

Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy

Ivica Letunic and Peer Bork*

EMBL, Meyerhofstrasse 1, 69117 Heidelberg, Germany

Received January 28, 2011; Revised February 23, 2011; Accepted March 21, 2011

ABSTRACT

Interactive Tree Of Life (<http://itol.embl.de>) is a web-based tool for the display, manipulation and annotation of phylogenetic trees. It is freely available and open to everyone. In addition to classical tree viewer functions, iTOL offers many novel ways of annotating trees with various additional data. Current version introduces numerous new features and greatly expands the number of supported data set types. Trees can be interactively manipulated and edited. A free personal account system is available, providing management and sharing of trees in user defined workspaces and projects. Export to various bitmap and vector graphics formats is supported. Batch access interface is available for programmatic access or inclusion of interactive trees into other web services.

INTRODUCTION

With the exponential increase of sequence data being generated by various classical and next-generation sequencing studies (1), phylogenetics and phylogenetic trees are becoming an integral part of various biological studies. Through the years, many tools have been developed, covering all possible aspects of classical tree visualization. Tools like TREEVIEW (2), Archaeopteryx (3), TreeDyn (4), Dendroscope (5), PhyloWidget (6) and others provide many useful features and are used regularly by scientists worldwide. However, an important aspect missing from most visualization tools is the ability to annotate phylogenetic trees with various external data sets. We have, therefore, developed an in-house tool that included these functions, and made it public after significant interest from the small evaluating community (7). Here, we report on the current state of the tool and recent developments which have greatly expanded iTOL's functionality, tested by and implemented through the suggestions of several thousands active users. A complete overview of features added since the original publication is available

in iTOL's version history page (<http://itol.embl.de/help/history.shtml>).

FEATURES

iTOL is an online tool, accessible with any modern web browser (Figure 1). Trees are displayed using a combination of Adobe Flash and Javascript. The Flash object which displays the tree and provides interactivity is created dynamically by the server using a set of Perl scripts and the Ming library (<http://ming.sourceforge.net>).

INPUT, INTERFACE AND BASIC FUNCTIONS

iTOL works with two most commonly used phylogenetic tree formats, Newick and Nexus (8). Recently proposed new tree format, phyloXML (9) is also supported. Various additional data used for tree annotation is uploaded through plain text files.

iTOL provides most common functions available in any phylogenetic tree viewer. Trees can be displayed in various formats: as phylograms or cladograms, rooted or unrooted and rectangular or radial. Several tree customization options are available, giving users a great amount of control over the tree display. Tree clades can be pruned or collapsed, either manually or automatically, based on various parameters (such as associated bootstrap values or average branch length distance). Trees can be re-rooted manually on any node, or automatically using the mid-point rooting method. Basic editing functions allow users to interactively delete or move single nodes or whole clades.

EXPORT

One of iTOL's major functions is the creation of high-quality figures for publication or inclusion into other documents. It is therefore possible to export any tree with its annotations into several graphical formats, both vector and bitmap. Currently supported formats are Scalable Vector Graphics (svg), Portable Network Graphics (png),

*To whom correspondence should be addressed. Tel: +49 6221 387 8526; Fax: +49 6221 387 8517; Email: bork@embl.de

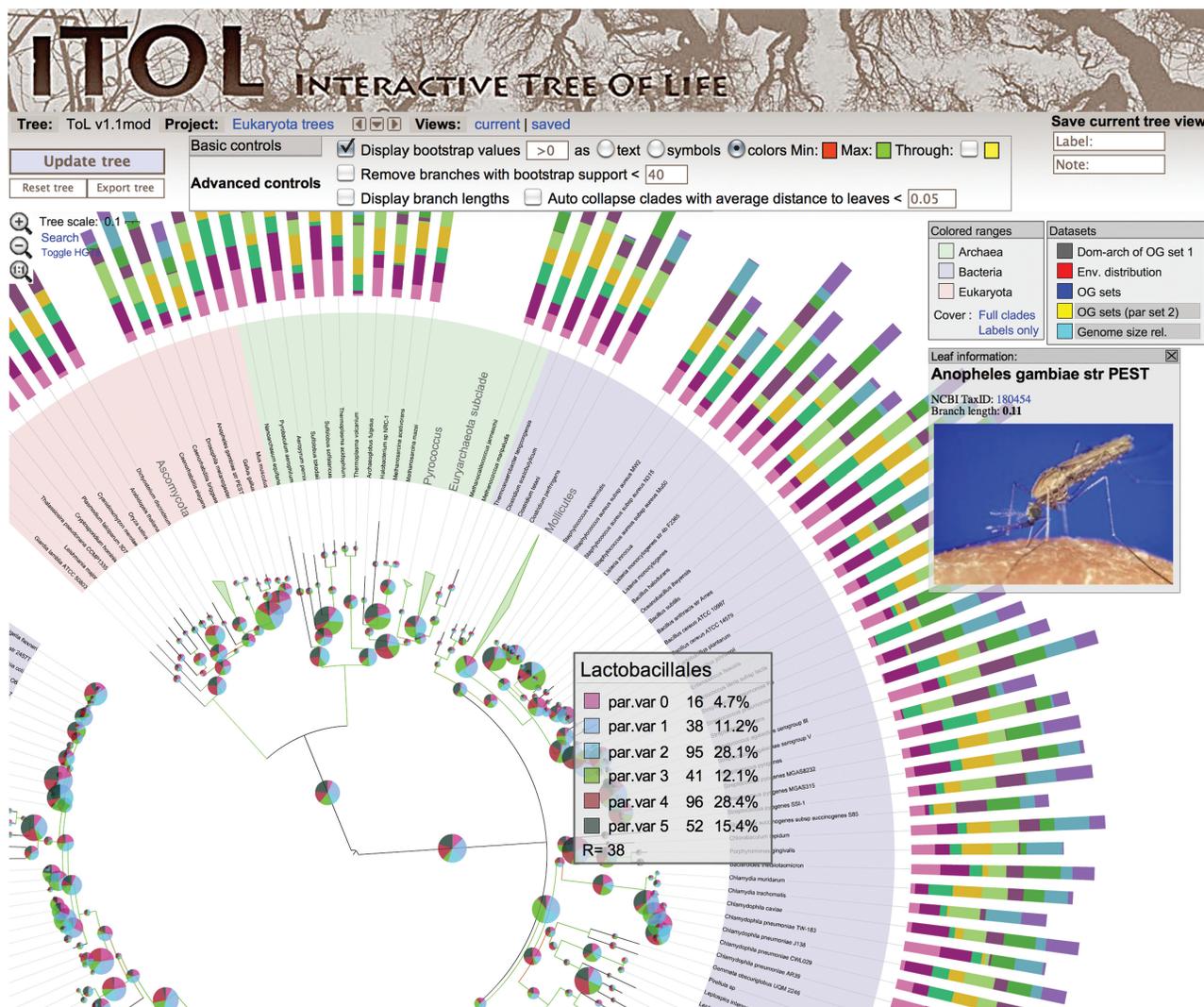


Figure 1. iTOL's user interface. Tree of life (12) annotated with various data sets is shown. Several branches are collapsed and displayed as triangles. Tree branches are colored according to the clade bootstrap value. Popup windows provide detailed information on clades/leaves, such as bootstrap values, taxonomical categories or associated data set values. Alternatively, popup windows can display user defined, HTML formatted text, including external links and images.

Encapsulated Postscript (eps), Postscript (ps) and Portable Document Format (pdf). In addition, pruned trees can be exported in all supported plain text formats (Newick, Nexus or phyloXML).

ANNOTATING TREES

Our primary goal when developing iTOL was to offer numerous ways of annotating phylogenetic trees with external data. Currently, there are two major groups of tree annotations in iTOL:

- (1) automatic annotation based on intrinsic node features (bootstrap values, branch lengths) and
- (2) user provided annotation associated with different nodes (data sets and other information)

DATA SETS

At the moment, iTOL supports 11 different data set types (Figure 2). All data sets are uploaded as plain text files. Various data set parameters can be customized during upload. Due to Flash performance constraints, each phylogenetic tree in iTOL can have a maximum of 10 data sets associated. An overview of supported data set types is shown in Figure 2. Detailed description and information on required formats for each data set is available in iTOL's online help pages.

NODE INFORMATION

When navigating interactive tree displays in iTOL, users can hover mouse cursor over any node to display popup windows with various bits of information associated with the node.

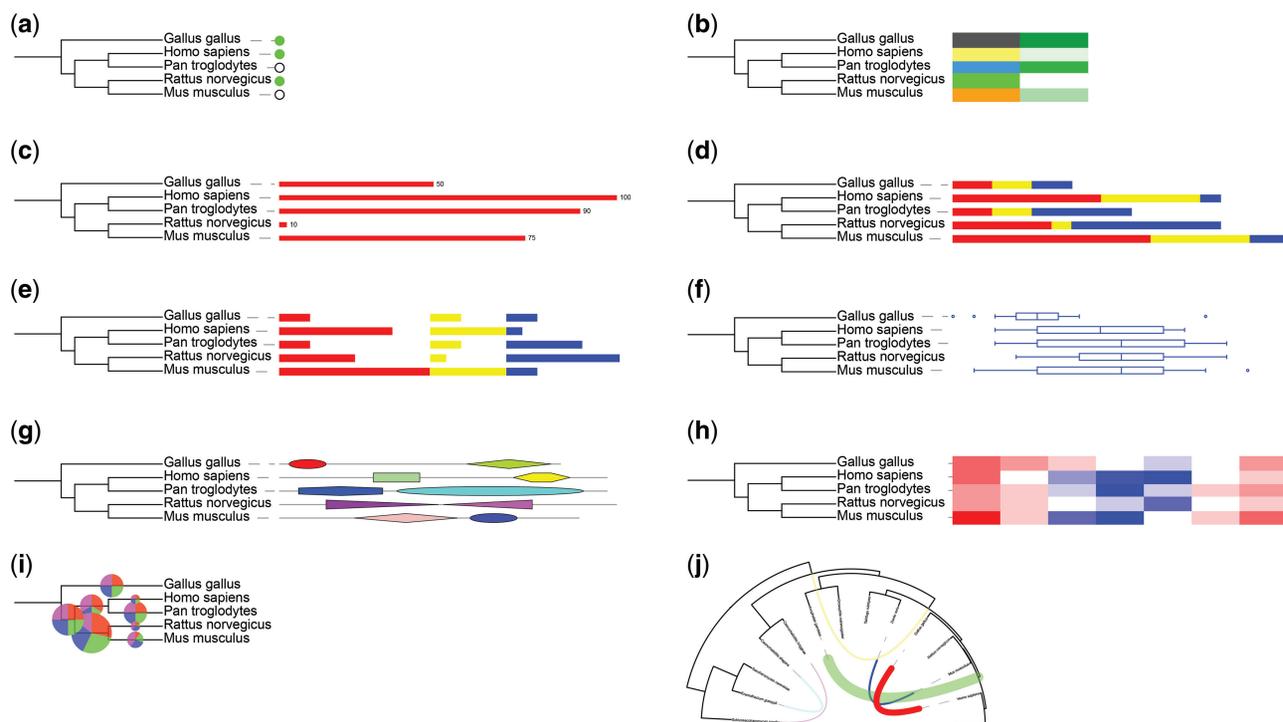


Figure 2. Data set types supported by iTOL. Each tree in iTOL can be annotated with up to 10 different data sets. All data sets are uploaded as numeric and/or color values associated with tree nodes. (a) Binary: binary data associated with leaves. (b) Color strip and gradient: color gradients are automatically calculated from numeric data, while color strips allow direct definition of colors. (c) Bar chart and time series: set of single values associated with leaf nodes displayed as a bar chart; multiple single bar charts can be combined into an animated time series data set. (d) Stacked bar chart: a stacked set of bar charts. (e) Aligned multi value bar chart: multiple aligned bar charts. (f) Boxplot: boxplots can be uploaded as pre-calculated five-number summaries or as raw data. (g) Protein domain architecture: numerous protein domain architectures can be easily defined using 13 different geometric shapes. (h) Heatmap: colors of the map are calculated from numerical data and user defined colors gradients. (i) Pie chart: different-sized pie charts can be displayed on any node in the tree. (j) Connections: any two nodes in the tree can be connected by lines of user defined colors, widths and transparency.

By default, these will include bootstrap values, branch lengths and taxonomic data (if present). However, iTOL allows complete customization of the information shown in the popups, which can contain any user-defined HTML formatted text, including URL links and external images.

DISPLAYING BOOTSTRAP VALUES

Clade support values (bootstraps) offer important clues regarding the quality of the tree and allow easy identification of problematic areas. iTOL can display bootstrap values in three different ways:

- (1) direct display of numeric values above the branches in the tree;
- (2) display of symbols whose size is proportional to the bootstrap values; and
- (3) coloring of tree branches using user specified color gradients, with actual colors calculated from the associated bootstrap value.

COLORING THE TREE

Colors used for the display of trees in iTOL are fully customizable. Both branch and text color of each node can be

independently defined. In addition, users can specify background colors for various clades. Colors can be specified through an interactive color editor, or defined in plain text files uploaded together with the tree.

DISPLAYING HORIZONTAL GENE TRANSFERS

iTOL is the first visualization tool which supports the display of horizontal gene transfers (HGTs) in phylogenetic trees. These can be annotated directly in the original Newick tree file, using an extended format described in iTOL's online help pages. Alternatively, HGTs can also be defined and uploaded in separate plain text files. HGTs are displayed as arrows with user defined colors and labels.

USER ACCOUNTS

iTOL provides a freely available personal account system. It allows registered users to access their uploaded trees from anywhere, organize them into workspaces and projects and easily manage data sets and other tree features. Trees are kept private and stored locally in iTOL.

TREE MANAGEMENT

Trees in an iTOL account can be organized into unlimited number of user-defined projects. Projects can be further separated into various workspaces allowing easy navigation even in accounts with a large number of trees.

DATA SHARING AND COLLABORATION

Any user project in iTOL can be set as 'shared'. Other users can simply access shared projects, allowing them to display, export or copy the trees contained within. Through shared projects, users can simply keep their colleagues up to date with various tree versions and other developments.

BATCH ACCESS

When working with hundreds or even thousands of phylogenetic trees, automation capabilities become an important factor in any tool's usefulness. iTOL offers programmatic access to both its tree upload and export interfaces. All iTOL features available interactively through the web interface can also be accessed using the batch access. Example upload/download scripts are provided in several programming languages. Various tools and servers [examples include SmashCommunity (10) and metaTIGER (11)] already use iTOL as the backend engine for phylogenetic tree visualization.

FUNDING

Funding for open access charge: European Molecular Biology Laboratory

Conflict of interest statement. None declared.

REFERENCES

1. Mardis, E.R. (2011) A decade's perspective on DNA sequencing technology. *Nature*, **470**, 198–203.
2. Page, R.D. (1996) TreeView: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.*, **12**, 357–358.
3. Zmasek, C.M. and Eddy, S.R. (2001) ATV: display and manipulation of annotated phylogenetic trees. *Bioinformatics*, **17**, 383–384.
4. Chevenet, F., Brun, C., Banuls, A.L., Jacq, B. and Christen, R. (2006) TreeDyn: towards dynamic graphics and annotations for analyses of trees. *BMC Bioinformatics*, **7**, 439.
5. Huson, D.H., Richter, D.C., Rausch, C., DeZulian, T., Franz, M. and Rupp, R. (2007) Dendroscope: An interactive viewer for large phylogenetic trees. *BMC Bioinformatics*, **8**, 460.
6. Jordan, G.E. and Piel, W.H. (2008) PhyloWidget: web-based visualizations for the tree of life. *Bioinformatics*, **24**, 1641–1642.
7. Letunic, I. and Bork, P. (2007) Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics*, **23**, 127–128.
8. Wilgenbusch, J.C. and Swofford, D. (2003) Inferring evolutionary trees with PAUP*. *Curr. Protoc. Bioinformatics*, Chapter 6, Unit 6.4.
9. Han, M.V. and Zmasek, C.M. (2009) phyloXML: XML for evolutionary biology and comparative genomics. *BMC Bioinformatics*, **10**, 356.
10. Arumugam, M., Harrington, E.D., Foerster, K.U., Raes, J. and Bork, P. (2010) SmashCommunity: a metagenomic annotation and analysis tool. *Bioinformatics*, **26**, 2977–2978.
11. Whitaker, J.W., Letunic, I., McConkey, G.A. and Westhead, D.R. (2009) metaTIGER: a metabolic evolution resource. *Nucleic Acids Res.*, **37**, D531–D538.
12. Ciccarelli, F.D., Doerks, T., von Mering, C., Creevey, C.J., Snel, B. and Bork, P. (2006) Toward automatic reconstruction of a highly resolved tree of life. *Science*, **311**, 1283–1287.