1   DateLife: leveraging databases and analytical tools to reveal the dated Tree of Life

2   Luna L. Sánchez Reyes[1,2], Emily Jane McTavish[1], & Brian O'Meara[2]

3   [1] University of California, Merced

4   [2] University of Tennessee, Knoxville

5   Author Note

6   School of Natural Sciences, University of California, Merced, Science and Engineering

7   Building 1.

8   Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville,

9   425 Hesler Biology Building, Knoxville, TN 37996, USA.

15   Correspondence concerning this article should be addressed to Luna L. Sánchez Reyes, .

16   E-mail: sanchez.reyes.luna@gmail.com

₁₇   Abstract

₁₈   Time of evolutionary origin is fundamental for research in the natural sciences, as well as for

₁₉ education, science communication and policy. Despite an increased availability of fossil and

₂₀ molecular data, and time-efficient analytical techniques, achieving a high-quality

₂₁ reconstruction of time of evolutionary origin as a phylogenetic tree with branch lengths

₂₂ proportional to absolute time (chronogram), is still a difficult and time-consuming task for a

₂₃ majority of interested parties. Yet, the amount of published chronograms has increased

₂₄ significantly in the past two decades, and a non-negligeable proportion of these data have

₂₅ been steadily accumulating in public, open databases such as TreeBASE and Open Tree of

₂₆ Life, exposing a wealth of expertly-curated and peer-reviewed data on time of evolutionary

₂₇ origin in a programatic and reusable way, for a large quantity and diversity of organisms.

₂₈ This trend results from intensive and localized efforts for improving data sharing practices,

₂₉ as well as incentivizing open science in biology. Despite these trends, accessibility for time

₃₀ data is not that good. R has become a widely used element of the biological data analyst

₃₁ toolkit. Hence,to improve accessibility of time data we developed an R package that access

₃₂ these data and eases interaction, reanalysis and reuse of it, incorporation of these data into

₃₃ the evolutionary workflow.

₃₄   Here we present `datelife`, a service implemented as an R package and a web site

₃₅ (www.datelife.org) for increased accessibility, efficient reuse, summary and reanalysis of

₃₆ expert, peer-reviewed, public data on time of evolutionary origin.

₃₇   Main results: 1. blah 2. blah 3. blah

₃₈   Results in a general context: All source and summary chronograms can be saved in

₃₉ formats that permit easy reuse and reanalysis. Summary and newly generated trees are

₄₀ potentially useful to evaluate evolutionary hypothesis in different areas of research in biology.

₄₁ How well this trees work for this purpose still needs to be tested.

⁴² `datelife` will be useful to increase awereness on the existing variation in expert time

⁴³ of divergence data, and might foster exploration of the effect of alternative divergence time

⁴⁴ hypothesis on the results of analyses, nurturing a culture of more cautious interpretation of

⁴⁵ evolutionary results.

⁴⁶ *Keywords:* Tree; Phylogeny; Scaling; Dating; Ages; Divergence times; Open Science;

⁴⁷ Congruification; Supertree; Calibrations

⁴⁸ Word count: 55

DateLife: leveraging databases and analytical tools to reveal the dated Tree of Life

```
## Warning in data(opentree_chronograms): data set 'opentree_chronograms' not found
```

Introduction

Description

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study.

Benchmark

Example

We used R (Version 4.1.0; R Core Team, 2021) and the R-package *papaja* (Version 0.1.0.9997; Aust & Barth, 2020) for all our analyses.

Discussion

Conclusions

Availability

`datelife` is free and open source and it can be used through its current website http://www.datelife.org/query/, through its R package, and through Phylotastic's project web portal http://phylo.cs.nmsu.edu:3000/. `datelife`'s website is maintained using RStudio's shiny server and the shiny package open infrastructure, as well as Docker. `datelife`'s R package stable version will be available for installation from the CRAN repository (https://cran.r-project.org/package=datelife) using the command `install.packages(pkgs = "datelife")` from within R. Development versions are available from the GitHub repository (https://github.com/phylotastic/datelife) and can be installed using the command `devtools::install_github("phylotastic/datelife")`.

Supplementary Material

Code used to generate all versions of this manuscript, the biological examples, as well as the benchmark of functionalities are available at datelifeMS1, datelife_examples, and datelife_benchmark repositories in LLSR's GitHub account.

Funding

Acknowledgements

References

Aust, F., & Barth, M. (2020). *papaja: Create APA manuscripts with R Markdown.* Retrieved from https://github.com/crsh/papaja

Barker, F. K., Burns, K. J., Klicka, J., Lanyon, S. M., & Lovette, I. J. (2012). Going to extremes: Contrasting rates of diversification in a recent radiation of new world passerine birds. *Systematic Biology, 62*(2), 298–320.

Barker, F. K., Burns, K. J., Klicka, J., Lanyon, S. M., & Lovette, I. J. (2015). New insights into new world biogeography: An integrated view from the phylogeny of blackbirds, cardinals, sparrows, tanagers, warblers, and allies. *The Auk: Ornithological Advances, 132*(2), 333–348.

Burns, K. J., Shultz, A. J., Title, P. O., Mason, N. A., Barker, F. K., Klicka, J., . . . Lovette, I. J. (2014). Phylogenetics and diversification of tanagers (passeriformes: Thraupidae), the largest radiation of neotropical songbirds. *Molecular Phylogenetics and Evolution, 75*, 41–77.

Claramunt, S., & Cracraft, J. (2015). A new time tree reveals earth history's imprint on the evolution of modern birds. *Science Advances, 1*(11), e1501005.

Gibb, G. C., England, R., Hartig, G., McLenachan, P. A., Taylor Smith, B. L., McComish, B. J., . . . Penny, D. (2015). New zealand passerines help clarify the diversification of major songbird lineages during the oligocene. *Genome Biology and Evolution, 7*(11), 2983–2995.

Hedges, S. B., Marin, J., Suleski, M., Paymer, M., & Kumar, S. (2015). Tree of life reveals clock-like speciation and diversification. *Molecular Biology and Evolution, 32*(4), 835–845. https://doi.org/10.1093/molbev/msv037

Hooper, D. M., & Price, T. D. (2017). Chromosomal inversion differences correlate with range overlap in passerine birds. *Nature Ecology & Evolution*, *1*(10), 1526.

Jetz, W., Thomas, G., Joy, J. J. B., Hartmann, K., & Mooers, A. (2012). The global diversity of birds in space and time. *Nature*, *491*(7424), 444–448. https://doi.org/10.1038/nature11631

Price, T. D., Hooper, D. M., Buchanan, C. D., Johansson, U. S., Tietze, D. T., Alström, P., . . . others. (2014). Niche filling slows the diversification of himalayan songbirds. *Nature*, *509*(7499), 222.

R Core Team. (2021). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/