

Phylogenetics

Paloverde: an OpenGL 3D phylogeny browser

Michael J. Sanderson

Section of Evolution and Ecology, University of California, Davis, CA, USA

Received on December 11, 2005; revised on January 26, 2006; accepted on February 3, 2006

Advance Access publication February 24, 2006

Associate Editor: Keith A Crandall

ABSTRACT

Summary: *Paloverde* is a new program designed to help visualize the phylogenetic structure of moderately large trees—trees on the scale of 100–2500 leaf nodes. The program embeds the user in an interactive virtual 3D world in which a large tree presented in various layouts can be manipulated through a mouse interface. The program implements radial 2D layouts, and true 3D spiral, conical and hemispherical (i.e. truly ‘tree’-like) layouts. Subclades can be defined in the input file (using standard node-based definitions) and displayed collapsed as new leaf nodes, or left intact but annotated with names around the periphery of the tree. A search tool lets the user zoom to any selected leaf node. *Paloverde* is an open source project written in ANSI C using the OpenGL library for 3D visualization.

Availability: Source code, makefiles for Mac OS X and Linux and a compiled binary for Mac OS X are available at <http://ginger.ucdavis.edu/paloverde/paloverde.html>, along with a sample dataset.

Contact: mjsanderson@ucdavis.edu

1 INTRODUCTION

The increasing size of phylogenetic trees (e.g. Hibbett *et al.*, 2005) and recent efforts to assemble the ‘tree of life’ (<http://www.phylo.org/beta/AToL/index.html>) have stimulated interest in new tree visualization tools that can scale well and convey useful information about relationships (Munzner *et al.*, 2003). Widely used 2D ‘flat screen’ tools such as those provided in PAUP* (Swofford, 2002), MacClade (Maddison and Maddison, 2000), TreeView (Page, 1996), ATV (Zmasek and Eddy, 2001), TreeIllustrator (Trooskens *et al.*, 2005) and others become cumbersome when trees have more than ~100 leaf nodes. Collapsing subtrees via hyperlinking can rescue this layout to some extent (e.g. the very large tree encapsulated in the Tree of Life Web project: <http://tolweb.org/tree/phylogeny.html>), as can more elegant distortion of the 2D space to compress parts of the tree, as in TreeJuxtaposer (Munzner *et al.*, 2003). Transformation of tree structures to non-Euclidean ‘hyperbolic’ space can permit very large trees to be displayed in two dimensions by interactively magnifying ‘nearby’ parts of a tree and condensing more distant areas (Munzner, 1998; Bingham and Sudarsanam, 2000), providing ‘focus plus context’. These hyperbolic trees can also be laid out in a 3D space, allowing even larger trees to be displayed (Hughes *et al.*, 2004), such as NCBI’s manually constructed taxonomy tree for all GenBank accessions, currently containing over 150 000 leaves (<http://www.ncbi.nlm.nih.gov/Taxonomy/taxonomyhome.html>).

Although these tools show great promise for scalability, there is a need for tools tailored to the specific goals of phylogeneticists, which include rapid navigation among clades of trees that are

often nearly binary, without losing sight of leaf names or names of clades defined by the user. Moreover, although the visual metaphor of hyperbolic space is intuitive, and other distortion-based displays are promising, it may not be necessary to forego the even more intuitive metaphor of Euclidean space for trees that are somewhat smaller than the very large trees that are the targets of these programs. Features available in mature software graphics libraries like OpenGL, such as realistic lighting, perspective and other attributes that add realism have rarely been applied to phylogeny visualization tools to improve their utility. *Paloverde* is a new OpenGL-based program aimed at visualizing trees that are moderately large (100–2500 terminal taxa).

2 FEATURES**2.1 Tree layouts and annotations**

Paloverde takes as input a ‘Nexus’-formatted text file that includes a description of the tree, possibly with user-defined branch lengths. Such files can be created by most widely used phylogenetic inference tools. *Paloverde* constructs trees as 2D and 3D solid objects, but in all cases embeds the tree in a 3D space, keeping the metaphor of Euclidean geometry intact. This virtual world has natural lighting to add perspective and ‘realism’. The tree itself can be modified by several tunable parameters related to its layout and appearance (e.g. branch thickness, size of taxon labels). A mouse interface lets the user move and rotate the tree and zoom rapidly to parts of it. An animation-driven search tool lets the user zoom directly to any leaf node. Several layouts are implemented: 2D circular or partial circle; 3D spiral; 3D cone and 3D hemisphere. The latter is closest in spirit to a true ‘botanical’ tree but is not always the most informative presentation.

Several features of the program are aimed at displaying user-defined or taxonomically defined information about subtrees. Users can easily define names for clades using so-called node-based definitions in the input file. A node-based definition requires listing at least two leaf labels whose most recent common ancestor defines the clade of interest. This convention corresponds closely with one emerging view of a standard for phylogenetic taxon definitions (<http://www.phylocode.org>). The user can then optionally collapse these clades to single leaf nodes that are presented in separate windows, or annotate the original complete tree with the clade names in a visually appealing fashion. A final option, useful when leaf names obey Linnean taxonomic naming conventions, such as the binomial given by ‘genus_species’, will collapse all monophyletic collections of leaves that have the same ‘genus’ (Fig. 1). When a genus is not monophyletic, the program displays all subclades of those genera that are.

et al., 2003). In the middle range, there is more room for esthetic choice and a focus on features tailored to the specific information needed by the user. *Paloverde* emphasizes options for displaying labels attached to leaf nodes and those associated with user- or taxonomy-defined clades, while embedding the information in an intuitive virtual world. Even its essentially 2D circular layout, when imbedded in a 3D world, offers visual cues that aid in navigation and display that is unavailable in strictly 2D versions of the same layout.

ACKNOWLEDGEMENTS

Paloverde is part of the Phylota project (<http://www.phylota.org>), supported by the NSF AToL program. Thanks to Gordon Burleigh, Oliver Eulenstein, David Fernandez-Baca, Amy Driskell, Taum Hanlon, Junhyong Kim, Shelley McMahon, Brian O'Meara, Cam Webb and Marty Wojciechowski for feedback on the program design.

Conflict of Interest: none declared.

REFERENCES

- Bingham, J. and Sundarsanam, S. (2000) Visualizing large hierarchical clusters in hyperbolic space. *Bioinformatics*, **16**, 660–661.
- Hibbett, D.S. *et al.* (2005) Automated phylogenetic taxonomy: an example in the Homobasidiomycetes (mushroom-forming fungi). *Syst. Biol.*, **54**, 660–668.
- Hughes, T. *et al.* (2004) Visualizing very large phylogenetic trees in three-dimensional hyperbolic space. *BMC Bioinformatics*, **5**, 1–6.
- Maddison, W.P. and Maddison, D.R. (2000) *MacClade 4: Analysis of Phylogeny and Character Evolution*. Sinauer, Sunderland, MA.
- Munzner, T. (1998) Exploring large graphs in 3D hyperbolic space. *IEEE Comput. Graphics Appl.*, **18**, 18–23.
- Munzner, T. *et al.* (2003) TreeJuxtaposer: scalable tree comparison using focus+context with guaranteed visibility. *ACM Trans. Graphics*, **22**, 453–462.
- Page, R.D.M. (1996) TreeView: an application to display phylogenetic trees on personal computers. *Comp. Appl. Biosci.*, **12**, 357–358.
- Swofford, D.L. (2002) *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Sinauer, Sunderland, MA.
- Trooskens, G. *et al.* (2005) Phylogenetic trees: visualizing, customizing and detecting incongruence. *Bioinformatics*, **21**, 3801–3802.
- Zmasek, C.M. and Eddy, S.R. (2001) ATV: display and manipulation of annotated phylogenetic trees. *Bioinformatics*, **17**, 383–384.