last option. DensiTree will attempt to open that file on start up.

The following virtual machine options (`[vmoptions]` above) can be useful:
`-Dcolor.1=0XFF0022` to specify color of most popular topology,
`-Dcolor.2=0XFF0022` for next most popular topology,
`-Dcolor.3=0XFF0022` for third most popular topology,
`-Dcolor.default=0XFF0022` for remainder of topologies,
`-Dcolor.cons=0XFF0022` for consensus trees,
`-Dcolor.label=0XFF0022` for color of labels,
`-Dcolor.height=0XFF0022` for color of height grid and height bar,
`-Dcolor.bg=0XFF0022` for back ground color.

The color is specified as RGB color in Hexadecimal (the number 0XFF0022 listed in all options above).

# 7 Acknowledgements

Joseph Heled's enthusiasm, bug reports and suggestions helped a lot in improving DensiTree in a short period of time. Furthermore, he made tree sets available for improving the gallery in this document. Simon Greenhill kindly made tree sets available for the gallery.

# References

[1] BEAST: Bayesian Evolutionary Analysis Sampling Trees. Alexei J. Drummond and Andrew Rambaut `http://beast.bio.ed.ac.uk/`

[2] FigTree. Andrew Rambaut `http://tree.bio.ed.ac.uk/software/figtree/`

[3] D. H. Huson and D. Bryant, Application of Phylogenetic Networks in Evolutionary Studies, Mol. Biol. Evol., 23(2):254-267, 2006. `http://www.splitstree.org/`

[4] Greenhill, S.J., Blust. R, Gray, R.D. The Austronesian Basic Vocabulary Database: From Bioinformatics to Lexomics. Evolutionary Bioinformatics, 4:271-283. 2008.

[5] Nina Amenta, Jeff Klingner. Case Study: Visualizing Sets of Evolutionary Trees. University of Texas at Austin. 2002 IEEE Symposium on Information Visualization (InfoVis), pp.71, 2002.

[6] Hillis, D.M., T.A. Heath, and K. St. John. Analysis and Visualization of Tree Space. Systematic Biology 54(3):471-82, Jun 2005.

[7] Maddison, D. R. The discovery and importance of multiple islands of most-parsimonious trees. Syst. Zool. 40:315328. 1991

[8] Huelsenbeck, J. P., F. Ronquist, R. Nielsen and J. P. Bollback. 2001. Bayesian inference of phylogeny and its impact on evolutionary biology. Science 294: 2310-2314. `http://mrbayes.csit.fsu.edu/`