Check for updates

Received: 3 November 2019 | Accepted: 5 January 2020

DOI: 10.1111/2041-210X.13358

APPLICATION

Methods in Ecology and Evolution | BRITISH ECOLOGICAL SOCIETY

VISUAL TREECMP: Comprehensive Comparison of Phylogenetic Trees on the Web

Tomasz Goluch | Damian Bogdanowicz | Krzysztof Giaro

Department of Algorithms and System Modeling, Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology, Gdańsk, Poland

Correspondence
Tomasz Goluch
Email: goluch@eti.pg.edu.pl

Handling Editor: Samantha Price

Abstract

1. We present VISUAL TREECMP—a package of applications for comparing phylogenetic tree sets.
2. VISUAL TREECMP includes a graphical web interface allowing the visualization of compared trees and command line application extended by comparison methods recently proposed in the literature.
3. The phylogenetic tree similarity analysis in VISUAL TREECMP can be performed using eighteen metrics, of which 11 are dedicated to rooted trees and seven to the unrooted ones. Additionally, in the case of different sets of labels, the compared trees can be pruned so that the sets are identical.
4. The VISUAL TREECMP is dedicated to people who need a universal and easy-to-use tool designed to compare arbitrary (not necessary binary) phylogenetic trees. An example of its use in a supertree accuracy analysis is described in section 3.
5. The VISUAL TREECMP package Web application is available at: (https://eti.pg.edu.pl/TreeCmp), and its source codes available under the GNU license can be reached through several repositories located at: (https://github.com/TreeCmp).

KEYWORDS

Matching Cluster metric, Matching Pair metric, Matching Split metric, phylogenetic tree, phylogenetics, supertree accuracy analysis, tree comparison, tree metrics

1 | INTRODUCTION

The construction of phylogenetic trees helps to understand evolutionary processes as mechanisms leading to genetic diversity. These trees allow an easy visualization of genetic relationships between populations or biological sequence sets. There are many methods for constructing such trees (e.g. maximum likelihood, maximum parsimony or Bayesian method) and most of them return not one but numerous sets of phylogenetic relations for the same input data. The comparison of the obtained set of trees requires the use of methods determining the distance between them, i.e. metrics. An alternative application of these metrics is related to the assessment of the accuracy of phylogenetic tree reconstruction, especially when developing a new method (Price, Dehal, & Arkin, 2010; Sul, Matthews, & Williams, 2009). The phylogenetic metrics are also used to build supertrees (Bansal, Burleigh, Eulenstein, & Fernández-Baca, 2010) and to define the consensus and median point of tree sets (Bryant, 1997). Another use of these methods is an efficient search for phylogenetic databases (Chen, Burleigh, Bansal, & Fernández-Baca, 2008). It is difficult to find a tool to compare large-size phylogenetic trees easily and efficiently in a comprehensive manner. The VISUAL TREECMP Web application, which is a successor of the command line utility TREECMP (Bogdanowicz, Giaro, & Wróbel, 2012), has been designed to fill this gap.

2 | FUNCTIONALITY OF VISUAL TREECMP

The VISUAL TREECMP package is a comprehensive toolset for comparing phylogenies that implements a rich set of phylogenetic tree comparison metrics. Metrics are divided into rooted or unrooted and topological or weighted, see Table 1. Weighted metrics use branch lengths which must be defined and positive. This assumption can be relaxed to non-negative values using 'Zero weights allowed' option, then the weight zero means that there is no edge between vertices and the vertices are suppressed into one vertex.

TABLE 1 List of the metrics available in the TREECMP package. The second column contains information whether the metric is rooted, the third column—whether it is weighted and the fourth—the authors of that metric

Metric name	Root.	Weig.	Cite
Triples metric	Yes	No	Critchlow et al. (1996)
Robinson–Foulds metric based on clusters	Yes	No	Robinson and Foulds (1981)
Matching Pair metric	Yes	No	Bogdanowicz and Giaro (2017)
Nodal Splitted metric with L^2 norm	Yes	No	Cardona et al. (2010)
Matching Cluster metric	Yes	No	Bogdanowicz and Giaro (2013)
Rooted maximum agreement subtree distance	Yes	No	Farach-Colton and Thorup (1995)
Cophenetic Metric with L^2 norm	Yes	No	Cardona, Mir, Rosselló, Rotger, and Sánchez (2013)
Weighted Robinson–Foulds metric based on clusters	Yes	Yes	Robinson and Foulds (1979)
Weighted Nodal Splitted metric with L^2 norm	Yes	Yes	Cardona et al. (2010)
Weighted Geodesic (BHV) rooted metric	Yes	Yes	Owen and Provan (2011)
Weighted Cophenetic Metric with L^2 norm	Yes	Yes	Cardona et al. (2013)
Quartet distance	No	No	Estabrook et al. (1985)
Path difference distance	No	No	Steel and Penny (1993)
Robinson–Foulds distance	No	No	Robinson and Foulds (1981)
Matching Split distance	No	No	Bogdanowicz and Giaro (2012)
Unrooted maximum agreement subtree distance	No	No	Farach-Colton and Thorup (1995)
Weighted Robinson–Foulds distance	No	Yes	Robinson et al. (1979)
Weighted Geodesic (BHV) Unrooted distance	No	Yes	Owen and Provan (2011)

According to our knowledge, this tool is the only one that implements the Matching Split (MS), Matching Cluster (MC) and Matching Pair (MP) metrics. All distances are implemented using polynomial time algorithms and all of them fulfil classic mathematical metric space axioms. It has been implemented in Java based on the Spring framework and is available in three forms, as a publicly available hosted web-based, stand-alone web-based and command line application. In this part, we will describe the main features of each version and describe the random phylogenetic trees generator added to VISUAL TREECMP package.

Web application is available at: https://eti.pg.edu.pl/treecmp, and can compare phylogenetic trees directly from the website. The input trees can be compared in one of four modes (overlapping pair, window, matrix, reference trees to all trees). The most useful are: a matrix comparison mode—allowing for a simple comparison of every two inputted trees and reference trees to all trees comparison mode—allowing for a comparison of each inputted tree to all reference trees. For instance, reference trees to all trees comparison modes are suited for analysing the accuracy of the Subtree Prune-and-Regraft (SPR)-supertree method; an example will be presented in Section 3.

The entered trees are in the NEXUS (Maddison, Swofford, & Maddison, 1997) or NEWICK (Felsenstein, 1989) format. In the comparison process, we have to select from 1 to 18 available metrics (11 for rooted and 7 for unrooted trees). Newick trees are interpreted as rooted, even if there is a multifurcation at the base. The unrooted metrics modify each tree by removing the root indication, and if a node of degree 2 is left behind, it is also suppressed.

The range of values returned by different metrics can vary considerably, moreover, the codomain of some metrics can be integers or arbitrary real numbers. It is therefore difficult to interpret the distance, and consequently the similarity of the two compared trees. To facilitate such analysis, the 'Normalized distances' option has been added for trees with less than 1,000 taxa (applies only to the use of topological metrics). It compares the obtained result with the averaged value of 1,000 randomly generated binary tree pairs of the same number of taxa. When the result is closer to one, the distance between the two examined trees is closer to the average value of the randomly generated trees in a given model.

In the case when label sets of compared trees differ, the option 'Prune trees' can be used to remove leaves that have no equivalent in the opposite compared tree. Since all the implemented metrics take as input two trees on the same set of leaves, trees on a different leaf set are pruned to subtrees that have the same set of leaves. Extra leaves are only pruned if 'Prune trees' option is selected, otherwise an error is generated.

The report of the comparisons is displayed in a table. Each row contains information about one pair of compared trees, see Figure 1. The succeeding columns contain the result for subsequent metrics. From the generated report, the user can easily sort against the content of any column, filter out the results containing a given string of characters, print the results or save them to the

2041210x, 2020, 4, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.13358 by Gdask University Of Technology, Wiley Online Library on [15/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

MOST WIEDZY Downloaded from mostwiedzy.pl

No	Tree1	Tree2	Tree1_taxa	Tree2_taxa	Common_taxa	R-F_Cluster	MatchingPair	MatchingClust
1	1	2	17	14	14	2.0000	13.0000	9.0000
2	1	3	17	17	17	1.0000	1.0000	2.0000
3	1	4	17	14	14	2.0000	11.0000	9.0000
4	1	5	17	14	14	1.0000	1.0000	2.0000
5	1	6	17	17	17	1.0000	1.0000	2.0000
6	2	3	14	17	14	2.0000	13.0000	9.0000
7	2	4	14	14	14	1.0000	3.0000	4.0000
8	2	5	14	14	14	2.0000	13.0000	9.0000
9	2	6	14	17	14	2.0000	13.0000	9.0000
10	3	4	17	14	14	1.0000	10.0000	7.0000

Show 10 entries

Showing 1 to 10 of 15 entries Previous 1 2 Next

Copy CSV Excel PDF Print

Name	Avg	Std	Min	Max	Count
MatchingCluster	5.066666666666666	3.60493488922554	0.0	9.0	15
MatchingPair	6.6	5.511200716117436	0.0	13.0	15
R-F_Cluster(0.5)	1.1333333333333333	0.7180219742846007	0.0	2.0	15

FIGURE 1 The report of phylogenies comparison

system clipboard or to a file in one of the following formats: csv, excel, pdf.

Additionally, by clicking at any row, the corresponding pair of compared trees can be displayed. The Phylo.io application implemented in the HTML5/JS technology (Robinson, Dylus, & Dessimoz, 2016) is used for visualization.

Stand-alone application with WEB GUI is launched locally on the user's computer. The GUI is available at the local address: http://localhost:<port_number>/TreeCmp/. In addition, the information on the calculation progress, possible warnings and errors is generated in a console window.

Command line application is a typical Java library simply called TreeCmp, and thus its functionality can be used in other projects. As a command line tool, it allows one to perform exactly the same calculations. Moreover when calculating MC and MS metrics additional files containing aligned splits or clusters can be generated. It is well-suited for use in scripts, e.g. with the below-described generator of random phylogenetic trees.

Phylogenetic random trees generator allows generating a number of binary trees given by the user. The trees can be generated according to the Yule (Harding, 1971) or uniform model. Also, all possible binary trees for a fixed number of leaves can be generated. However, for a larger number of leaves huge amounts of disc spaces are required.

Visual TreeCmp package allows one to normalize metric values. If the number of tree pairs included in a random sample (currently 1,000) is too small, the calculations can be repeated locally by generating a larger number of tree pairs and calculating more accurate mean and standard deviation values. This makes the interpretation of the normalized result more reliable.

3 | AN EXAMPLE OF SUPERTREE ACCURACY ANALYSING

Supertree methods allow constructing trees (called supertrees) that combine phylogenetic information represented by a set of smaller phylogenies with partially overlapped taxa. We used Visual TreeCmp to analyse the accuracy of two variants of the proposed SPR-supertree method (Whidden, Zeh, & Beiko, 2014) that is based on computing SPR distances between binary trees. The idea of supertree distance methods, which SPR-supertree belongs to, is that we are trying to find a tree that minimizes the sum of its distances in a particular metric to source phylogenies. Since a supertree contains more taxa than each of

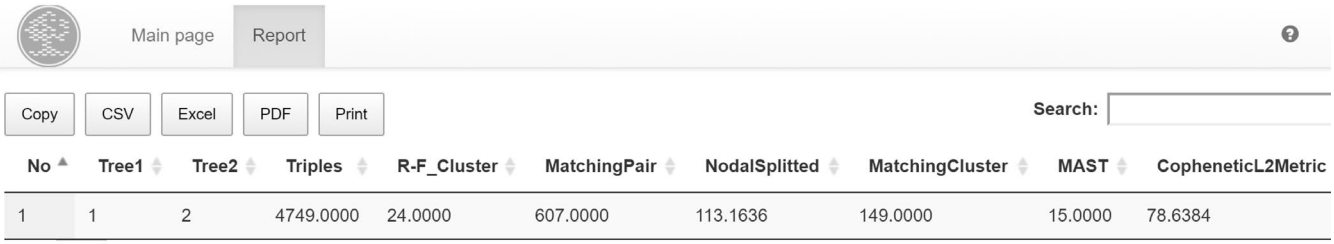
2041210x, 2020, 4, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.13358 by Gdask University Of Technology, Wiley Online Library on [15/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

MOST WIEDZY Downloaded from mostwiedzy.pl

the source trees and because most of the phylogenetic metrics require trees with the same leaf sets, before applying a particular method to compute the distance, the supertree has to be pruned to fulfil the requirement. This operation can be easily accomplished in VISUAL TREECMP by simply enabling 'Prune trees' in the user interface.

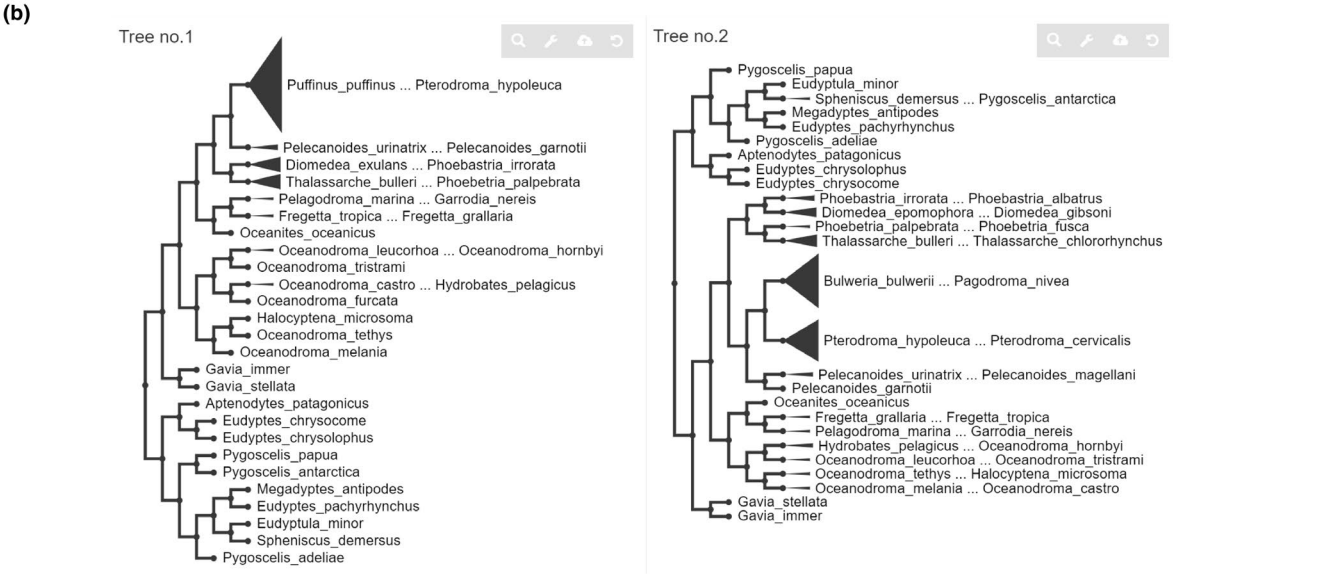
We then analysed in detail the accuracy of seabirds supertrees computed by two variants of SPR-supertree method. SPR-RF-TIES and SPR seeded with an Matrix Representation with Parsimony starting tree (SPR-MRP) proposed in Whidden et al. (2014), see Figure 2. The seabirds dataset (Kennedy & Page, 2002) contains 121 different taxa and consists of seven source trees having from 14 to 90 leaves. Based on the analysis presented in Whidden et al. (2014) we know that an average SPR distance of both the supertrees to source trees equals 17/7≈2.43 and they receive the same parsimony score of 208. SPR-MRP is slightly better (avg. dist. approx. 4.36) than SPR-RF-TIES (4.5) according to RF metric. However, there are many other metrics defined in the literature whose use may lead to a different conclusion. A comparison of the two above-mentioned trees made by using VISUAL TREECMP is presented in Figure 2. In Figure 2a we see the distance between SPR-RF-TIES and SPR-MRP supertrees in seven metrics. The graphical representation of

(a)



No	Tree1	Tree2	Triples	R-F_Cluster	MatchingPair	NodalSplitted	MatchingCluster	MAST	CopheneticL2Metric
1	1	2	4749.0000	24.0000	607.0000	113.1636	149.0000	15.0000	78.6384

(b)



(c)

Name	Avg	Std
CopheneticL2Metric	11.211589759136334	6.719925440804543
MAST	4.714285714285714	4.025429372458677
MatchingCluster	20.571428571428573	14.802095825129848
MatchingPair	22.0	17.062908812475587
NodalSplitted	26.684369194743688	28.07798091222347
R-F_Cluster	4.357142857142857	3.934411243977261
Triples	52.57142857142857	45.04872645371943

(d)

Name	Avg	Std
CopheneticL2Metric	14.337238608724332	13.264372924360366
MAST	5.0	4.105745103771404
MatchingCluster	26.285714285714285	25.946176628410072
MatchingPair	49.785714285714285	106.59012187643212
NodalSplitted	26.65195974374596	26.52818256852839
R-F_Cluster	4.428571428571429	3.5949227234747476
Triples	238.5	663.243410392121

FIGURE 2 Supertree accuracy analysis in VISUAL TREECMP. (a) The report presenting Subtree Prune-and-Regraft seeded with an Matrix Representation with Parsimony starting tree (SPR-MRP) (Tree1) and SPR-RF-TIES (Tree2) supertrees distance in seven metrics. (b) Graphical representation of analysed trees available after clicking on a suitable row in the report. (c) Average distance of SPR-MRP supertree to source trees—a part of summary report computed in Ref-to-all mode. (d) Analogical report to the previous one for SPR-RF-TIES supertree

2041210x, 2020, 4, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.13358 by Gdask University Of Technology, Wiley Online Library on [15/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

MOST WIEDZY Downloaded from mostwiedzy.pl

the trees, obtained using embedded Phylo.io library (Robinson et al., 2016), is presented in Figure 2b. When analysing summary reports containing average distances of analysed supertrees to source trees (Figure 2c,d) we noticed that according to almost all metrics (six out of seven) the SPR-MRP tree is better. Although, in the case of Nodal Splitted metric (Cardona, Llabrés, Rosselló, & Valiente, 2010) the order is reversed, the relative difference between the results is quite small (in comparison to std. deviation for example), i.e. ≈26.68 for SPR-MRP and ≈26.62 for SPR-RF-TIES.

4 | COMPARISON WITH SIMILAR APPLICATIONS

According to OMICtools metadatabase (Henry, Bandrowski, Pepin, Gonzalez, & Desfeux, 2014), there are 10 applications related to Tree Comparison category. Most of them are desktop programs which, unlike web applications, need to be installed on a user's computer. The installation of such software is often associated with additional requirements for runtime environments, available compilers and interpreters and limitations on supported operating systems. That is why web-based tools are becoming increasingly popular. We would also like to review briefly the tools that are similar to Visual TreeCmp in their functionality.

One of three applications described in OMICtools that provides its functionality directly on the website is **Phylo.io** (Robinson et al., 2016). It is a convenient tool for the visualization of phylogenetic trees. It uses such technologies as HTML5, CSS, Ajax, jQuery, and the D3 JavaScript visualization library which ensure multiplatforming of Phylo.io. Moreover, it can be embedded easily in other websites. We have included this feature in our application, i.e. we have applied Phylo.io to visualize the compared trees. In the Tree Comparison category, this application has less to offer. It implements only a variant of Jaccard Index, which can specify the best matching node in the compared tree.

The second web application, implemented as a Java applet, is **Compare2Trees** (Nye, Liò, & Gilks, 2006). The tool can compare two trees at the same time. It calculates one metric, which is based on computing the best matching between branches. The newest version (2011) of this applet is not supported by many browsers, which is confirmed by the message on the project website. Launching a stand-alone version of the program with the latest version of Java virtual machine is not possible either.

TreeVersity2 (Guerra-Gomez, Pack, Plaisant, & Shneiderman, 2013) is the third tool with a web interface. It provides five types of comparisons described by numbers from 0 to 4. Type 0, which is defined as 'Topological differences between two trees where the nodes contain only a label' and corresponds to the problem of phylogeny, is not supported by this tool. TreeVersity2 is dedicated to a richer set of problems, including label weights in external and internal vertices (Types 1–4). This application has also been included in the Phylogenetic Tree Visualization category. However, this tool is not dedicated to the visualization of phylogenetic trees.

TOPD/FMTS 'TOPological Distance/From Multiple To Single' (Puigbò, Garcia-Vallvé, & McInerney, 2007) is a perl script allowing calculation of Robinson–Foulds (RF; Robinson & Foulds, 1981), Path Difference (PD; Steel & Penny, 1993), Quartet (QT; Estabrook, McMorris, & Meacham, 1985) and Triples (TT; Critchlow, Pearl, & Qian, 1996) metrics. However, it works much slower than Visual TreeCmp, e.g. the comparison of two rooted trees with 100 leaves takes more than 30 min when using TT metric, while in TreeCmp it takes less than 1 s (Bogdanowicz et al., 2012).

5 | CONCLUSIONS

The Visual TreeCmp is dedicated to people who need a universal and easy-to-use tool designed to compare different arbitrary (not necessary binary) phylogenetic trees. For instance, it has been used during research on a new subpopulation of *Helicobacter pylori* from Australia (Lamichhane, Wise, Chua, Marshall, & Tay, 2019). It is characterized by a rich collection of metrics, which enable the compared trees visualization and an easy generation of reports. In the visual version, the ease of its use is ensured by a graphical interface and the access to this application by a WEB service. The application is accessible in an open repository and with GNU open source license.

ACKNOWLEDGEMENT

The authors declare no conflict of interest to the publication of this work.

AUTHORS' CONTRIBUTIONS

T.G. and D.B. implemented the Visual TreeCmp application. T.G. and D.B. implemented the TreeCmp 2.0 application. D.B. and K.G. are the authors of the MC, MS and MP metrics. T.G., D.B. and K.G. implemented metrics and are the authors of the documentation and the manuscript in the proportions mentioned within parenthesis.

DATA AVAILABILITY STATEMENT

Data available from the Github repositories: https://doi.org/10.5281/zenodo.3608800 (TreeCmp) (KNAG, TreeCmp, & Wong, 2020); https://doi.org/10.5281/zenodo.3608806 (TreeCmpWEB) (KNAG & TReeCmp, 2020a); https://doi.org/10.5281/zenodo.3608808 (PRTGen) (KNAG & TreeCmp, 2020b).

ORCID

Tomasz Goluch https://orcid.org/0000-0003-0544-7947
Damian Bogdanowicz https://orcid.org/0000-0002-5051-9287
Krzysztof Giaro https://orcid.org/0000-0001-8741-860X

REFERENCES

Bansal, M. S., Burleigh, J. G., Eulenstein, O., & Fernández-Baca, D. (2010). Robinson-foulds supertrees. *Algorithms for Molecular Biology*, *5*(1), 18. https://doi.org/10.1186/1748-7188-5-18

Bogdanowicz, D., & Giaro, K. (2012). Matching split distance for unrooted binary phylogenetic trees. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *9*(1), 150–160. https://doi.org/10.1109/tcbb.2011.48

2041210x, 2020, 4, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.13358 by Gdask University Of Technology, Wiley Online Library on [15/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

MOST WIEDZY Downloaded from mostwiedzy.pl

Bogdanowicz, D., & Giaro, K. (2013). On a matching distance between rooted phylogenetic trees. *International Journal of Applied Mathematics and Computer Science*, *23*, 669–684. https://doi.org/10.2478/amcs-2013-0050

Bogdanowicz, D., & Giaro, K. (2017). Comparing phylogenetic trees by matching nodes using the transfer distance between partitions. *Journal of Computational Biology*, *24*(5), 422–435. https://doi.org/10.1089/cmb.2016.0204

Bogdanowicz, D., Giaro, K., & Wróbel, B. (2012). TreeCmp: Comparison of trees in polynomial time. *Evolutionary Bioinformatics*, *8*, 9657. https://doi.org/10.4137/EBO.S9657

Bryant, D. (1997). *Building trees, hunting for trees, and comparing trees: Theory and methods in phylogenetic analysis*. PhD thesis, University of Canterbury, Christchurch, New Zealand. Retrieved from https://ir.canterbury.ac.nz/handle/10092/7914

Cardona, G., Llabrés, M., Rosselló, F., & Valiente, G. (2010). Nodal distances for rooted phylogenetic trees. *Journal of Mathematical Biology*, *61*(2), 253–276. https://doi.org/10.1007/s00285-009-0295-2

Cardona, G., Mir, A., Rosselló, F., Rotger, L., & Sánchez, D. (2013). Cophenetic metrics for phylogenetic trees, after Sokal and Rohlf. *BMC Bioinformatics*, *14*(1), 3. https://doi.org/10.1186/1471-2105-14-3

Chen, D., Burleigh, J. G., Bansal, M. S., & Fernández-Baca, D. (2008). PhyloFinder: An intelligent search engine for phylogenetic tree databases. *BMC Evolutionary Biology*, *8*(1), 90. https://doi.org/10.1186/1471-2148-8-90

Critchlow, D. E., Pearl, D. K., & Qian, C. (1996). The triples distance for rooted bifurcating phylogenetic trees. *Systematic Biology*, *45*(3), 323–334. https://doi.org/10.1093/sysbio/45.3.323

Estabrook, G. F., McMorris, F. R., & Meacham, C. A. (1985). Comparison of undirected phylogenetic trees based on subtrees of four evolutionary units. *Systematic Biology*, *34*(2), 193–200. https://doi.org/10.2307/sysbio/34.2.193

Farach-Colton, M., & Thorup, M. (1995). Fast comparison of evolutionary trees. *Inform Comput*, *123*, 29–37. https://doi.org/10.1145/314464.314629

Felsenstein, J. (1989). PHYLIP-Phylogeny inference package (Version 3.2). *Cladistics*, *5*, 164–166.

Guerra-Gomez, J. A., Pack, M. L., Plaisant, C., & Shneiderman, B. (2013). Visualizing change over time using dynamic hierarchies: TreeVersity2 and the StemView. *IEEE Transactions on Visualization and Computer Graphics*, *19*(12), 2566–2575. https://doi.org/10.1109/TVCG.2013.231

Harding, E. F. (1971). The probabilities of rooted tree-shapes generated by random bifurcation. *Advances in Applied Probability*, *3*(1), 44–77. https://doi.org/10.2307/1426329

Henry, V. J., Bandrowski, A. E., Pepin, A., Gonzalez, B. J., & Desfeux, A. (2014). OMICtools: an informative directory for multi-omic data analysis. *Database*, *2014*, bau069. https://doi.org/10.1093/database/bau069

Kennedy, M., & Page, R. D. M. (2002). Seabird supertrees: Combining partial estimates of procellariiform phylogeny. *The Auk*, *119*(1), 88–108. https://doi.org/10.1093/auk/119.1.88

KNAG, & TreeCmp. (2020a). Data from: TreeCmp/TreeCmpWEB v1.0.89 (version v1.0.89). *Zenodo*, https://doi.org/10.5281/zenodo.3608806

KNAG, & TreeCmp. (2020b). Data from: TreeCmp/PRTGen v1.0.31 (version v1.0.31). *Zenodo*, https://doi.org/10.5281/zenodo.3608808

KNAG, TreeCmp, & Wong, Y. (2020). Data from: TreeCmp/TreeCmp v2.0.76 (version v2.0.76). *Zenodo*, https://doi.org/10.5281/zenodo.3608800

Lamichhane, B., Wise, M. J., Chua, E. G., Marshall, B. J., & Tay, C. Y. (2019). A novel taxon selection method, aimed at minimizing recombination, clarifies the discovery of a new sub-population of *Helicobacter pylori* from Australia. *Evolutionary Applications*. https://doi.org/10.1111/eva.12864

Maddison, D. R., Swofford, D. L., & Maddison, W. P. (1997). Nexus: An extensible file format for systematic information. *Systematic Biology*, *46*(4), 590–621. https://doi.org/10.1093/sysbio/46.4.590

Nye, T. M. W., Liò, P., & Gilks, W. R. (2006). A novel algorithm and web-based tool for comparing two alternative phylogenetic trees. *Bioinformatics*, *22*(1), 117–119. https://doi.org/10.1093/bioinformatics/bti720

Owen, M., & Provan, J. S. (2011). A fast algorithm for computing geodesic distances in tree space. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *8*(1), 2–13. https://doi.org/10.1109/TCBB.2010.3

Price, M. N., Dehal, P. S., & Arkin, A. P. (2010). FastTree 2 – Approximately maximum-likelihood trees for large alignments. *PLoS ONE*, *5*(3), 1–10. https://doi.org/10.1371/journal.pone.0009490

Puigbò, P., Garcia-Vallvé, S., & McInerney, J. O. (2007). TOPD/FMTS: A new software to compare phylogenetic trees. *Bioinformatics*, *23*(12), 1556–1558. https://doi.org/10.1093/bioinformatics/btm135

Robinson, D. F., & Foulds, L. R. (1979). Comparison of weighted labelled trees. In A. F. Horadam & W. D. Wallis (Eds.), *Combinatorial mathematics VI* (pp. 119–126). Berlin, Heidelberg: Springer.

Robinson, D. F., & Foulds, L. R. (1981). Comparison of phylogenetic trees. *Mathematical Biosciences*, *53*(1), 131–147. https://doi.org/10.1016/0025-5564(81)90043-2

Robinson, O., Dylus, D., & Dessimoz, C. (2016). Phylo.io: Interactive viewing and comparison of large phylogenetic trees on the web. *Molecular Biology and Evolution*. *33*, 2163–2166. https://doi.org/10.1093/molbev/msw080

Steel, M. A., & Penny, D. (1993). Distributions of tree comparison metrics—Some new results. *Systematic Biology*, *42*(2), 126–141. https://doi.org/10.1093/sysbio/42.2.126

Sul, S., Matthews, S., & Williams, T. L. (2009). Using tree diversity to compare phylogenetic heuristics. *BMC Bioinformatics*, *10*. https://doi.org/10.1186/1471-2105-10-S4-S3

Whidden, C., Zeh, N., & Beiko, R. G. (2014). Supertrees based on the subtree prune-and-regraft distance. *Systematic Biology*, *63*(4), 566–581. https://doi.org/10.1093/sysbio/syu023

How to cite this article: Goluch T, Bogdanowicz D, Giaro K. Visual TreeCmp: Comprehensive Comparison of Phylogenetic Trees on the Web. *Methods Ecol Evol*. 2020;11:494–499. https://doi.org/10.1111/2041-210X.13358

MOST WIEDZY Downloaded from mostwiedzy.pl

2041210x, 2020, 4, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.13358 by Gdask University Of Technology, Wiley Online Library on [15/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License