

## Phylogenetics

## PhyloWidget: web-based visualizations for the tree of life

Gregory E. Jordan<sup>1,\*</sup> and William H. Piel<sup>2</sup><sup>1</sup>EMBL-EBI, Hinxton, Cambridge, CB10 1SD, UK and <sup>2</sup>Peabody Museum of Natural History, Yale University, P.O. Box 208118, New Haven, CT 06520, USA

Received on March 16, 2008; revised on May 13, 2008; accepted on May 14, 2008

Advanced Access publication May 16, 2008

Associate Editor: Burkhard Rost

## ABSTRACT

**Summary:** PhyloWidget is a web-based tool for the visualization and manipulation of phylogenetic tree data. It can be accessed online or downloaded as a standalone application. A simple URL-based API allows databases to easily link to and customize PhyloWidget for interactively viewing medium- to large-sized trees.

**Availability:** PhyloWidget is available for online use or download at <http://www.phylowidget.org/>. Its source code is released under the GNU General Public License.

**Contact:** [phylowidget@treebase.org](mailto:phylowidget@treebase.org)

## 1 INTRODUCTION

Ever since Charles Darwin published *On the Origin of Species* with a drawing of a ‘tree of life’ as its sole figure (Darwin, 1859), the phylogenetic tree has been a fundamental data structure throughout biology, playing an important role in diverse areas of study.

Reconstructing the Tree of Life is one such field, which requires a combination of biodiverse sequencing projects and automated methods for inferring phylogenetic trees from molecular sequence data. Phylogenetic databases have become an important tool in these efforts, providing centralized repositories for the massive amounts of published phylogenetic data (Morell, 1996; Nakhleh *et al.*, 2003; Sanderson *et al.*, 1993). The trees stored within these databases may then be combined parsimoniously into larger structures by a variety of supertree methods (Sanderson *et al.*, 1998). One such method is the ‘prune and graft’ approach, in which, for example, species-level trees are grafted on to a genus-level tree. Unfortunately, software does not yet exist to assist in the manual cutting and pasting of subtrees into a larger tree structure.

Large-scale phylogenetic analyses are also used in many fields apart from systematics. For example, phylogenetic analyses have been employed by researchers studying the evolution of gene families, (Ruan *et al.*, 2007), organismal coevolution (Huelsenbeck and Rannala, 1997) and the spread of virus strains. For most of these applications, web-accessible databases were used for storing the large number of phylogenetic trees produced by the analysis; thus, phylogenetic databases have proven generally useful beyond the Tree of Life initiative.

A large number of programs have already been written for inferring and viewing phylogenetic trees. Popular programs for phylogenetic inference include Phylip (Felsenstein, 1989), PAUP (Swofford, 2003) and Mesquite (Maddison and Maddison, 2008).

For viewing and editing phylogenies, popular applications include ATV (Zmasek and Eddy, 2001), TreeView (Page, 1996) and Dendroscope (Huson *et al.*, 2007); an extensive listing is maintained by Felsenstein (2008). Unfortunately, most of the available phylogenetic software is ill-suited for online use and integration with phylogenetic databases. The ideal program for online tree visualization would be intuitive to use, extremely configurable and tightly integrated with the web environment. As no such program currently exists, we created a new design from scratch. PhyloWidget was developed as a ‘Phyloinformatics Summer of Code 2007’ project, hosted by NESCent and sponsored by Google.

## 2 FEATURES

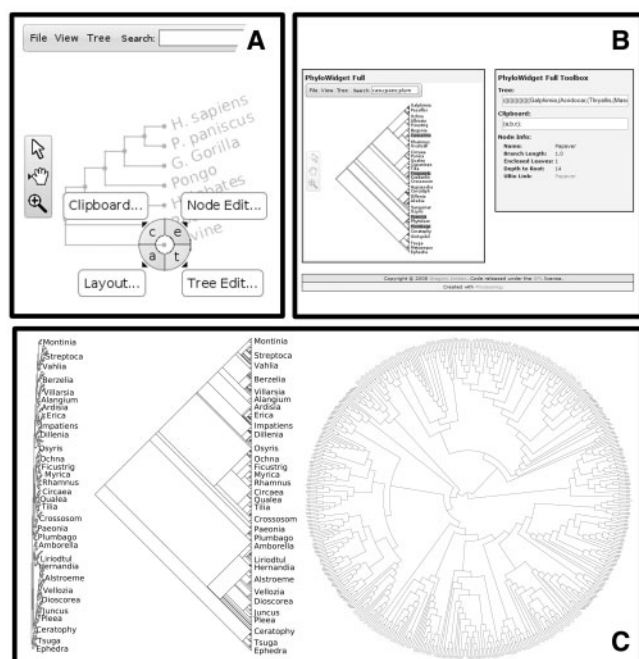
## 2.1 User interface

PhyloWidget is based on a novel user interface for navigating and manipulating phylogenetic trees. A menubar sits at the top of the window and a tool palette floats to the side. Three simple tools, shown in Figure 1, provide access to the complete navigation and manipulation functionality. The pan and zoom tools were designed for quick navigation through phylogenetic trees regardless of size; a single click-and-drag motion can magnify the field of view of a 1000-node tree from the full tree down to a single node. Finally, the arrow tool is used for performing actions on the tree. Clicking on any node brings up a contextual menu of available actions, such as rerooting the tree, editing the node label or branch length, or copying and pasting the subtree enclosed by the selected node. PhyloWidget also provides dynamic web links, which let the user quickly visit a database or web site of relevance to the selected node. In the web interface, two HTML text fields are continuously updated with the Newick representation of the tree and clipboard contents, allowing for easy cutting and pasting from and to external programs. Keyboard shortcuts are provided for the toolbar and context menu, allowing advanced users to quickly perform complex tasks.

## 2.2 Rendering engine

The tree rendering engine contains some notable features. Primary among them is its speed and responsiveness; on a moderately fast computer, trees with upwards of 3000 leaf nodes can be displayed at 30 frames per second. Such speed allows for interactive navigation through the tree and is achieved by rendering only a skeletonized version of the tree if the drawing step becomes too time intensive. Furthermore, the rendering algorithm can optionally ensure that node labels are drawn with a ‘guaranteed minimum text size’.

\*To whom correspondence should be addressed.



**Fig. 1.** (A) PhyloWidget's user interface includes a toolbar (top), tool palette (left), and radial context menu (bottom). Users click on any node in the tree to activate the context menu. (B) The PhyloWidget web applet sits embedded within a browser window, and two text fields contain the current Newick representations of the tree contents and clipboard. Also shown is the search functionality, whereby leaves that match the search query are highlighted. (C) A 1000-node tree is shown using PhyloWidget's rectangular, diagonal and circular renderers.

consistently presenting the user with a readable sampling of leaf node labels regardless of tree size. For example, the first two trees in Figure 1C show a 1000-node tree with a guaranteed minimum text size of 10 pixels. When the renderer has to occlude nodes to obey the minimum text size, display priority is given to the node with a shorter distance to the root node.

## 2.3 Customization

PhyloWidget is highly customizable. Although the default configuration is suitable for many applications, particular databases or applications may customize the program by a variety of means. The toolbar, tool palette and context menu are all defined within a simple XML file, so any undesired tools or actions may be removed. In addition, many parameters may be changed from their default values, either by (a) altering the source code, (b) using Javascript to change settings in real-time or (c) configuring parameters using a simple URL-based API. The third option is ideal for databases which desire a customized view without having to host the applet on their own site.

## 2.4 Input and output

PhyloWidget can parse trees in the Newick, NHX and Nexus formats and can output trees in Newick or NHX format. The tree image can be exported as a JPEG, PNG or PDF file. PDF output

is completely vector and text-based, making it ideal for creating publication-quality figures or for high-resolution printing.

## 3 IMPLEMENTATION

PhyloWidget is written in Java and may either be run offline or online as a browser applet. Although Flash is widely used for web-based graphics and was successfully employed by the Interactive Tree of Life (Letunic and Bork, 2007), we chose Java due to its preferable speed, code portability and acceptance among the bioinformatics community. Furthermore, many useful open-source libraries are written in Java; PhyloWidget makes use of the Processing library for graphical operations and the iText library for PDF output.

PhyloWidget should work on all operating systems with a Java Runtime Environment version 5 or newer. Some OS and browser combinations are known to cause problems with the web interface; these issues are noted in a dedicated troubleshooting page. The website also provides extensive documentation demonstrating the use and deployment of PhyloWidget.

## ACKNOWLEDGEMENTS

We thank Hilmar Lapp for his support, advice and encouragement, Perry Miller and Cynthia Brandt for their comments on the article, and three anonymous reviewers for their helpful comments.

**Funding:** This work was supported by Google's Summer of Code 2007, the National Evolutionary Synthesis Center (NESCent), and NSF grants IIS-0629702 and EF-0331654 awarded to W.P. G.J. is a member of Darwin College, University of Cambridge.

**Conflict of interest:** none declared.

## REFERENCES

- Darwin, C. (1859) *On the Origin of Species By Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. John Murray, London.
- Felsenstein, J. (1989) Phylip – phylogeny inference package (version 3.2). *Cladistics*, **5**, 164–166.
- Felsenstein, J. (2008) Phylogeny programs. Available at <http://evolution.genetics.washington.edu/phylip/software.html>. (last accessed date May 12, 2008).
- Huelsenbeck, J.P. and Rannala, B. (1997) Phylogenetic methods come of age: testing hypotheses in an evolutionary context. *Science*, **276**, 227–232.
- Huson, D.H. et al. (2007) Dendroscope: an interactive viewer for large phylogenetic trees. *BMC Bioinformatics*, **8**, doi: 10.1186/1471-2105-8-460.
- Letunic, I. and Bork, P. (2007) Interactive tree of life (itol): an online tool for phylogenetic tree display and annotation. *Bioinformatics*, **23**, 127–128.
- Maddison, W.P. and Maddison, D.R. (2008) Mesquite: a modular system for evolutionary analysis (version 2.5). Available at <http://mesquiteproject.org> (last accessed date May 12, 2008).
- Morell, V. (1996) Treebase: the roots of phylogeny. *Science*, **273**, 569.
- Nakhleh, L. et al. (2003) Requirements of phylogenetic databases. In *Proceedings of the Third IEEE Symposium on Bioinformatics and Bioengineering*. IEEE Computer Society, Washington DC, USA, pp. 141–148.
- Page, R.D. (1996) Treeview: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.*, **12**, 357–358.
- Ruan, J. et al. (2007) Treefam: 2008 update. *Nucleic Acids Res.*, **36**, (Database Issue) D735–D740.
- Sanderson, M.J. et al. (1993) The growth of phylogenetic information and the need for a phylogenetic data base. *Syst. Biol.*, **42**, 562–568.
- Sanderson, M.J. et al. (1998) Phylogenetic supertrees: assembling the trees of life. *Trends in Ecol. & Evol.*, **13**, 105–109.
- Swofford, D.L. (2003) Paup\*: phylogenetic analysis using parsimony, version 4.0b10.
- Zmasek, C.M. and Eddy, S.R. (2001) Atv: display and manipulation of annotated phylogenetic trees. *Bioinformatics*, **17**, 383–384.