Dryad Data Package for study “DateLife: leveraging databases and analytical tools to reveal the dated Tree of Life”

This README file was generated on 2022-07-17 by Luna L. Sánchez Reyes, <https://orcid.org/0000-0001-7668-2528>

GENERAL INFORMATION

1. Title of Dataset

* Data from: DateLife: leveraging databases and analytical tools to reveal the dated Tree of Life.

1. Author Information

* Corresponding Researcher  
  name: Luna L. Sanchez Reyes  
  institution: University of California, Merced, USA  
  email: [sanchez.reyes.luna@gmail.com](mailto:sanchez.reyes.luna@gmail.com)
* Co-researcher 1  
  name: Emily Jane McTavish  
  institution: University of California, Merced, USA
* Co-researcher 2  
  name: Brian C. O’Meara  
  institution: University of Tennessee, Knoxville, USA

1. Date of data collection: 2022-01-28
2. Geographic location of data collection: Online
3. Funding sources that supported the collection of the data: National Science Foundation, USA
4. Recommended citation for this dataset:

* Sánchez-Reyes LL, McTavish EJ & O’Meara BC. (2019). DateLife: leveraging databases and analytical tools to reveal the dated Tree of Life. BioRxiv, 782094.
* Data from: DateLife: leveraging databases and analytical tools to reveal the dated Tree of Life. Dryad, Dataset

DATA & FILE OVERVIEW

1. Description of dataset

These data were generated to investigate and showcase the performance of the datelife R package (<https://github.com/phylotastic/datelife>). We showcased the application of the package with one mock example and two different biological examples.

The first biological example uses datelife on a small sample of bird species. The second one uses datelife on bird species belonging to the family Fringillidae of “true finches”, following the NCBI taxonomy. The mock example was created by replacing species names from the small example with letters, and reducing the number of source chronograms.

We evaluated the performance of the package datelife with two analysis: a benchmarking analysis to measure computing time of functions, and a cross validation analysis to test the accuracy and precision of the functions.

1. File List:

File 1 Name: Sanchez-Reyes\_etal\_2022\_table\_1.csv  
File 1 Description: CSV (comma separated values) file with node age results from datelife’s small example, presented in Table 1 of the manuscript. It contains node names, taxon names defining the nodes, the corresponding node ages (in Million years), and references of source chronograms that were congruified to the chosen tree topology to extract node ages.

File 2 Name: Sanchez-Reyes\_etal\_2022\_table\_2.csv  
File 2 Description: CSV (comma separated values) file summarizing age data (in Million years) from table 1, per node. These data are presented in Table 2 of the manuscript.

File 3 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_table\_S1.csv  
File 3 Description: CSV (comma separated values) file with node age results from datelife’s Fringillidae example. It contains node names, taxon names defining the nodes, the corresponding node ages (in Million years), and references of source chronograms that were congruified to the chosen tree topology (shown in Figure 4B of the manuscript) to extract node ages.

File 4 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_table\_S1.pdf  
File 4 Description: PDF version of file Sanchez-Reyes\_etal\_2022\_supplementary\_table\_S1.csv

File 5 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_table\_S2.csv  
File 5 Description: CSV (comma separated values) file summarizing age data (in Million years) from table S1, per node.

File 6 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_table\_S2.pdf  
File 6 Description: PDF file version of file Sanchez-Reyes\_etal\_2022\_supplementary\_table\_S2.csv.

File 7 Name: Sanchez-Reyes\_etal\_2022\_figure\_1\_chronogram\_mock\_example.tre  
File 7 Description: Newick version of chronogram from the mock example shown in Figure 1 of the manuscript (<https://www.biorxiv.org/content/10.1101/782094v2>).

File 8 Name: Sanchez-Reyes\_etal\_2022\_figure\_3\_chronogram\_small\_example.tre  
File 8 Description: Newick version of chronogram from the small example shown in Figure 3 of the manuscript (<https://www.biorxiv.org/content/10.1101/782094v2>).

File 9 Name: Sanchez-Reyes\_etal\_2022\_figure\_4A\_topology\_finches\_mrca.tre  
File 9 Description: Newick version of topology extracted from the Open Tree of Life synthetic tree, shown in Figure 4A of the manuscript (<https://www.biorxiv.org/content/10.1101/782094v2>).

File 10 Name: Sanchez-Reyes\_etal\_2022\_figure\_4B\_topology\_finches\_ncbi.tre  
File 10 Description: Newick version of topology extracted from the Open Tree of Life synthetic tree, shown in Figure 4B of the manuscript (<https://www.biorxiv.org/content/10.1101/782094v2>).

File 11 Name: Sanchez-Reyes\_etal\_2022\_figure\_5\_chronogram\_finches\_example.tre  
File 11 Description: Newick version of chronogram from the finch example shown in Figure 5 of the manuscript (<https://www.biorxiv.org/content/10.1101/782094v2>).

File 12 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S1.pdf  
File 12 Description: PDf file showing results of cross validation analysis of datelife’s chronogram generating functions.

File 13 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S2.jpg  
File 13 Description: JPG file of results of cross validation analysis of datelife’s chronogram generating functions, using a tree topology from Barker et al. 2013. Comparison of original chronogram (black) and the chronogram obtained using datelife (gray).

File 14 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S2.tre  
File 14 Description: Newick file of chronogram obtained with datelife, shown in supplementary Figure S2 (gray).

File 15 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S3.jpg  
File 15 Description: JPG file of results of cross validation analysis of datelife’s chronogram generating functions, using a tree topology from Barker et al. 2015, chronogram 1. Comparison of original chronogram (black) and the chronogram obtained using datelife (gray).

File 16 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S3.tre  
File 16 Description: Newick file of chronogram obtained with datelife, shown in supplementary Figure S3 (gray).

File 17 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S4.jpg  
File 17 Description: JPG file of results of cross validation analysis of datelife’s chronogram generating functions, using a tree topology from Barker et al. 2015, chronogram 2. Comparison of original chronogram (black) and the chronogram obtained using datelife (gray).

File 18 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S4.tre  
File 18 Description: Newick file of chronogram obtained with datelife, shown in supplementary Figure S4 (gray).

File 19 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S5.jpg  
File 19 Description: JPG file of results of cross validation analysis of datelife’s chronogram generating functions, using a tree topology from Burns et al. 2015. Comparison of original chronogram (black) and the chronogram obtained using datelife (gray).

File 20 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S5.tre  
File 20 Description: Newick file of chronogram obtained with datelife, shown in supplementary Figure S5 (gray).

File 21 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S6.jpg  
File 21 Description: JPG file of results of cross validation analysis of datelife’s chronogram generating functions, using a tree topology from Claramunt et al. 2015. Comparison of original chronogram (black) and the chronogram obtained using datelife (gray).

File 22 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S6.tre  
File 22 Description: Newick file of chronogram obtained with datelife, shown in supplementary Figure S6 (gray).

File 23 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S7.jpg  
File 23 Description: JPG file of results of cross validation analysis of datelife’s chronogram generating functions, using a tree topology from Gibb et al. 2015. Comparison of original chronogram (black) and the chronogram obtained using datelife (gray).

File 24 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S7.tre  
File 24 Description: Newick file of chronogram obtained with datelife, shown in supplementary Figure S7 (gray).

File 25 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S8.jpg  
File 25 Description: JPG file of results of cross validation analysis of datelife’s chronogram generating functions, using a tree topology from Hedges et al. 2015, chronogram 1. Comparison of original chronogram (black) and the chronogram obtained using datelife (gray).

File 26 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S8.tre  
File 26 Description: Newick file of chronogram obtained with datelife, shown in supplementary Figure S8 (gray).

File 27 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S9.jpg  
File 27 Description: JPG file of results of cross validation analysis of datelife’s chronogram generating functions, using a tree topology from Hedges et al. 2015, chronogram 2. Comparison of original chronogram (black) and the chronogram obtained using datelife (gray).

File 28 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S9.tre  
File 28 Description: Newick file of chronogram obtained with datelife, shown in supplementary Figure S9 (gray).

File 29 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S10.jpg  
File 29 Description: JPG file of results of cross validation analysis of datelife’s chronogram generating functions, using a tree topology from Hooper et al. 2017. Comparison of original chronogram (black) and the chronogram obtained using datelife (gray).

File 30 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S10.tre  
File 30 Description: Newick file of chronogram obtained with datelife, shown in supplementary Figure S10 (gray).

File 31 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S11.jpg  
File 31 Description: JPG file of results of cross validation analysis of datelife’s chronogram generating functions, using a tree topology from Jetz et al. 2012, chronogram 1. Comparison of original chronogram (black) and the chronogram obtained using datelife (gray).

File 32 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S11.tre  
File 32 Description: Newick file of chronogram obtained with datelife, shown in supplementary Figure S11 (gray).

File 33 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S12.jpg  
File 33 Description: JPG file of results of cross validation analysis of datelife’s chronogram generating functions, using a tree topology from Jetz et al. 2012, chronogram 2. Comparison of original chronogram (black) and the chronogram obtained using datelife (gray).

File 34 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S12.tre  
File 34 Description: Newick file of chronogram obtained with datelife, shown in supplementary Figure S12 (gray).

File 35 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S13.jpg  
File 35 Description: JPG file of results of cross validation analysis of datelife’s chronogram generating functions, using a tree topology from Kimball et al. 2019, chronogram 1. Comparison of original chronogram (black) and the chronogram obtained using datelife (gray).

File 36 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S13.tre  
File 36 Description: Newick file of chronogram obtained with datelife, shown in supplementary Figure S13 (gray).

File 37 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S14.jpg  
File 37 Description: JPG file of results of cross validation analysis of datelife’s chronogram generating functions, using a tree topology from Kimball et al. 2019, chronogram 2. Comparison of original chronogram (black) and the chronogram obtained using datelife (gray).

File 38 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S14.tre  
File 38 Description: Newick file of chronogram obtained with datelife, shown in supplementary Figure S14 (gray).

File 39 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S15.jpg  
File 39 Description: JPG file of results of cross validation analysis of datelife’s chronogram generating functions, using a tree topology from Oliveros et al. 2019. Comparison of original chronogram (black) and the chronogram obtained using datelife (gray).

File 40 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S15.tre  
File 40 Description: Newick file of chronogram obtained with datelife, shown in supplementary Figure S15 (gray).

File 41 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S16.jpg  
File 41 Description: JPG file of results of cross validation analysis of datelife’s chronogram generating functions, using a tree topology from Price et al. 2014, chronogram 1. Comparison of original chronogram (black) and the chronogram obtained using datelife (gray).

File 42 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S16.tre  
File 42 Description: Newick file of chronogram obtained with datelife, shown in supplementary Figure S16 (gray).

File 43 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S17.jpg  
File 43 Description: JPG file of results of cross validation analysis of datelife’s chronogram generating functions, using a tree topology from Price et al. 2014, chronogram 2. Comparison of original chronogram (black) and the chronogram obtained using datelife (gray).

File 44 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S17.tre  
File 44 Description: Newick file of chronogram obtained with datelife, shown in supplementary Figure S17 (gray).

File 45 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S18.jpg  
File 45 Description: JPG file of results of cross validation analysis of datelife’s chronogram generating functions, using a tree topology from Roquet et al. 2014, chronogram 1. Comparison of original chronogram (black) and the chronogram obtained using datelife (gray).

File 46 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S18.tre  
File 46 Description: Newick file of chronogram obtained with datelife, shown in supplementary Figure S18 (gray).

File 47 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S19.jpg  
File 47 Description: JPG file of results of cross validation analysis of datelife’s chronogram generating functions, using a tree topology from Roquet et al. 2014, chronogram 2. Comparison of original chronogram (black) and the chronogram obtained using datelife (gray).

File 48 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S19.tre  
File 48 Description: Newick file of chronogram obtained with datelife, shown in supplementary Figure S19 (gray).

File 49 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S20.jpg  
File 49 Description: JPG file of results of cross validation analysis of datelife’s chronogram generating functions, using a tree topology from Uyeda et al. 2017. Comparison of original chronogram (black) and the chronogram obtained using datelife (gray).

File 50 Name: Sanchez-Reyes\_etal\_2022\_supplementary\_figure\_S20.tre  
File 50 Description: Newick file of chronogram obtained with datelife, shown in supplementary Figure S20 (gray).

METHODOLOGICAL INFORMATION

All age data and chronograms were obtained using the R package datelife v0.6.5 <https://CRAN.R-project.org/package=datelife> and the chronogram database OpenTreeChronograms v2022.1.28

DATA-SPECIFIC INFORMATION FOR: Sanchez-Reyes\_etal\_2022\_table\_1.csv . Number of variables/columns: 5 . Number of cases/rows: 28 . Variable List Node Name: Node names in tree topology, as defined by taxon A and taxon B taxon A: Scientific name of the first element of a taxon pair that defines a node from a tree topology taxon B: Scientific name of the second element of a taxon pair that defines a node from a tree topology Node Age: Ages of nodes defined by taxon A and taxon B, in Million years Study chronogram: Reference of the study that published the chronogram from which the node ages were extracted . Missing data codes: None . Abbreviations used: NA = not applicable DATA-SPECIFIC INFORMATION FOR: Sanchez-Reyes\_etal\_2022\_table\_2.csv . Number of variables/columns: 9 . Number of cases/rows: 5 . Variable List Node Name: Names of nodes from the tree topology used for the dating analysis Min Age: Minimum node age, in Million years Q1: First quartile of node age distributions, in Million years Median Age: Mean node age, in Million years Mean Age: Median node age, in Million years Q3: Third quartile of node age distributions, in Million years Max Age: Maximum node age, in Million years Variance: Variance of node age, in Million years SD: Standard deviation of node age, in Million years . Missing data codes: None . Abbreviations used: NA = not applicable DATA-SPECIFIC INFORMATION FOR: Sanchez-Reyes\_etal\_2022\_supplementary\_table\_S1.csv . Number of variables/columns: 5 . Number of cases/rows: 818 . Variable List Node Name: Node names in tree topology, as defined by taxon A and taxon B taxon A: Scientific name of the first element of a taxon pair that defines a node from a tree topology taxon B: Scientific name of the second element of a taxon pair that defines a node from a tree topology Node Age: Ages of nodes defined by taxon A and taxon B, in Million years Study reference: Reference of the study that published the chronogram from which the node ages were extracted . Missing data codes: None . Abbreviations used: NA = not applicable DATA-SPECIFIC INFORMATION FOR: Sanchez-Reyes\_etal\_2022\_supplementary\_table\_S2.csv . Number of variables/columns: 9 . Number of cases/rows: 194 . Variable List Node Name: Names of nodes from the tree topology used for the dating analysis Min Age: Minimum node age, in Million years Q1: First quartile of node age distributions, in Million years Median Age: Mean node age, in Million years Mean Age: Median node age, in Million years Q3: Third quartile of node age distributions, in Million years Max Age: Maximum node age, in Million years Variance: Variance of node age, in Million years SD: Standard deviation of node age, in Million years . Missing data codes: None . Abbreviations used: NA = not applicable