

Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation

Ivica Letunic^{1,*} and Peer Bork^{2,3,4}

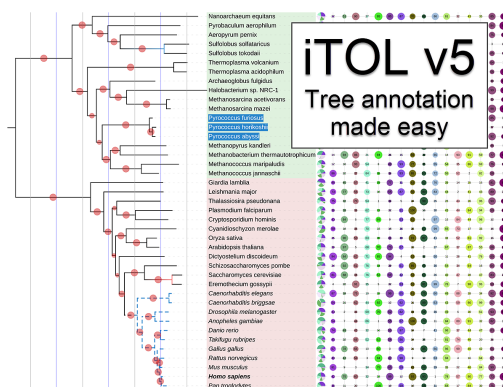
¹Biobyte Solutions GmbH, Bothestr 142, 69126 Heidelberg, Germany, ²EMBL, Meyerhofstrasse 1, 69117 Heidelberg, Germany, ³Department of Bioinformatics, Biocenter, University of Würzburg, 97074 Würzburg, Germany and ⁴Yonsei Frontier Lab (YFL), Yonsei University, Seoul 03722, South Korea

Received February 15, 2021; Revised March 29, 2021; Editorial Decision April 12, 2021; Accepted April 12, 2021

ABSTRACT

The Interactive Tree Of Life (<https://itol.embl.de>) is an online tool for the display, manipulation and annotation of phylogenetic and other trees. It is freely available and open to everyone. iTOL version 5 introduces a completely new tree display engine, together with numerous new features. For example, a new dataset type has been added (MEME motifs), while annotation options have been expanded for several existing ones. Node metadata display options have been extended and now also support non-numerical categorical values, as well as multiple values per node. Direct manual annotation is now available, providing a set of basic drawing and labeling tools, allowing users to draw shapes, labels and other features by hand directly onto the trees. Support for tree and dataset scales has been extended, providing fine control over line and label styles. Unrooted tree displays can now use the equal-daylight algorithm, proving a much greater display clarity. The user account system has been streamlined and expanded with new navigation options and currently handles >1 million trees from >70 000 individual users.

GRAPHICAL ABSTRACT



INTRODUCTION

Phylogenetics and phylogenetic trees are heavily used in a wide variety of biological and other scientific studies, and classical tree visualization is supported by many software tools (1,2), including Interactive Tree Of Life (iTOL) (3), which introduced the annotation of trees with various types of additional data. Nowadays, various software packages offer additional tree annotation features, both online and as stand-alone packages or libraries, for example ETE toolkit (4), ggtree (5), Dendroscope (6), PhyD3 (7) or Evolvew (8). Here, we report on the current state of iTOL as well as recent developments, which have expanded and streamlined its functionality.

TREE DISPLAY

iTOL is an online tool, accessible with any modern web browser (Figure 1). The tree display engine is implemented in pure JavaScript and uses the HTML5 Canvas element for visualization. Most of the display calculations and functionality are performed by the client web browser, allowing fine-grained interactive control over various display parameters. In iTOL v5, the display engine was completely rewritten and optimized with support for high-resolution displays. The low-level drawing functionality is now implemented through the paper.js library (<https://paperjs.org>).

Input types and basic functions

iTOL supports commonly used phylogenetic tree formats such as Newick, Nexus (9) and phyloXML (10). Phylogenetic placement files created by EPA (11) and pplacer (12), as well as QIIME 2 trees and annotation files (13), are also supported.

All additional data used for various types of tree annotation are provided in plain text files, and simply dragged and dropped onto the trees visualized in the user's web browser.

iTOL provides the most common functions available in any phylogenetic tree viewer. Various tree display formats are supported, such as phylograms or cladograms, rooted or unrooted, rectangular or radial. iTOL v5 introduces the

*To whom correspondence should be addressed. Tel: +49 6221 3878534; Email: letunic@biobyte.de

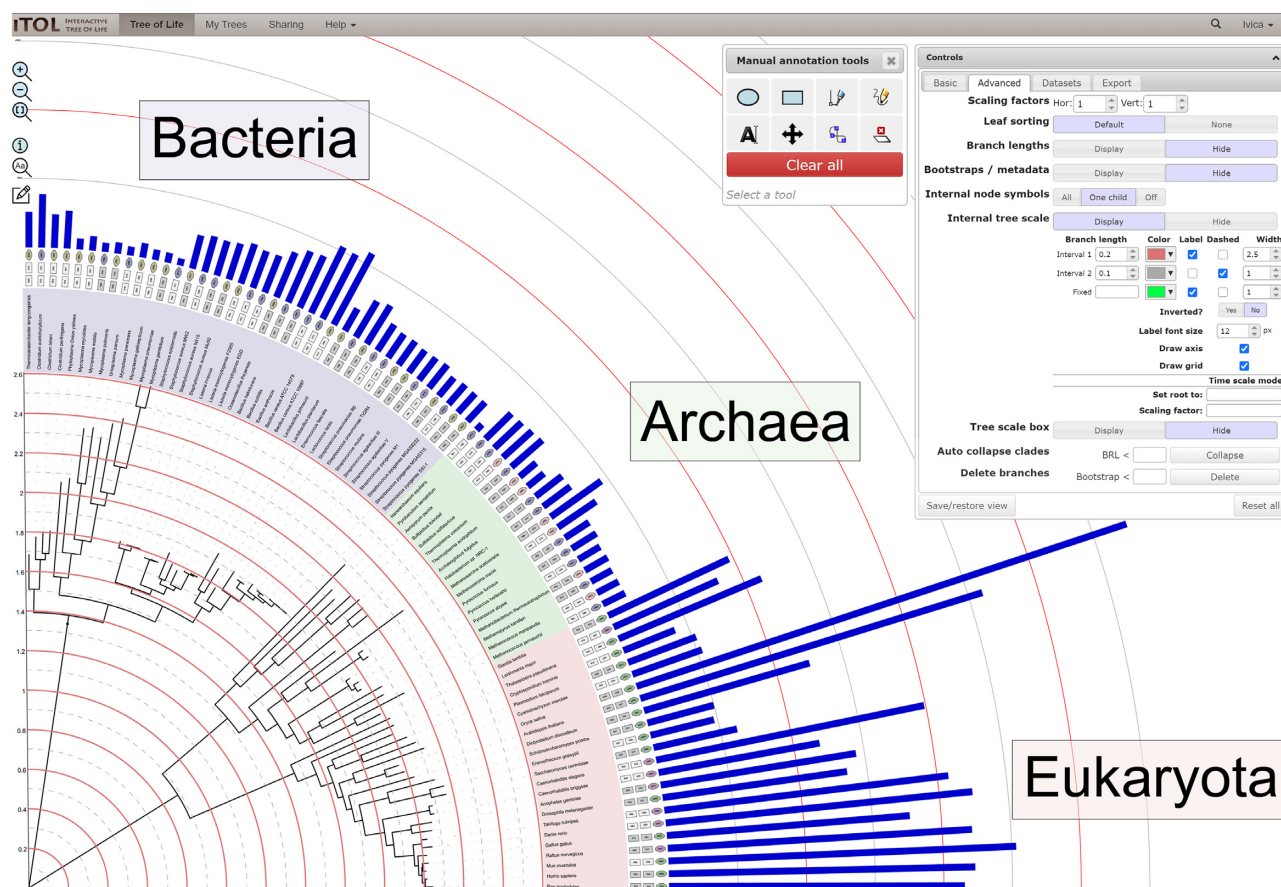


Figure 1. iTOL's user interface. A phylogenetic tree annotated with several datasets is displayed, highlighting several new features. Internal tree scale level lines were individually styled, and an axis was added. Manual annotation drawing tools were used to add the taxonomic domain labels outside the tree. Manual annotations can be interactively scaled, rotated and repositioned.

'equal-daylight' algorithm for the unrooted display mode, which tries to equalize the sizes of angular gaps between the tree's clades. The resulting display shows a much better separation of nearby tree nodes and increases the overall clarity of the tree structure. On the other hand, it can cause internal subtree overlaps, as well as significant overlaps of leaf text labels. Users can select how many iterations of the equal-daylight algorithm to apply, or switch it off completely, and use the default 'equal-angle' unrooted tree display algorithm.

iTOL can manipulate the trees in various ways, and basic editing functions allow users to interactively delete or move single nodes or whole clades. Clades can also be pruned or collapsed, either manually or automatically, based on various parameters (such as associated bootstrap values or average branch length distances). iTOL v5 newly supports tree pruning through a simple text file with a list of tree nodes to include. Users can drop such files onto the display to automatically create a pruned tree representation.

Trees can be re-rooted manually on any node or automatically using the midpoint rooting method. Tree leaves can be sorted in various ways, either manually or automatically.

A complete overview of changes and functions added since the last publication (14) is listed on the iTOL's version history page (https://itol.embl.de/version_history.cgi).

TREE ANNOTATION

iTOL v5 offers various new annotation features, extended functionality in the visualization of existing dataset types, and a new dataset type, MEME motifs (Figure 2). MEME datasets are used to visualize sequence motifs detected by the MEME suite (15). The datasets are created automatically by simply dropping the XML result files generated by the MEME suite onto the phylogenetic tree display in iTOL. The visualization style is similar to the default output generated by the MEME suite and uses the same colors and layout. Individual motifs can be displayed as MEME sequence logos, showing the individual residue probabilities as differently sized letters (Figure 2B).

Displaying branch length, support and metadata values

iTOL v5 expands the node metadata visualization options with several new features. MRBAYES (16) and the New Hampshire X (NHX) formatted metadata support has been extended to support non-numerical values, which can be visualized as branch text labels or used to color the tree branches with user-defined category colors.

Bootstrap support and other numerical metadata values can now be displayed as percentages and scaled by a user-defined factor. In addition to the precise horizontal posi-

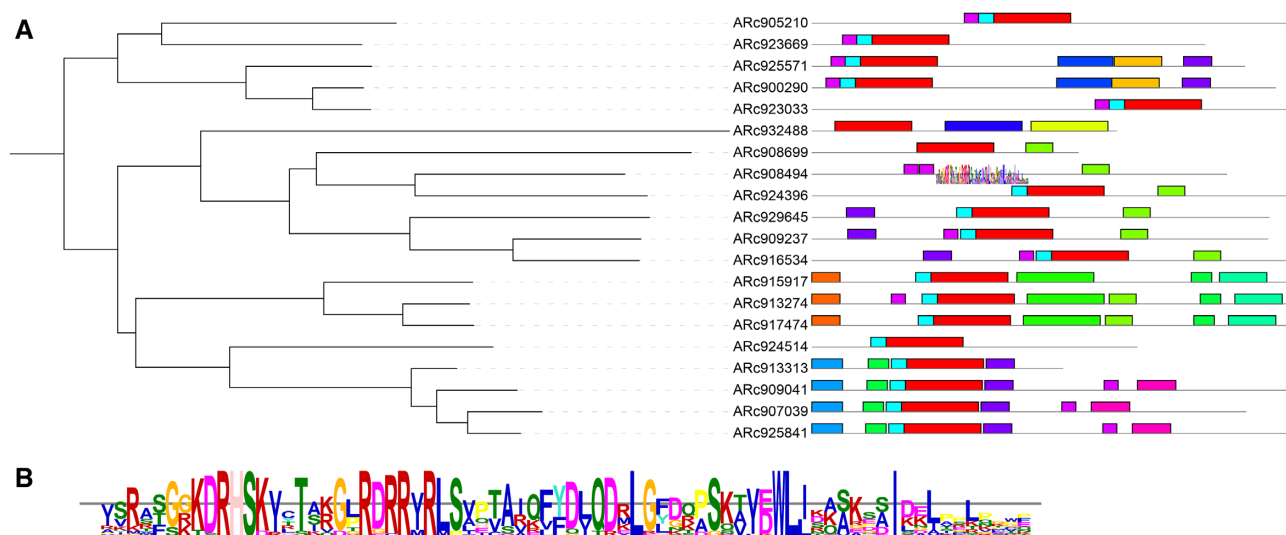


Figure 2. MEME motif dataset visualization in iTOL. (A) Dataset display: an XML file created by the MEME suite has been dropped onto the tree and visualized with a style similar to the default MEME results page. (B) MEME motif logo: zoomed in detail of the dataset, with two motifs displayed as MEME sequence logos, showing the individual residue probabilities as different sized letters.

tion along the branch, metadata text labels can now also be shifted vertically by a user-defined amount.

If multiple metadata values are available in the tree, these can all be displayed simultaneously as slash delimited text labels on the tree branches.

Datasets and the interactive editor

iTOL currently supports the visualization of 19 different dataset types. Most of these datasets can be uploaded as plain text files, which use simple predefined templates that are available through iTOL's help pages. The support for the external data types has been extended and now includes QI-ME FeatureData[AlignedSequence] QZA files, which can be dropped onto the trees to automatically create multiple sequence alignment visualizations.

In addition to the template files, the most commonly used dataset types can be generated and edited directly in the interactive editor, within the iTOL web user interface. In iTOL v5, the editor has been streamlined and expanded with support for new dataset types. Raw data, as well as dataset legends, scales and other information, can be edited directly through a spreadsheet-like interface, with dynamically updating tree visualization.

Dataset support in the unrooted tree display mode has been extended, and now includes the text label datasets, as well as the newly introduced MEME motifs. Even though the individual dataset entries cannot be aligned in this mode, they still provide users with valuable annotation options.

Tree and dataset scales and legends

Each dataset in iTOL can have an associated legend displayed next to the tree. In iTOL v5, all the legends are now draggable and can be manually positioned anywhere in the display, or their position can be predefined in the user-prepared dataset template files. Individual symbols in the

legends can now be scaled. In addition, an automatically generated legend can be added to the tree when support values/metadata are visualized as symbols or branch colors.

Many dataset types in iTOL (like bar charts or protein domain architectures) can have a user-defined scale. The individual levels in any scale now support custom display labels with variable size, while the scale lines can be dashed and displayed in different widths.

iTOL's internal tree scale display options have also been extended, with support for fixed position levels, in addition to the intervals. Individual scale level lines can be dashed and displayed in different widths. Internal tree scale can now be visualized as a simple axis and customized to display not only any values but also branch lengths. Users can specify a custom value for the tree root and a branch length scaling factor that are then used to calculate the scale text labels. This allows a simple display of timescales or similar features associated with the tree.

Manual annotation tools

iTOL v5 introduces a set of basic manual drawing and labeling tools, allowing users to add simple shapes, polygons, lines and text labels anywhere on the tree display (Figure 1). These tools are not a replacement for a proper drawing or illustration software, but can help in adding simple additional annotations to the trees, simplifying the user's figure creation workflow and lowering the need for postprocessing.

EXPORT

One of iTOL's primary uses is the creation of high-quality figures for publication or inclusion into other documents. Due to the constantly increasing number of active users, the backend server has been extended to make the tree export faster. In peak usage times, the export can still be relatively

slow, so we are actively developing a queueing system, which will prevent the overloading of the export server and make the process more stable for all users. We expect the queueing system to be fully implemented in the next major release of iTOL.

Sharing exported tree figures

In addition to the standard tree export that downloads the created figure to the user computer, iTOL includes a system for the simple sharing of an unlimited number of figures. For each of their trees, users can create shared exported figures with individual descriptions and share a simple iTOL URL with their collaborators. The URL will display a table with an overview of all available shared exports for a tree, together with their format and description, allowing collaborators to simply access different tree figures created by the user. This feature can also be accessed through iTOL's main data sharing page (<https://itol.embl.de/sharing>). Shared tree figures remain stored on the iTOL server and permanently accessible, unless removed by the user.

SERVER ACCESS

iTOL is a self-sustaining tool and has been maintained and extended without any substantial public funding. In view of the constantly growing number of active users and uploaded trees, we have tried to find a sustainable model to maintain and further develop iTOL, serve the storage and CPU power needs, and provide timely technical support to our user base. Therefore, with version 5, albeit all tree annotation features remain freely available, certain account managing features of iTOL require an active subscription. Most user account management features also remain freely available apart from the saving of tree annotations to the iTOL server and the batch upload mode. User-uploaded trees and annotations remain permanently accessible, regardless of their subscription status.

FUNDING

This work has been partially supported by the German Federal Ministry of Education and Research with the project number 031A537B. Funding for open access charge: European Molecular Biology Laboratory.

Conflict of interest statement. None declared.

REFERENCES

1. Page, R.D. (1996) TreeView: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.*, **12**, 357–358.
2. Zmasek, C.M. and Eddy, S.R. (2001) ATV: display and manipulation of annotated phylogenetic trees. *Bioinformatics*, **17**, 383–384.
3. Letunic, I. and Bork, P. (2007) Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics*, **23**, 127–128.
4. Huerta-Cepas, J., Serra, F. and Bork, P. (2016) ETE 3: reconstruction, analysis, and visualization of phylogenomic data. *Mol. Biol. Evol.*, **33**, 1635–1638.
5. Yu, G. (2020) Using ggtree to visualize data on tree-like structures. *Curr. Protoc. Bioinform.*, **69**, e96.
6. Huson, D.H. and Scornavacca, C. (2012) Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. *Syst. Biol.*, **61**, 1061–1067.
7. Kreft, L., Botzki, A., Coppens, F., Vandepoele, K. and Van Bel, M. (2017) PhyD3: a phylogenetic tree viewer with extended phyloXML support for functional genomics data visualization. *Bioinformatics*, **33**, 2946–2947.
8. Subramanian, B., Gao, S., Lercher, M.J., Hu, S. and Chen, W.H. (2019) Evolview v3: a webserver for visualization, annotation, and management of phylogenetic trees. *Nucleic Acids Res.*, **47**, W270–W275.
9. Wilgenbusch, J.C. and Swofford, D. (2003) Inferring evolutionary trees with PAUP*. *Curr. Protoc. Bioinform.*, doi:10.1002/0471250953.bi0604s00.
10. Han, M.V. and Zmasek, C.M. (2009) phyloXML: XML for evolutionary biology and comparative genomics. *BMC Bioinform.*, **10**, 356.
11. Berger, S.A., Krompass, D. and Stamatakis, A. (2011) Performance, accuracy, and Web server for evolutionary placement of short sequence reads under maximum likelihood. *Syst. Biol.*, **60**, 291–302.
12. Matsen, F.A., Kodner, R.B. and Armbrust, E.V. (2010) pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinform.*, **11**, 538.
13. Lawley, B. and Tannock, G.W. (2017) Analysis of 16S rRNA gene amplicon sequences using the QIIME software package. *Methods Mol. Biol.*, **1537**, 153–163.
14. Letunic, I. and Bork, P. (2019) Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.*, **47**, W256–W259.
15. Bailey, T.L., Johnson, J., Grant, C.E. and Noble, W.S. (2015) The MEME suite. *Nucleic Acids Res.*, **43**, W39–W49.
16. Ling, C., Hamada, T., Gao, J., Zhao, G., Sun, D. and Shi, W. (2016) MrBayes tgMC(3)++: a high performance and resource-efficient GPU-oriented phylogenetic analysis method. *IEEE/ACM Trans. Comput. Biol. Bioinform.*, **13**, 845–854.