Evolview v2: an online visualization and management tool for customized and annotated phylogenetic trees

Zilong He[1,2], Huangkai Zhang[1,2], Shenghan Gao[1], Martin J. Lercher[3], Wei-Hua Chen[1,4,*] and Songnian Hu[1,2,*]

[1]CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics (BIG), Chinese Academy of Sciences (CAS), No.7 Beitucheng West Road, Chaoyang District, 100029 Beijing, PR China, [2]University of Chinese Academy of Sciences, Beijing 100049, China, [3]Institute for Computer Science and Cluster of Excellence on Plant Sciences CEPLAS, Heinrich Heine University, 40225 Düsseldorf, Germany and [4]BIG Data Center, Beijing Institute of Genomics (BIG), Chinese Academy of Sciences (CAS), No.7 Beitucheng West Road, Chaoyang District, 100029 Beijing, PR China

ABSTRACT

Evolview is an online visualization and management tool for customized and annotated phylogenetic trees. It allows users to visualize phylogenetic trees in various formats, customize the trees through built-in functions and user-supplied datasets and export the customization results to publication-ready figures. Its 'dataset system' contains not only the data to be visualized on the tree, but also 'modifiers' that control various aspects of the graphical annotation. Evolview is a single-page application (like Gmail); its carefully designed interface allows users to upload, visualize, manipulate and manage trees and datasets all in a single webpage. Developments since the last public release include a modern dataset editor with keyword highlighting functionality, seven newly added types of annotation datasets, collaboration support that allows users to share their trees and datasets and various improvements of the web interface and performance. In addition, we included eleven new 'Demo' trees to demonstrate the basic functionalities of Evolview, and five new 'Showcase' trees inspired by publications to showcase the power of Evolview in producing publication-ready figures. Evolview is freely available at: http://www.evolgenius.info/evolview/.

INTRODUCTION

Phylogenetic trees are an essential tool for displaying inferred evolutionary relationships among different biological species or other entities. Researchers can use phylogenetic trees to analyze species and gene family evolution (1,2) and identify horizontal gene transfers (3,4). In recent years, trees are increasingly used to visualize comparative information in an evolutionary context, e.g. when comparing the copy number of a gene in selected species or the number of genes involved in certain biological processes in each species. A wide variety of tree visualization software and online tools have thus been developed, such as the ETE toolkit (5), TreeView (6), TreeVector (7), Dendroscope (8), iTol (9), jsPhyloSVG (10), PhyloWidget (11), T-REX (12), Treelink (13), TreeGraph 2 (14) and many others (see https://en.wikipedia.org/wiki/List_of_phylogenetic_tree_visualization_software for a comprehensive list); each tool has its own advantages and disadvantages (15).

In 2012, we introduced Evolview, an online visualization and management tool for customized and annotated phylogenetic trees (16). Evolview takes phylogenetic trees of various formats such as Newick/Phylip, Nexus, Nhx and PhyloXML, and visualizes them as phylograms and cladograms, each in either rectangular or circular layout. It allows users to customize and annotate the trees with built-in functions as well as user-supplied datasets. In total, seven types of annotation datasets were supported in the first version; users can use datasets to change the display styles of leafs and internal branches or add additional information as pie and bar charts and as color strips to the tree. Its 'dataset system' contains not only the data to be visualized on the tree, but also 'modifiers' that control various aspects of the graphical annotation. The customized results can be exported into a wide range of graphical formats, including SVG, PDF, PNG, TIFF and JPEG. Exporting phylogenetic trees to textual formats such as Newick, Nexus, Nhx and PhyloXML is also supported. With a carefully de-

---

*To whom correspondence should be addressed. Tel: +86 108 299 5362; Fax: +86 108 299 5373; Email: chenwh550@gmail.com
Correspondence may also be addressed to Songnian Hu. Tel: +86 108 299 5362; Fax: +86 108 299 5373; Email: husn@big.ac.cn
Present address: Wei-Hua Chen, Department of Genetic Medicine and Development, University of Geneva Medical School, 1211 Geneva, Switzerland.
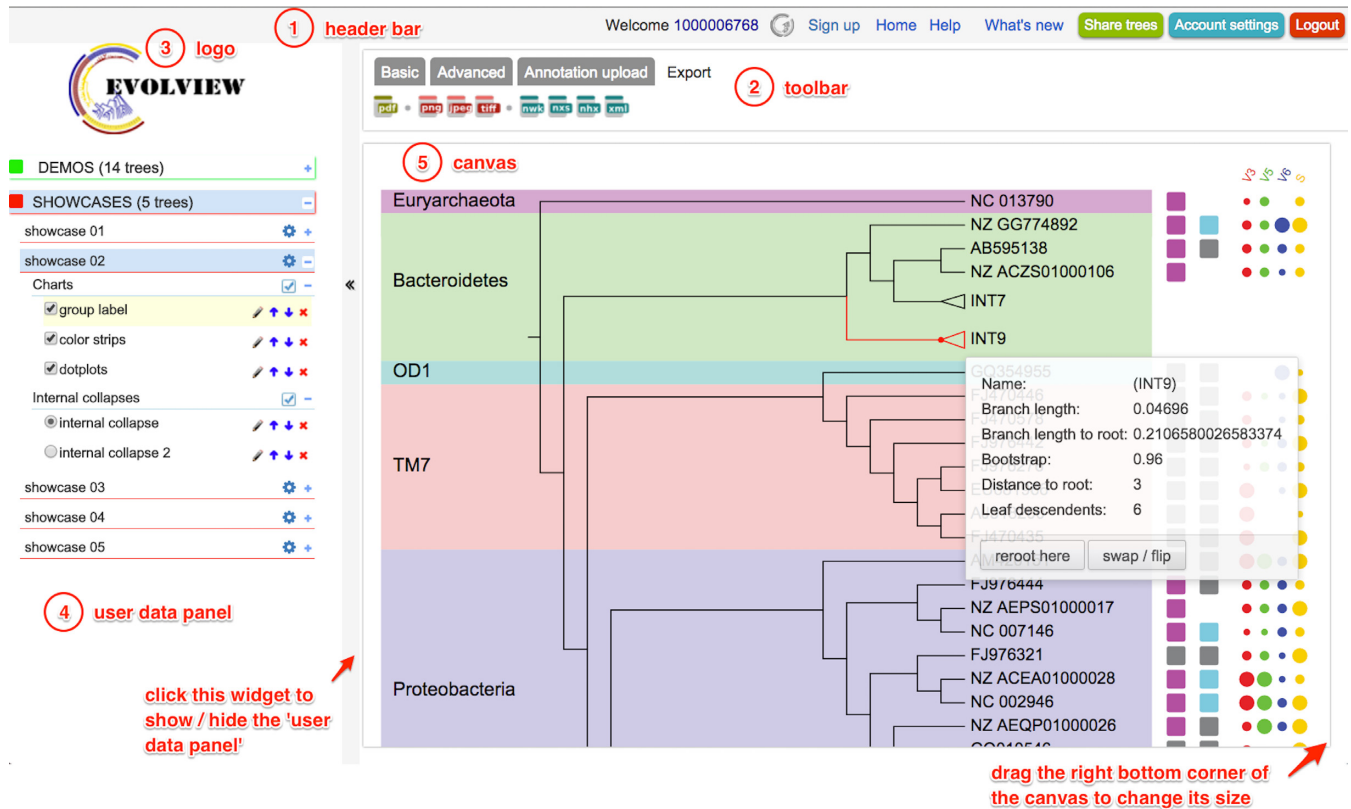
Figure 1. The user interface of Evolview. It consists of five parts. (1) A header bar. (2) A multi-tab toolbar consisting of sets of icons for tree and dataset manipulation, data upload and exporting to various text and graphical formats. (3) A logo. Clicking on the logo shows statistics on Evolview usage in a popup panel. (4) A user data panel that lists user-uploaded trees and datasets. Trees can be organized into projects. Datasets of a tree are organized into sub-categories. Trees and datasets can be easily manipulated using the icons. Users can choose to hide the user data panel to give more space to the canvas by clicking the widget to the right of this panel. (5) A canvas that shows the visualization results. Mouse-over on different elements of the tree and datasets opens up additional information. The size of the canvas can be changed by using the corresponding icons of the toolbar and by dragging the right bottom corner of the canvas.

signed user interface, users can view, manage and manipulate uploaded trees and datasets easily and efficiently (16).

Evolview is capable of handling large trees. For example, based on our tests on a 4-year old laptop with 2.2 GHz Intel Core i7 CPU and 8GB of memory, Evolview could parse and visualize two NCBI species trees of 2000 and 4000 leaf nodes in ∼7 and ∼17 s, respectively (for more information on the performance of Evolview, please consult https://evolview.codeplex.com/wikipage?title=evolviewperformance).

About 2456 users have uploaded 12 131 trees and 8605 datasets since 2012. More importantly, monthly usage has been increasing steadily over the years; on average, users upload more than 400 trees and 400 datasets per month in 2016.

In the past several years, we kept adding new functionalities to Evolview on a regular basis (see the 'What's new' page of our online documentation at https://evolview.codeplex.com). Major updates since the 2012 version include (i) a modern text editor that highlights 'modifiers' and predefined keywords in different colors and font styles, allowing users to easily compose and validate their datasets before submitting to Evolview; (ii) seven newly added types of annotation datasets that were often requested by our users; (iii) tree sharing functionality to allow users to share their

trees and datasets for collaborative purposes and (iv) various improvements of the web interface and performance. In addition, we added eleven new 'Demo' trees to demonstrate the basic functionalities of Evolview, and five new 'Showcase' trees to showcase the power of Evolview in producing complex and publication-ready figures. All 'Showcase' trees replicate figures from published papers, some of which were initially created using Evolview.

USER INTERFACE: MAIN FEATURES AND LATEST IMPROVEMENTS

Evolview is equipped with an intuitive user interface that provides access to most of its functionalities through clickable icons, such as uploading and managing trees and datasets, changing tree representation modes and tweaking visualization details (Figure 1). Evolview is designed as a single-page web application, where viewing, editing and manipulating of trees and datasets can be performed on the same web page (Figure 1).

Evolview supports 14 types of annotation datasets that allow users to apply complex customizations of the tree, twice as many as in the previous version. Datasets can be uploaded from text files or entered by hand using a built-in text editor. We updated the text editor so that it now sup-
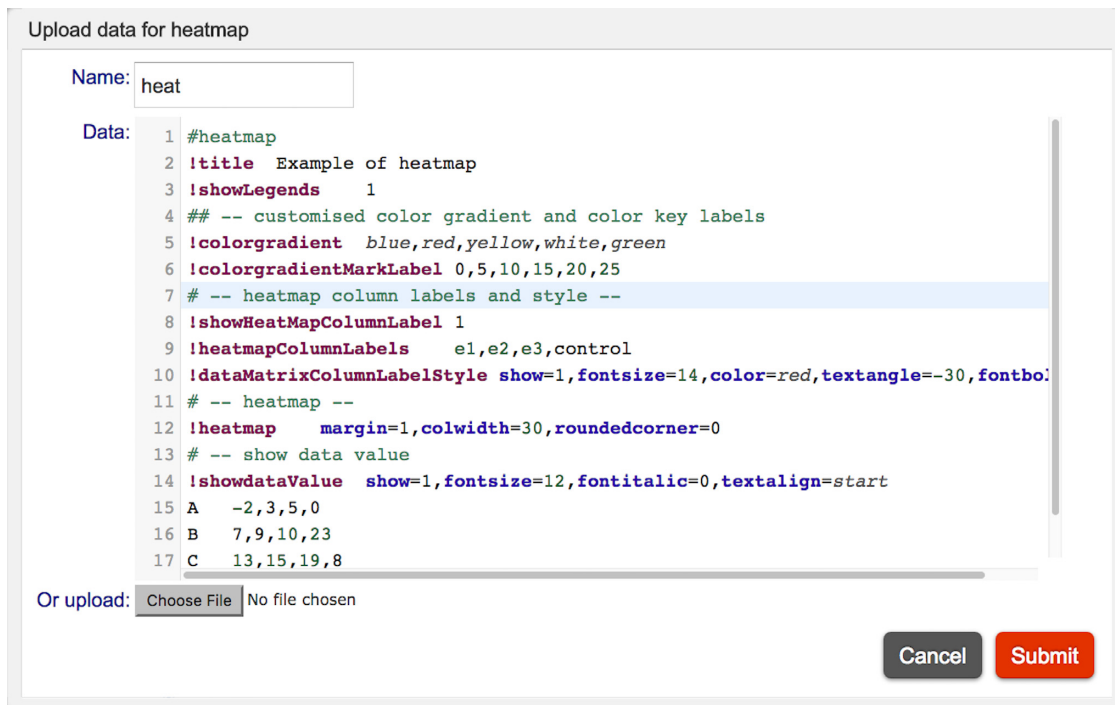
Figure 2. A modern text editor with syntax highlighting. This editor highlights different elements of a dataset with colors and font styles. In this new editor, the annotation lines are shown in green, supported 'modifiers' are shown in purple with bold font, predefined keywords are shown in dark blue with bold font, numbers are highlighted in light green and supported HTML color names are shown in light green with italic font. In addition, line numbers are shown to the left of the text and the current line is highlighted with a light blue background. The new editor is powered by the JavaScript library CodeMirror (https://codemirror.net/). Shown here is an example of a 'Heatmap' annotation; for more information on and examples of this annotation type please consult our online documentation at : https://evolview.codeplex.com/wikipage?title=DatasetHeatmap.

ports syntax highlighting. As shown in Figure 2, this new editor highlights different elements of a dataset with different colors and font styles. This modern text editor allows users to validate a dataset before submitting to Evolview.

Due to limited space of the webpage, the icons for uploading datasets are often quite small. We thus added graphical tooltips to these icons. Mouse-over on these icons triggers the graphical tooltips, on which a much larger and detailed illustration of the visualization is shown, with a link to the online documentation of the corresponding dataset type.

NEWLY ADDED ANNOTATION DATASET TYPES

Group labels

Group labels can be used to group genes or species in a tree into subgroups and annotate them, as shown in Figure 3A. Group labels can be applied to individual as well as multiple left nodes. Evolview provides five visualization styles for Group labels (Figure 3A, styles 1–5). For more examples, please consult the corresponding example trees of the 'Demos' project and 'Showcases' trees #1–3.
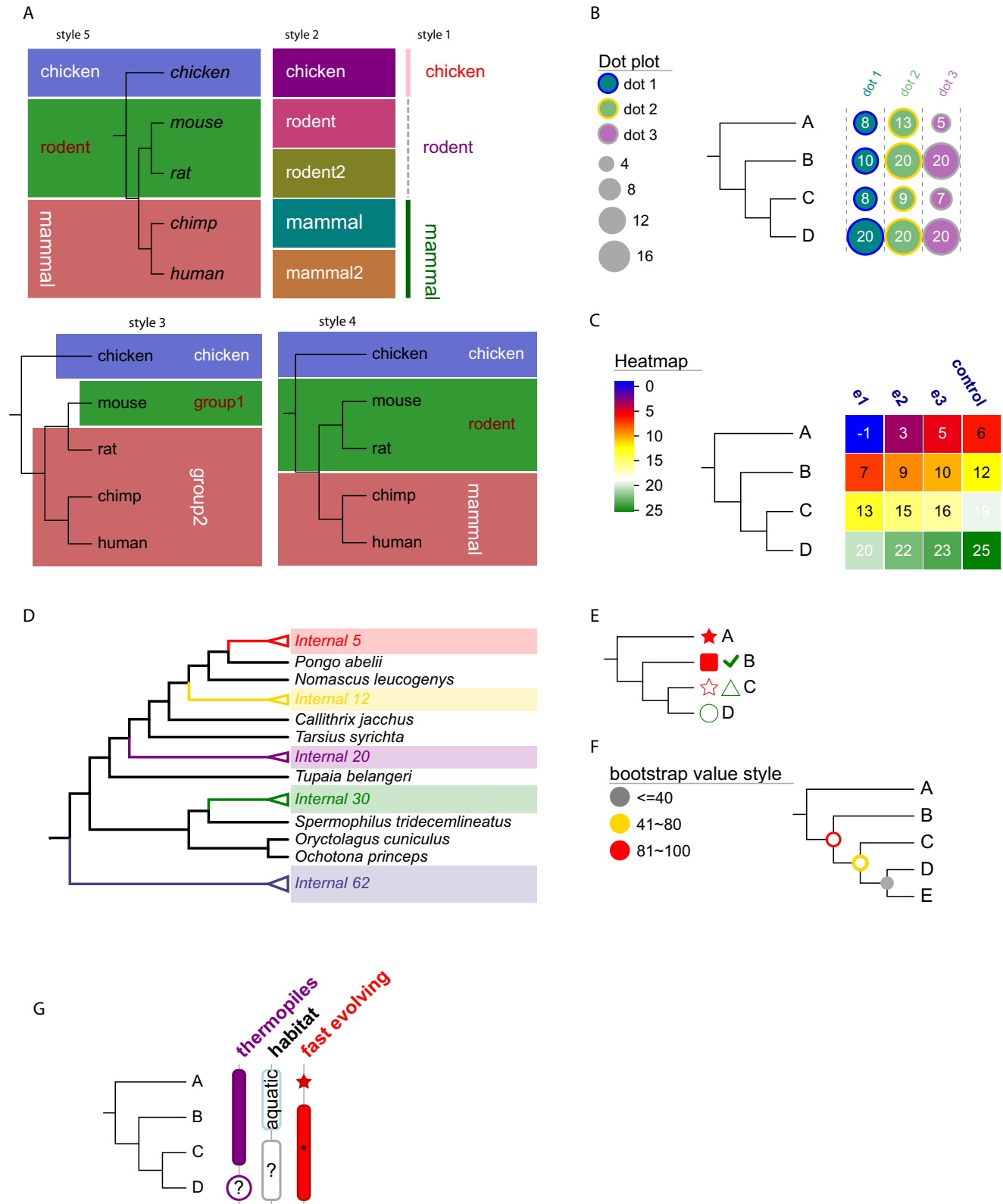
Dot plot

Dot plot displays user-supplied numerical data as dots (circles) of different sizes, with the diameters or the areas of the dots corresponding to the numerical values (Figure 3B). It is a variant of bar chart and can be used to show the variation of a particular attribute across leaf nodes. For example, Warinner *et al.* used dot plots to represent targeted 16S rRNA hypervariable regions in different species in the context of a phylogenetic tree (17) (see the 'Showcases' tree #2). The legend system has been updated so that in addition to the regular legend items, four dots of different sizes will be plotted to show the scaling of the user-supplied data (Figure 3B). See also the tree 'dotplots' of the 'Demos' project for examples.

Heatmap

Heatmap can be used to show gene and protein expression abundances in different experiments/samples or correlations among the samples (Figure 3C); it uses a multicolor gradient to show the relative strengths of the user-supplied numerical values. For example, Hornsby *et al.* used Heatmap to show the signal strengths of ELISA experiments using Evolview (18). Evolview provides multiple 'modifiers' for controlling various aspects of Heatmap, including the number and choices of the colors in the gradient and the names of experiments/samples and their display styles (Figure 3C). The color key, i.e. a legend item containing the color gradient and corresponding numeric values, is supported by the legend system and can be customized through multiple 'modifiers'. See the tree 'heatmaps' in the 'Demos' project and our online documentation for details and examples.

Figure 3. Seven newly supported types of annotation datasets. (**A**) Group labels of styles 1–5, (**B**) Dot plot, (**C**) Heatmap, (**D**) Collapsing at internal nodes, (**E**) Leaf label decoration, (**F**) Bootstrap value style and (**G**) Column plot.

Collapsing at internal nodes

Evolview supports collapsing a tree at internal nodes with user-supplied datasets. Users can collapse a tree at specified internal nodes, resulting in a small triangle at the collapsed nodes that is proportional in size to the number of contained leaf descendants. As shown in Figure 3D, customization of collapsed nodes is also supported.

Leaf label decorations

Leaf label decorations are color shapes shown in front of leaf labels (Figure 3E); we provide five different shapes ('star', 'rect' (for rectangle), 'triangle', 'circle' and 'check') and two display styles (filled and outlined). Multiple shapes can be added to a single leaf label, as shown in Figure 3E. See the trees 'leaf decoration and stroke' of the 'Demos' project and 'showcase 04' of the 'Showcases' project for examples.

Bootstrap value styles

Bootstrap value style allows users to tweak the displaying styles of bootstrap values (Figure 3F).

Column plots

Column plot provides a flexible way of visualizing summary information of species such as traits, lifestyles and habitats in the context of a phylogenetic tree (Figure 3G); it was inspired by a figure in a recent publication by Laumer *et al.* (19).

TREE SHARING

Users can share their trees with others, allowing them to view and manipulate the trees. Shared trees are available to all people (even non-registered users) that possess the generated web links (also known as URLs, Uniform Resource Locators). Evolview supports three modes of sharing; each provides the owners and others with different privileges with respect to the shared trees and associated datasets. The first of these is the 'Share' mode. This is the default mode for tree sharing. Other users, including non-registered users, can view and interact with the shared trees and datasets, but they cannot make any changes (or the changes will not be saved to the server. The owner of the trees will be able to change the shared trees and datasets; any changes will be available to others. The second mode is the 'Collaborate' mode. Shared trees under this mode can be manipulated by anyone with the URLs; all changes will be saved to the server. The third mode is the 'Publish' mode. Under this mode, a copy of selected trees and associated datasets will be shared. No one, including the owner, can make any changes to the shared trees and datasets.

NEW 'DEMOS' AND 'SHOWCASES' TREES

'Demos' and 'Showcases' are two projects that are available to all users. Trees and datasets in the 'Demos' project are used to show the basic functionality of Evolview. There are 14 trees in 'Demos' in the current version.
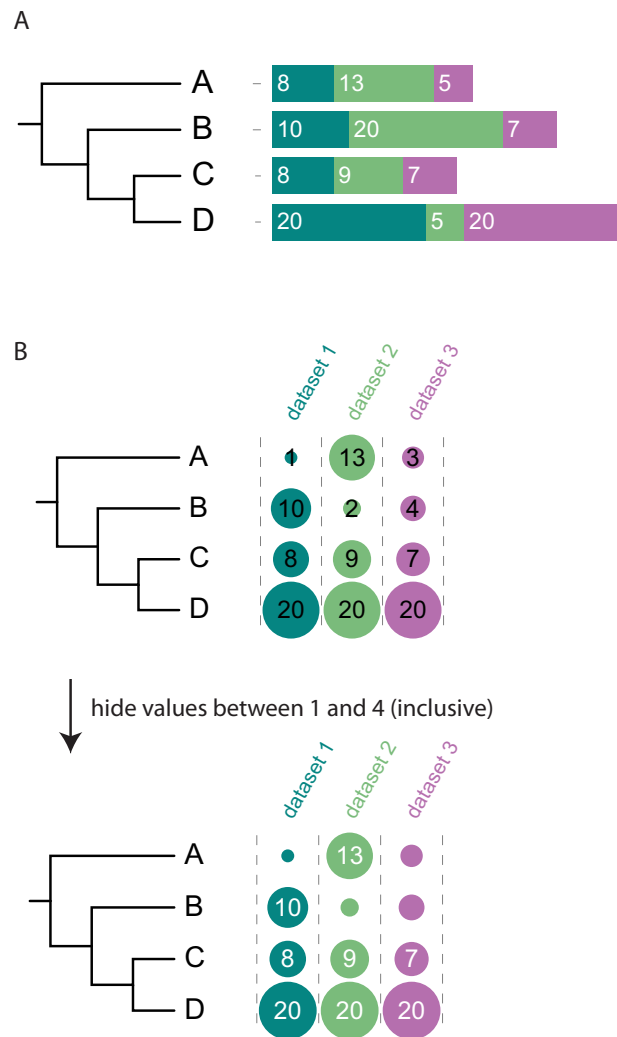
Figure 4. Displaying data values is supported for Bar plot (**A**), Pie charts (not shown), Dot plot (**B**) and Heatmap (see Figure 3C). Excluding certain values from the display is also supported (**B**).

Trees in the 'Showcases' project are used to demonstrate the power of Evolview in producing complex and publication-ready figures; these trees all reproduce figures from published papers (17,20–22), many of which were created using Evolview (20–22). Currently, there are five trees in the 'Showcases' project.

OTHER UPDATES

Many annotation datasets, such as 'Pie', 'Bar', 'Dotplot' and 'Heatmap' take numerical values as input. Now Evolview allows users to show these values at the corresponding positions of the visualization (Figure 3B and C; Figure 4). In 'Dotplot' and 'Heatmap', users can also choose to hide a single value or a range of values, as shown in Figure 4B.

Datasets for 'Dotplot', 'Heatmap' and 'Column' plots often contain column labels (Figure 3B and C). Evolview allows users to tweak the visualization styles of the column

labels by changing their color, style, size and rotation (see also Figure 3B and C).

CONCLUSION

In this article, we introduced Evolview v2, a comprehensive web application for visualizing, annotating and managing phylogenetic trees. Updates since the last version include support for new visualization styles, an improved user interface and better performance. In the future, we aim to keep Evolview up-to-date and state-of-the-art by supporting new annotation types that meet the developing needs of our broad range of users.

AVAILABILITY

Evolview is accessible using modern browsers such as Firefox, Chrome, Opera and Safari. However, due to some compatibility issues, Internet Explorer (IE) is currently not supported. The Evolview website is free and open to all users and there is no login requirement. However, if users want to save their data on our server, an account is required. Registration and the account are free.

Evolview is freely available at: http://www.evolgenius.info/evolview/. Its source codes are available upon request. Users who wish to install Evolview locally can find the compiled codes and instructions at: https://github.com/evolgeniusteam/evolivew-local-deploy. The documentation is hosted at: https://evolview.codeplex.com/documentation.

FUNDING

REFERENCES

1. Thornton,J.W. and DeSalle,R. (2000) Gene family evolution and homology: genomics meets phylogenetics. *Annu. Rev. Genomics Hum. Genet.*, **1**, 41–73.
2. Zardoya,R. (2005) Phylogeny and evolution of the major intrinsic protein family. *Biol. Cell*, **97**, 397–414.
3. Bapteste,E., Boucher,Y., Leigh,J. and Doolittle,W.F. (2004) Phylogenetic reconstruction and lateral gene transfer. *Trends Microbiol.*, **12**, 406–411.
4. Schonknecht,G., Chen,W.H., Ternes,C.M., Barbier,G.G., Shrestha,R.P., Stanke,M., Brautigam,A., Baker,B.J., Banfield,J.F., Garavito,R.M. *et al.* (2013) Gene transfer from bacteria and archaea facilitated evolution of an extremophilic eukaryote. *Science*, **339**, 1207–1210.
5. Huerta-Cepas,J., Dopazo,J. and Gabaldon,T. (2010) ETE: a python Environment for Tree Exploration. *BMC Bioinformatics*, **11**, 24.
6. Page,R.D. (2002) Visualizing phylogenetic trees using TreeView. *Curr. Protoc. Bioinformatics*, Chapter 6, Unit 6.2.
7. Pethica,R., Barker,G., Kovacs,T. and Gough,J. (2010) TreeVector: scalable, interactive, phylogenetic trees for the web. *PLoS One*, **5**, e8934.
8. Huson,D.H. and Scornavacca,C. (2012) Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. *Syst. Biol.*, **61**, 1061–1067.
9. Letunic,I. and Bork,P. (2011) Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. *Nucleic Acids Res.*, **39**, W475–W478.
10. Smits,S.A. and Ouverney,C.C. (2010) jsPhyloSVG: a javascript library for visualizing interactive and vector-based phylogenetic trees on the web. *PLoS One*, **5**, e12267.
11. Jordan,G.E. and Piel,W.H. (2008) PhyloWidget: web-based visualizations for the tree of life. *Bioinformatics*, **24**, 1641–1642.
12. Boc,A., Diallo,A.B. and Makarenkov,V. (2012) T-REX: a web server for inferring, validating and visualizing phylogenetic trees and networks. *Nucleic Acids Res.*, **40**, W573–W579.
13. Allende,C., Sohn,E. and Little,C. (2015) Treelink: data integration, clustering and visualization of phylogenetic trees. *BMC Bioinformatics*, **16**, 414.
14. Stover,B.C. and Muller,K.F. (2010) TreeGraph 2: combining and visualizing evidence from different phylogenetic analyses. *BMC Bioinformatics*, **11**, 7.
15. Pavlopoulos,G.A., Malliarakis,D., Papanikolaou,N., Theodosiou,T., Enright,A.J. and Iliopoulos,I. (2015) Visualizing genome and systems biology: technologies, tools, implementation techniques and trends, past, present and future. *Gigascience*, **4**, 38.
16. Zhang,H., Gao,S., Lercher,M.J., Hu,S. and Chen,W.H. (2012) EvolView, an online tool for visualizing, annotating and managing phylogenetic trees. *Nucleic Acids Res.*, **40**, W569–W572.
17. Warinner,C., Rodrigues,J.F., Vyas,R., Trachsel,C., Shved,N., Grossmann,J., Radini,A., Hancock,Y., Tito,R.Y., Fiddyment,S. *et al.* (2014) Pathogens and host immunity in the ancient human oral cavity. *Nat. Genet.*, **46**, 336–344.
18. Hornsby,M., Paduch,M., Miersch,S., Saaf,A., Matsuguchi,T., Lee,B., Wypisniak,K., Doak,A., King,D., Usatyuk,S. *et al.* (2015) A high through-put platform for recombinant antibodies to folded proteins. *Mol. Cell. Proteomics*, **14**, 2833–2847.
19. Laumer,C.E., Hejnol,A. and Giribet,G. (2015) Nuclear genomic signals of the 'microturbellarian' roots of platyhelminth evolutionary innovation. *Elife*, **4**, e05503.
20. Spandole,S., Cimponeriu,D., Berca,L.M. and Mihaescu,G. (2015) Human anelloviruses: an update of molecular, epidemiological and clinical aspects. *Arch. Virol.*, **160**, 893–908.
21. Chen,W.H., van Noort,V., Lluch-Senar,M., Hennrich,M.L., JA,H.W., Yus,E., Alibes,A., Roma,G., Mende,D.R., Pesavento,C. *et al.* (2016) Integration of multi-omics data of a genome-reduced bacterium: prevalence of post-transcriptional regulation and its correlation with protein abundances. *Nucleic Acids Res.*, **44**, 1192–1202.
22. Alam,M.T., Petit,R.A. 3rd, Crispell,E.K., Thornton,T.A., Conneely,K.N., Jiang,Y., Satola,S.W. and Read,T.D. (2014) Dissecting vancomycin-intermediate resistance in staphylococcus aureus using genome-wide association. *Genome Biol. Evol.*, **6**, 1174–1185.