

Generating a README.txt file for the Dryad submission of data from study “DateLife: leveraging databases and analytical tools to reveal the dated Tree of Life”

## Description of README file

```
line0 <- paste("This README file was generated on", Sys.Date(), "by Luna L. S  nchez Reyes, <https://orcid.org/0000-0001-7668-2528>")
```

## General Information on data set

```
line1 <- "
```

```
GENERAL INFORMATION
```

```
1. Title of Dataset
```

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

```
2. Author Information
```

```
Corresponding Researcher
```

```
name: Luna L. S  nchez Reyes
```

```
institution: University of California, Merced, USA
```

```
email: s  nchez.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
```

```
3. Date of data collection: 2022-01-28
```

```
4. Geographic location of data collection: Online
```

```
5. Funding sources that supported the collection of the data: National Science Foundation, USA
```

## 6. Recommended citation for this dataset :

Sanchez-Reyes LL, McTavish EJ & Mearns 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

"

## Original file names, rename

```
file.copy(from = "../tables/table-fringillidae-small-example.csv",
          to = "../dryad/Sanchez-Reyes_etal_2022_table_1.csv")
file.copy(from = "../tables/table-fringillidae-small-example-summary.csv",
          to = "../dryad/Sanchez-Reyes_etal_2022_table_2.csv")
file.copy(from = "../tables/table-fringillidae-all-congruified.csv",
          to = "../dryad/Sanchez-Reyes_etal_2022_supplementary_table_S1.csv")
file.copy(from = "../tables/table-fringillidae-all-congruified.pdf",
          to = "../dryad/Sanchez-Reyes_etal_2022_supplementary_table_S1.pdf")
file.copy(from = "../tables/table-fringillidae-all-summary.csv",
          to = "../dryad/Sanchez-Reyes_etal_2022_supplementary_table_S2.csv")
file.copy(from = ".../tables/table-fringillidae-all-summary.pdf",
          to = "../dryad/Sanchez-Reyes_etal_2022_supplementary_table_S2.pdf")
```

## Description of data files

line2 <- "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/phylostatic/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.

### 2. File List:

File1Name: 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\_et al\_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\_et al\_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\_et al\_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\_et al\_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\_et al\_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\_et al\_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\_et al\_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\_et al\_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\_et al\_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\_et al\_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\_et al\_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).

"

## Methods that generated the data files

```
line3 <- "  
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  
"
```

## Individual descriptions of each data file

```
var_description_1 <- c("Node names in tree topology , as defined by taxon A and  
taxon B. ",  
"Scientific name of the first element of a taxon pair  
that defines a node from a tree topology. ",  
"Scientific name of the second element of a taxon pair  
that defines a node from a tree topology. ",
```

```

      "Ages of nodes defined by taxon A and taxon B, in
      Million years.",
      "Reference of the study that published the chronogram
      from which the node ages were extracted.")

var_description_2 <- c("Names of nodes from the tree topology used for the
      dating analysis.",
      "Minimum node age, in Million years.",
      "First quartile of node age distributions, in Million
      years.",
      "Mean node age, in Million years.",
      "Median node age, in Million years.",
      "Third quartile of node age distributions, in Million
      years.",
      "Maximum node age, in Million years.",
      "Variance of node age, in Million years.",
      "Standard deviation of node age, in Million years.")

csv_file_names <- c("Sanchez-Reyes_etal_2022_table_1.csv",
      "Sanchez-Reyes_etal_2022_table_2.csv",
      "Sanchez-Reyes_etal_2022_supplementary_table_S1.csv",
      "Sanchez-Reyes_etal_2022_supplementary_table_S2.csv")
tre_file_names <- c("Sanchez-Reyes_etal_2022_figure_1_chronogram_mock_example.
      tre")

text <- c()
var_descriptions <- list(var_description_1, var_description_2, var_description
      _1, var_description_2)
names(var_descriptions) <- csv_file_names

for (file_name in csv_file_names) {
  # read the csv file in:
  csv_data <- utils::read.csv(file = paste0("../dryad/", file_name), header =
      TRUE, row.names = NULL)
  # remove the column with row numbers:
  csv_data <- csv_data[, -1]
  # create the description text
  var_list <- gsub("[.] ", "_", colnames(csv_data))
  var_list <- paste0("\t", var_list, ":", var_descriptions[[file_name]])
  text <- c(text,
      c(paste("\nDATA-SPECIFIC INFORMATION FOR:", file_name, "\n"),
        paste("1. Number of variables/columns:", ncol(csv_data), "\n"),
        paste("2. Number of cases/rows:", nrow(csv_data), "\n"),
        "3. Variable List:", var_list,
        "\n4. Missing data codes: None",
        "\n5. Abbreviations used: NA = not applicable\n"))
}

line4 <- text

```

## Writing everything to TXT file

```
writeLines(text = c(line0,
```



```
line1 ,  
line2 ,  
line3 ,  
line4), con = ".. /dryad/README_Sanchez-Reyes_etal_2022_  
data.txt")
```