Phylo: A Phylogenetic Tree Viewer

The Jackson Laboratory

Aubrey A. Smith, Joel H. Graber, Gary A. Churchill
Center for Genome Dynamics at The Jackson Laboratory, Bar Harbor, Maine 04609 USA

Abstract

Phylogenetic tree viewers are typically used to show relationships and differences between strains or species across the entirety of their genomes. Recent studies into the ancestry of various laboratory and wild-derived mouse strains have produced novel phylogenies, using small spans of the mouse genome instead; the trees produced often have zero distance between certain strains. Previous phylogenetic tree viewers could not handle such trees, as their labels are not visually separated, making making manual editing and labeling necessary for display and analysis of these trees.

We present a new phylogenetic tree viewer designed to overcome these limitations. "*Phylo*" is a small, OpenGL driven tree viewer which can display trees in various forms, including an unrooted tree view with the option to "explode" the labels for clear reading. Rooted views are also available for trees with a defined root.

Phylo will open any New Hampshire (or Newick) format tree file, locally or from a web page. It was designed specifically to work with web browsers, to integrate it with a genome browser or similar tool, joining the trees with the genome data as a whole.

Purpose

Phylo is designed to display phylogenies encoded in the standard Newick (New Hampshire) file format. *Phylo* can display readable trees even when other tree viewers would render labels too close together, or even on top of one another, making the results hard to understand.

The user has control over nearly all aspects of the display in real time, including label distance in unrooted view, zoom and rotation, and in which style the tree should be displayed (including standard cladogram styles and circular view).

Phylo is also designed using a cross-platform API (GTK+), and thus will soon be made available for use on the three major platforms (Windows, Mac OS X, and Linux).

Software Framework

• GTK+, used as the underlying engine
• OpenGL for graphics rendering
• Pango and Cairo for handling fonts (underlying FreeType2 library to Pango)

The methods for reading the Newick (New Hampshire) format tree files are written from base principles, using Joseph Felsenstein's *Inferring Phylogenies* as a guide.

Phylo was specifically designed to allow interaction with and linking from browsers, such as GMOD's GBrowse, giving an interactive visualization for the relationships between strains in a given location on the chromosome.

Figure 1: Example View

A view from a partial genome phylogenies (Yang H, *et al*, *Nature Genetics*, 2007) at 100 Kbp spacing; this is from Chromosome 11, 75.6 Mbp to 75.7 Mbp, build 36. Thin lines are to connect labels to the tree, for ease of reading; thick lines are the tree connections themselves.
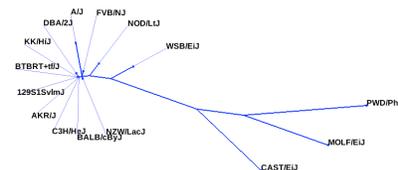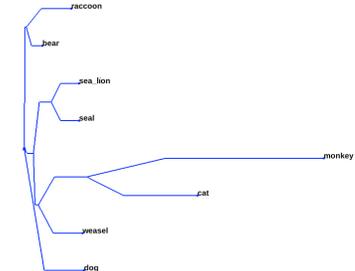

Figure 2 : Example Zoom

The same tree as in Figure 1, zoomed in to get a closer look at the relationships between closely related strains. The view is rotated nearly 180 degrees from the above, and shows zero-distance relationships between sets of three in this window; without spreading the labels, this would be unreadable.
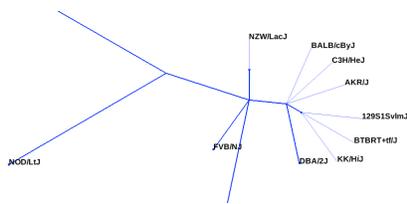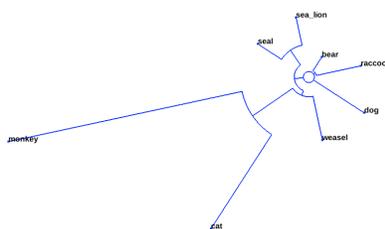

Figure 3 : Circular View

This is a circular rooted view, starting from the center and progressing outwards, where a normal cladogram would progress from one side of the image to the other. The map is drawn using example data extracted from:
http://evolution.genetics.washington.edu/phylip/newicktree.html

Figure 4 : European Cladogram

The same information as in Figure 3, set using a European-style cladogram.

Figure 5 : Newick File Format

The raw Newick (New Hampshire) tree format files, which *Phylo* interprets into the graphs seen in previous figures, are in this style. Spaces are inserted for easier reading.

A) From the sub-specific origin of the laboratory mouse, Chromosome 11 75.6-75.7 Mbp, build 36 (Yang H *et al*, *Nat Gen* 2007):

```
((KK/HlJ:0.00000, BTBRT+tf/J:0.00000, 129S1SvImJ:0.00000):0.00947,
AKR/J:0.00000, C3H/HeJ:0.00000, BALB/cByJ:0.00000, (NZW/LacJ:
0.01694, (((CAST/EiJ:0.57599, (MOLF/EiJ:0.58305, PWD/PhJ:0.80187):
0.30922):0.59197, WSB/EiJ:0.16659):0.14025, NOD/LtJ:0.10164):
0.04865, FVB/NJ:0.03388, A/J:0.22023):0.02098, DBA/2J:0.03388)
[0.0625];
```

B) From washington.edu, example Newick format file, carnivore phylogeny:

```
((raccoon:19.19959, bear:6.80041):0.84600, ((sea_lion:11.99700,
seal:12.00300):7.52973, ((monkey:100.85930, cat:47.14069):
20.59201,  weasel:18.87953):2.09460):3.87382, dog:25.46154);
```

Conclusions

• *Phylo* is a useful tool for rendering phylogenetic trees and interacting with them in real time, soon to be available on most major platforms.

• *Phylo* has the ability to display trees with close (or even zero-distance) relationships, and still have a readable image, a feature that is valuable for many applications, including Hyuna Yang's work on inferring the sub-specific origin of the lab mouse.