

1 Running head: DATELIFE: REVEALING THE DATED TREE OF LIFE

2 Title: DateLife: Leveraging databases and analytical tools to reveal the dated Tree of Life

3 Authors: Luna L. Sánchez-Reyes<sup>1</sup>, Brian C. O'Meara<sup>1</sup>

4 Correspondence address:

5 1. *Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, 425 Hesler Biology*  
6 *Building, Knoxville, TN 37996, USA*

7 Corresponding authors: [sanchez.reyes.luna@gmail.com](mailto:sanchez.reyes.luna@gmail.com), [bomeara@utk.edu](mailto:bomeara@utk.edu)

8 **abstract.-** Here goes the abstract.

9 **Keywords:** Tree; Phylogeny; Scaling; Open; Ages; Congruify; Supertree;

Time of lineage divergence constitutes in many ways a fundamental piece of information for evolutionary understanding. Coupled to species number and distribution, it allows studying the tempo and mode of speciation and extinction), central for understanding of how biodiversity patterns are shaped across space, time and clades (Morlon 2014). In the past two decades, the possibility to obtain good quality DNA sequences coupled to methodological developments in phylogenetic and dating inference, allowed the application of molecular dating methods on a very large amount and diversity of organisms, greatly increasing the quantity and quality of data on taxon ages across the tree of life. To date, there is a large amount of data on taxon ages and phylogenetic relationships in public repositories such as Dryad, TreeBASE and Open Tree of Life (OToL). OToL alone hosts now more than 200 chronograms. Methods to include living and fossil lineages are in continued development and increased usage by the community, which coupled to better sharing data practices, are greatly contributing to the accumulation in number and type of available data on taxon ages.

The TimeTree project (Hedges et al. 2006, 2015; Kumar et al. 2017) has aggregated chronograms from 3,163 studies, encompassing 97,085 species (Kumar et al. 2017), and continues to grow. However, even in this gold standard resource, the included taxa only encompass between 0.097 and 3.236% of total species diversity (following taxonomic expert opinion on the global, extant species numbers, which ranges from 3 to 100 million species [Mayr2010; Moran2011]). One advantage of TimeTree is that it includes taxa from across the tree of life, versus more specialized chronogram repositories focusing on plants (Webb and Donoghue 2005), birds (Jetz et al. 2012), and fish . Users can choose between a web interface or a mobile app to receive information on divergence times for the evolutionary history of a lineage, pairs of taxa, all lineages within a taxon, or a list of taxa. As a science communication tool, TimeTree project is very powerful: it has a friendly graphical interface, with informative and colorful outputs, that allows the general public to satisfy curiosity regarding a particular organism of interest or group of them. It is of limited utility for scientific studies, however. The thousands of trees that have been entered are unavailable for examination or reuse; according to the creators (see TimeTree web FAQ), methods for allowing data downloading have been under discussion for the past several years yet the primary data remain closed. Moreover, there is no Application Programming Interface (API) allowing programmatic access to any data, greatly impairing the possibility of large-scale, automated data-mining, which is not allowed under TimeTree website’s terms of use. The nearly hundred thousand

taxon summary chronogram generated from TimeTree resources is not available with its publication (Kumar et al. 2017) or the TimeTree website, though the still substantial chronogram from a previous publication (Hedges et al. 2015) was made available at OTOL.

Despite its great importance, analytical tools to summarize available information on taxon ages for the scientific community are still lacking. We identified several aspects that might have so far delayed the exploitation of existing data. First, original chronograms available publicly are scattered across repositories and journals supplementary data, usually with different formats too. Second, lineage names due to taxonomic idiosyncrasy can be different among studies and manual curation of that is usually necessary. Third, data curation Recent advances on this area (e.g., supersmart) aim to: Supersmart is an open tool but not easy to use, still requires a lot of curation and knowledge. Generate new dates using all available DNA sequence information; Perform one global analysis using all available information; Problems or downsides: This might be time consuming for large groups and a lot of data curation and knowledge on the group of interest is still necessary. For example, choosing correct fossils for calibration requires a lot of expertise and knowledge on the group. An incorrect use of fossils can generate severe bias in dating results (Sauquet et al. 2012). Hence, data curation is still an important part of any biological study. The research community considers it as an important or even crucial step before data analysis. Hence, automated processes for large data analysis are frequently received with skepticism.

**datelife** palliates this by only using information available from already published studies, which are ideally constructed using robust information, such as sequence data and thoughtfully curated fossil calibrations.

Rapidly increasing data on time of lineage divergence both from molecular and paleontological studies; the increasing importance of use of these data in distant areas of research, often not specialized enough to rapidly obtain data on their own; and the lack of an open (both the data sources and the code underlying the analyses) easy to use tool inspired the development of a prototype **datelife** service over a series of phylotastic hackathons (Stoltzfus et al. 2013) at the National Evolutionary Synthesis Center. In this paper we present the first formal description of **datelife**, featuring an improved database of chronograms, more methods to summarize trees, and new functions to visualize data, as well as comparisons of summary trees.

**datelife** is the main service for scaling phylogenetic trees in Phylotastic! system (Stoltzfus et al. 2013) It can be used through its R package , and a web interface (<http://www.datelife.org/query/>).

## DESCRIPTION

The basic **datelife** workflow is shown in figure 1 and consists of:

- 1) A user providing at least two taxon names as input, either as tip labels on a tree, or as a simple comma separated character string. The tree can be in newick or phylo format, and can be with or without branch lengths.
- 2) **datelife** then performs a search across its database of peer reviewed and curated chronograms; identifies and gets source trees with at least two matching input names; drops unmatching taxa from positively identified source trees; and finally transforms each source tree to a patristic matrix named by the citation of the original study. This format facilitates and greatly speeds up all further analyses and summarization algorithms.
- 3) The user can obtain different types of summaries from the source data including: a) all source chronograms, b) mrca ages of source chronograms, c) citations of studies where source chronograms were originally published, d) a summary table with all of the above, e) a single summary tree of all source chronograms, and f) a report of succesful matches per input taxon name across source chronograms.
- 4) At this point, users can choose to use all or some source data as calibration points to date a tree of their own making or choosing.
- 5) Users can also simulate age and/or phylogenetic data of input taxa not found in the database. A variety of algorithms are available for this purpose.
- 6) Finally, users can easily view results graphically as well as construct their own graphs using inbuilt **datelife** graphic generators.

**datelife**'s chronogram database is currently built from Open Tree of Life (OToL)'s (Hinchliff et al. 2015) tree repository. Among currently existing repositories (e.g., TreeBase, Dryad), OToL's metadata rich tree store is the only one meeting the requirements for proper/accurate automatized handling of trees. Input

taxon names accepted by **datelife** are binomial species names or clades. Taxon searches are performed at the species level, so when input names correspond to higher clades, **datelife** pulls all accepted species names within the clade from OToL’s reference taxonomy to perform the search. Currently, searches at the infrapsecies level are not allowed, so input names belonging to subspecies or any other infraspecific category are treated as species. **datelife** also processes input names with the taxon name resolution service (TNRS), which corrects potentially misspelled names and typos, and standardizes variation in spelling and synonyms (Boyle et al. 2013), increasing the probability to correctly find the queried taxa in **datelife**’s chronogram database.

Source chronogram summary tree can be assembled using the Super Distance Matrix (SDM) supertree construction approach (Criscuolo et al. 2006) or using the median of branch lengths and the hierarchical clustering method. Tree dating and simulation options are performed with various algorithms: Branch Length Adjuster (BLADJ) is a simple algorithm to distribute ages of undated nodes evenly, which minimizes age variance in the chronogram (Webb et al. 2008). PATHd8 is a non-clock, rate-smoothing method (Britton et al. 2007) to date trees. treePL, is a semi-parametric, rate-smoothing, penalized likelihood dating method (Smith and O’Meara 2012). MrBayes (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003) can be used when adding taxa at random, following a reference taxonomy or a topological constraint. It draws ages from a pure birth model, as implemented by Jetz and collaborators (2012). To apply calibrations to a tree, the eongruification algorithm described in (Eastman et al. 2013) is used to find shared nodes between trees (congruent nodes).

To gather, process, and present information, **datelife** builds up from functions available in several R packages including rotl (Michonneau et al. 2016), ape (Paradis et al. 2004), geiger (Harmon et al. 2008), paleotree (Bapst 2012), bold (Chamberlain 2018), phytools (Revell 2012), taxize (Chamberlain and Szöcs 2013; Chamberlain 2018), phyloch (Heibl 2008), phylocomr (Ooms and Chamberlain 2018) and rphylotastic (O’Meara et al. 2019).

## BENCHMARK

**datelife**'s code speed was tested on an Apple iMac with one 3.4 GHz Intel Core i5 processor. We registered variation in computing time of query processing and search through the database relative to number of queried taxon names. Query processing increases roughly linearly with number of input taxon names, and increases considerably if TNRS service is activated. Up to ten thousand names can be processed and searched in less than 30 minutes. A name search through the chronogram database with an already processed query can be performed in less than a minute, even with a very large number of taxon names (Fig. 2).

**datelife**'s code performance was evaluated with a set of unit tests designed and implemented with the R package `testthat` (R Core Team 2018). These tests were run both locally –using the `devtools` package (R Core Team 2018)– and on a public server –via GitHub– using the continuous integration tool Travis CI (<https://travis-ci.org>). At present, unit tests cover more than 50% of **datelife**'s code (<https://codecov.io/gh/phylostatic/datelife>).

## EXAMPLE

In this section we demonstrate the types of outputs that can be obtained with **datelife**, using as an example the bird family Fringillidae of true finches. We performed a higher-taxon search to obtain all data on lineage divergence available from **datelife**'s database for all recognised species within the Fringillidae (475 spp. according to the Open Tree of Life taxonomy). There are 13 chronograms containing at least two Fringillidae species, published in 9 different studies (Fig. 3). Data from these source chronograms was used to generate two types of summary chronograms, median and SDM. As explained in the **Description**, data from source chronograms was first summarised into a single distance matrix (using either the median or the SDM method) and then the available node ages were used as calibrations points over a consensus tree topology, to obtain a dated tree with the program BLADJ (Fig. 4). Median summary chronograms are older and have wider variation in maximum ages than chronograms obtained with SDM. In both cases, ages are coherent with source ages. It is not certain if these chronograms can be used to perform downstream evolutionary analyses. There is currently wide interest in determining this. However, it is certain that these chronograms are useful for...

Data from source chronograms was also used to date tree topologies with no branch length information and trees with branch lengths in relative substitution rates (Figs. 5 and 6). As a form of cross validation, we used tree topologies from each study and calibrated them using information from all other source chronograms. In the absence of branch length data, the ages of internal nodes were approximately recovered in almost all cases (except for studies 3, and 5; Fig. 5). Maximum tree ages were only approximately recovered in one case (study 2; Fig. 5). Branch lengths were successfully generated using the BOLD database for all source chronograms. However, dating with PATHd8 (using congruified calibrations) was only successful in three cases (studies 3, 5, and 9; Fig. 6). From these, two trees have a different sampling than the original source chronogram, mainly because DNA data for some species is absent from the BOLD. Maximum ages are quite different from source chronograms, but this might be explained also by the differences in sampling between source chronograms and BOLD trees. More examples and details can be consulted in [https://github.com/LunaSare/datelife\\_examples](https://github.com/LunaSare/datelife_examples).

## CONCLUSIONS

Taxon ages are key to many areas of evolutionary studies: trait evolution, species diversification, biogeography, macroecology and more. Obtaining these ages is difficult, especially for those who want to use phylogenies but who are not systematists, or do not have the time to develop the necessary knowledge and data curation skills to produce new chronograms. Knowledge on taxon ages is also important for non-biological studies and the non-academic community. The combination of new analytical techniques, availability of more fossil and molecular data, and better practices in data sharing has resulted in a steady accumulation of chronograms in public and open databases such as Dryad, TreeBASE or Open Tree of Life, for a large quantity and diversity of organisms. However, this information remains difficult to synthesize for many biologists and the non-academic community.

Here, we have shown that **datelife** allows an easy and fast obtention of all publicly available information on taxon ages, which can be used to generate new data. This information can be used to account for the effect of phylogenetic signal in studies of trait evolution; to explore potential speciation and extinction dynamics of interest within a clade; to obtain a time frame of biogeographical events; for science communication and



outreach, amongst others. Compared to similar platforms such as time tree of life and supermart, it offers several advantages. It is fast; source data is completely open; it requires no expert biological knowledge from users for any of its functionalities; it allows exploration of alternative taxonomic and phylogenetic schemes; it allows rapid exploration of the effect of alternative divergence time hypothesis; it allows rapid synthesis in a number of different formats; it facilitates reproducibility of analyses;

Improvements, short and long-term: \* fossils as calibrations: Using secondary calibrations can generate biased ages when using bayesian methods, mainly because we don't know what prior to give to secondary calibrations (Schenk 2016). \* bayesian congruification \* topological congruification

Problems and caveats: Not many databases, only OTOL Why TreeBase is not very useful for us? Be precise. Are these chronograms reliable to study evolutionary patterns, such as species diversification? **datelife** can be seen as an open resource to know the current state of knowledge on lineage divergence times. Whether chronograms obtained using this original data can be used reliably to study complicated patterns of evolution is still uncertain. If all, la facilidad para obtener hipotesis de tiempo de divergencia nos ayudará a evaluar la capacidad de los cronogramas para estudiar otros fenomenos evolutivos. Por ahora, no podemos aseverar que estos cronogramas puedan usarse para todo tipo de analisis.

#### 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 by RStudio's shiny server and the shiny package open infrastructure, as well as Docker. **datelife**'s R package stable version is 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 software benchmark can be found in GitHub repositories at [https://github.com/LunaSare/datelife\\_paper1](https://github.com/LunaSare/datelife_paper1), [https://github.com/LunaSare/datelife\\_examples](https://github.com/LunaSare/datelife_examples), and [https://github.com/LunaSare/datelife\\_benchmark](https://github.com/LunaSare/datelife_benchmark), respectively.

## FUNDING

Funding was provided by NSF grant 1458603

NESCent

Open Tree of Life

University of Tennessee, Knoxville

## ACKNOWLEDGEMENTS

We thank colleagues (students and postdocs) at the O’Meara Lab at the University of Tennessee Knoxville for suggestions, discussions and software testing. The late National Evolutionary Synthesis Center (NESCent), which sponsored hackathons that led to initial work on this project. The Open Tree of Life project that provides the open, metadata rich repository of trees used for **datelife**. The many scientists who publish their chronograms in an open, reusable form, and the scientists who curate them for deposition in OpenTree. The US National Science Foundation (NSF) for funding nearly all the above, in addition to the ABI grant that funded this project itself.

- 205 Bapst D.W. 2012. Paleotree: An R package for paleontological and phylogenetic analyses of evolution.  
206 *Methods in Ecology and Evolution*. 3:803–807.
- 207 Barker F.K., Burns K.J., Klicka J., Lanyon S.M., Lovette I.J. 2012. Going to extremes: Contrasting rates of  
208 diversification in a recent radiation of new world passerine birds. *Systematic biology*. 62:298–320.
- 209 Barker F.K., Burns K.J., Klicka J., Lanyon S.M., Lovette I.J. 2015. New insights into new world biogeography:  
210 An integrated view from the phylogeny of blackbirds, cardinals, sparrows, tanagers, warblers, and allies. *The*  
211 *Auk: Ornithological Advances*. 132:333–348.
- 212 Boyle B., Hopkins N., Lu Z., Raygoza Garay J.A., Mozzherin D., Rees T., Matasci N., Narro M.L., Piel  
213 W.H., McKay S.J., Lowry S., Freeland C., Peet R.K., Enquist B.J. 2013. The taxonomic name resolution  
214 service: An online tool for automated standardization of plant names. *BMC Bioinformatics*. 14.
- 215 Britton T., Anderson C.L., Jacquet D., Lundqvist S., Bremer K. 2007. Estimating Divergence Times in Large  
216 Phylogenetic Trees. *Systematic Biology*. 56:741–752.
- 217 Burns K.J., Shultz A.J., Title P.O., Mason N.A., Barker F.K., Klicka J., Lanyon S.M., Lovette I.J. 2014.  
218 Phylogenetics and diversification of tanagers (passeriformes: Thraupidae), the largest radiation of neotropical  
219 songbirds. *Molecular Phylogenetics and Evolution*. 75:41–77.
- 220 Chamberlain S. 2018. bold: Interface to Bold Systems API..
- 221 Chamberlain S.A., Szöcs E. 2013. taxize : taxonomic search and retrieval in R [version 2; referees: 3 approved].  
222 *F1000Research*. 2:1–29.
- 223 Claramunt S., Cracraft J. 2015. A new time tree reveals earth history’s imprint on the evolution of modern  
224 birds. *Science advances*. 1:e1501005.
- 225 Criscuolo A., Berry V., Douzery E.J., Gascuel O. 2006. SDM: A fast distance-based approach for (super)tree

226 building in phylogenomics. *Systematic Biology*. 55:740–755.

227 Eastman J.M., Harmon L.J., Tank D.C. 2013. Congruification: Support for time scaling large phylogenetic  
228 trees. *Methods in Ecology and Evolution*. 4:688–691.

229 Gibb G.C., England R., Hartig G., McLenachan P.A., Taylor Smith B.L., McComish B.J., Cooper A., Penny  
230 D. 2015. New zealand passerines help clarify the diversification of major songbird lineages during the oligocene.  
231 *Genome biology and evolution*. 7:2983–2995.

232 Harmon L., Weir J., Brock C., Glor R., Challenger W. 2008. GEIGER: investigating evolutionary radiations.  
233 *Bioinformatics*. 24:129–131.

234 Hedges S.B., Dudley J., Kumar S. 2006. TimeTree: A public knowledge-base of divergence times among  
235 organisms. *Bioinformatics*. 22:2971–2972.

236 Hedges S.B., Marin J., Suleski M., Paymer M., Kumar S. 2015. Tree of life reveals clock-like speciation and  
237 diversification. *Molecular Biology and Evolution*. 32:835–845.

238 Heibl C. 2008. PHYLOCH: R language tree plotting tools and interfaces to diverse phylogenetic software  
239 packages..

240 Hinchliff C.E., Smith S.A., Allman J.F., Burleigh J.G., Chaudhary R., Coghill L.M., Crandall K.A., Deng J.,  
241 Drew B.T., Gazis R., Gude K., Hibbett D.S., Katz L.A., Laughinghouse H.D., McTavish E.J., Midford P.E.,  
242 Owen C.L., Ree R.H., Rees J.A., Soltis D.E., Williams T., Cranston K.A. 2015. Synthesis of phylogeny and  
243 taxonomy into a comprehensive tree of life. *Proceedings of the National Academy of Sciences*. 112:12764–12769.

244 Hooper D.M., Price T.D. 2017. Chromosomal inversion differences correlate with range overlap in passerine  
245 birds. *Nature ecology & evolution*. 1:1526.

246 Huelsenbeck J.P., Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*.  
247 17:754–755.

248 Jetz W., Thomas G., Joy J.J., Hartmann K., Mooers A. 2012. The global diversity of birds in space and  
249 time. *Nature*. 491:444–448.

250 Kumar S., Stecher G., Suleski M., Hedges S.B. 2017. TimeTree: A Resource for Timelines, Timetrees, and  
251 Divergence Times. *Molecular biology and evolution*. 34:1812–1819.

252 Michonneau F., Brown J.W., Winter D.J. 2016. rotl: an R package to interact with the Open Tree of Life  
253 data. *Methods in Ecology and Evolution*. 7:1476–1481.

254 Morlon H. 2014. Phylogenetic approaches for studying diversification. *Ecology Letters*. 17:508–525.

255 O’Meara B., Md Tayeen A.S., Sanchez Reyes L.L. 2019. Rphylotastic: An r interface to ‘phylotastic’ web  
256 services..

257 Ooms J., Chamberlain S. 2018. Phylocomr: Interface to ‘phylocom’..

258 Paradis E., Claude J., Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R language.  
259 *Bioinformatics*. 20:289–290.

260 Price T.D., Hooper D.M., Buchanan C.D., Johansson U.S., Tietze D.T., Alström P., Olsson U., Ghosh-Harihar  
261 M., Ishtiaq F., Gupta S.K., others. 2014. Niche filling slows the diversification of himalayan songbirds.  
262 *Nature*. 509:222.

263 R Core Team. 2018. R: a language and environment for statistical computing. Vienna, Austria: R Foundation  
264 for Statistical Computing.

265 Revell L.J. 2012. Phytools: An r package for phylogenetic comparative biology (and other things). *Methods*  
266 *in Ecology and Evolution*. 3:217–223.

267 Ronquist F., Huelsenbeck J.P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models.  
268 *Bioinformatics*. 19:1572–1574.

269 Sauquet H., Ho S.Y.W., Gandolfo M. a, Jordan G.J., Wilf P., Cantrill D.J., Bayly M.J., Bromham L., Brown

270 G.K., Carpenter R.J., Lee D.M., Murphy D.J., Sniderman J.M.K., Udovicic F. 2012. Testing the impact  
271 of calibration on molecular divergence times using a fossil-rich group: the case of Nothofagus (Fagales).  
272 Systematic Biology. 61:289–313.

273 Schenk J.J. 2016. Consequences of secondary calibrations on divergence time estimates. PLoS ONE. 11.

274 Smith S.A., O’Meara B.C. 2012. TreePL: Divergence time estimation using penalized likelihood for large  
275 phylogenies. Bioinformatics. 28:2689–2690.

276 Stoltzfus A., Lapp H., Matasci N., Deus H., Sidlauskas B., Zmasek C.M., Vaidya G., Pontelli E., Cranston  
277 K., Vos R., Webb C.O., Harmon L.J., Pirrung M., O’Meara B., Pennell M.W., Mirarab S., Rosenberg M.S.,  
278 Balhoff J.P., Bik H.M., Heath T.A., Midford P.E., Brown J.W., McTavish E.J., Sukumaran J., Westneat M.,  
279 Alfaro M.E., Steele A., Jordan G. 2013. Phylotastic! Making tree-of-life knowledge accessible, reusable and  
280 convenient. BMC Bioinformatics. 14.

281 Webb C.O., Ackerly D.D., Kembel S.W. 2008. Phylocom: Software for the analysis of phylogenetic community  
282 structure and trait evolution. Bioinformatics. 24:2098–2100.

283 Webb C.O., Donoghue M.J. 2005. Phylomatic: Tree assembly for applied phylogenetics. Molecular Ecology  
284 Notes. 5:181–183.

FIGURE 1

Stylized DateLife workflow. This shows the general workflows and analyses that can be performed with **datelife**, via the R package or through the website. Details on the functions involved on each workflow are shown in **datelife**'s R package vignette.

FIGURE 2

Computation time of query processing and search across **datelife**'s chronogram database relative to number of input taxon names. We sampled N names from the class Aves for each cohort 100 times and then performed a search with query processing not using the Taxon Names Resoulution Service (TNRS; dark gray), and using TNRS (light gray). We also performed a search using the already processed query for comparison (light blue).

FIGURE 3

Lineage through time (LTT) plots of source chronograms containing all or a subset of species from the bird family Fringillidae of true finches. Arrows indicate maximum age of each chronogram. Numbers reference to chronograms' original publications 1: Barker et al. (2012), 2: Barker et al. (2015), 3: Burns et al. (2014), 4: Claramunt and Cracraft (2015), 5: Gibb et al. (2015), 6: Hedges et al. (2015), 7: Hooper and Price (2017), 8: Jetz et al. (2012), 9: Price et al. (2014).

FIGURE 4

LTT plots of median and Supermatrix Distance Method (SDM) chronograms summarising information from source chronograms found for the Fringillidae. Arrows indicate maximum age.

FIGURE 5

LTT plots showing results from the cross-validation analyses of trees without branch lengths dated using BLADJ. The dating analysis can only be performed in trees with more than 2 tips, thus excluding chronogram from study 4; its data was still used as calibration for the other source chronograms.

308 LTT plots showing results from the cross-validation analyses of trees with branch length reconstructed with  
309 data from the Barcode of Life Database (BOLD) dated using PATHd8. We could construct a tree with  
310 branch lengths for all source chronograms. However, dating with PATHd8 was only successful in three source  
311 chronograms shown here.



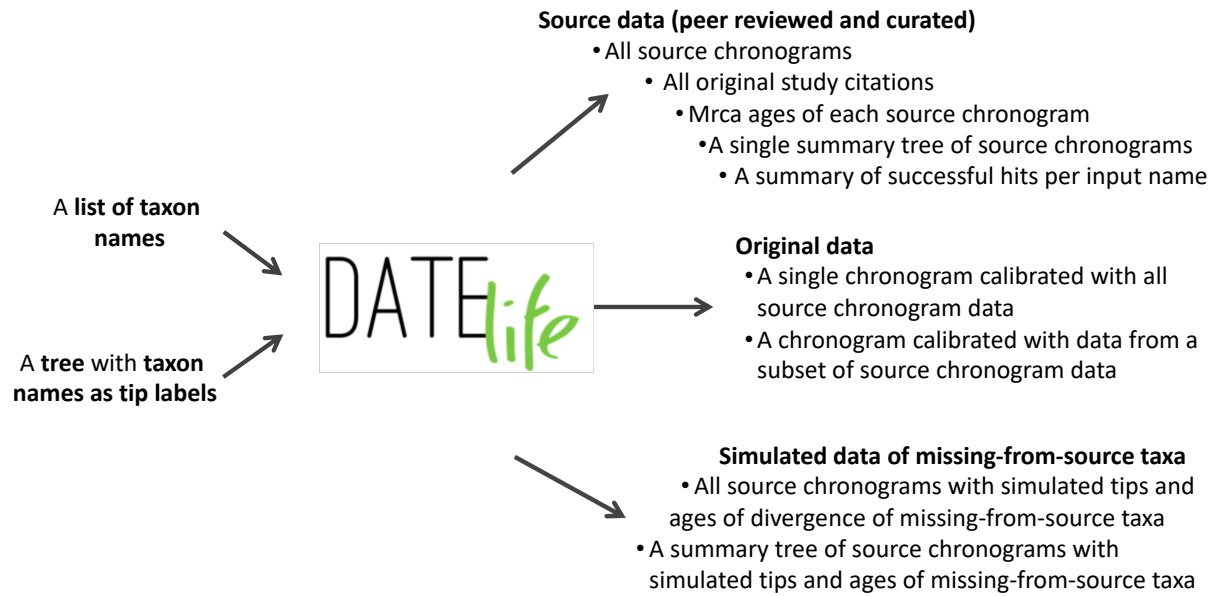


Figure 1:

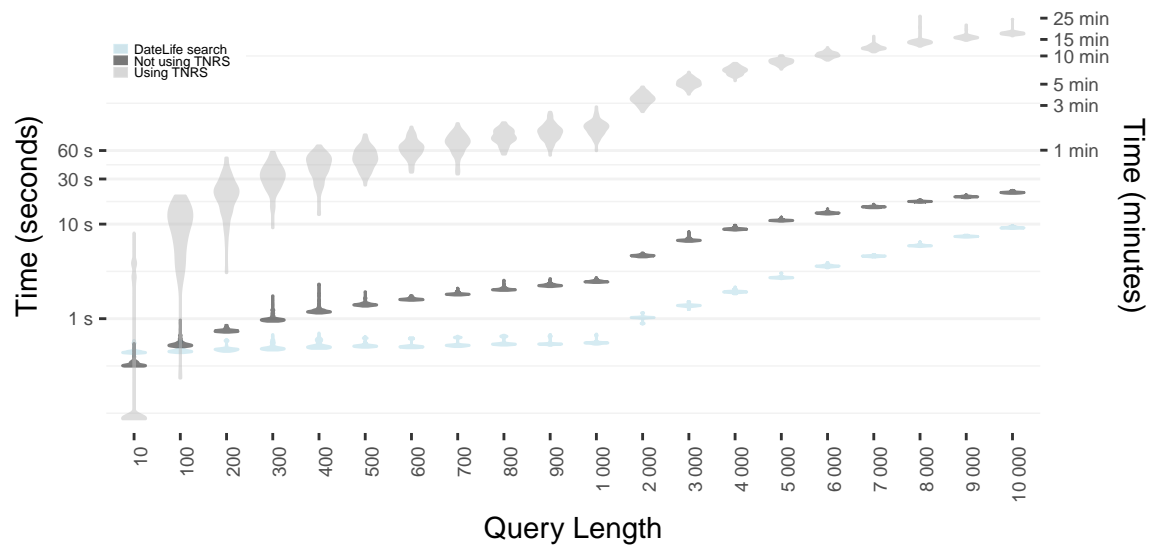


Figure 2:

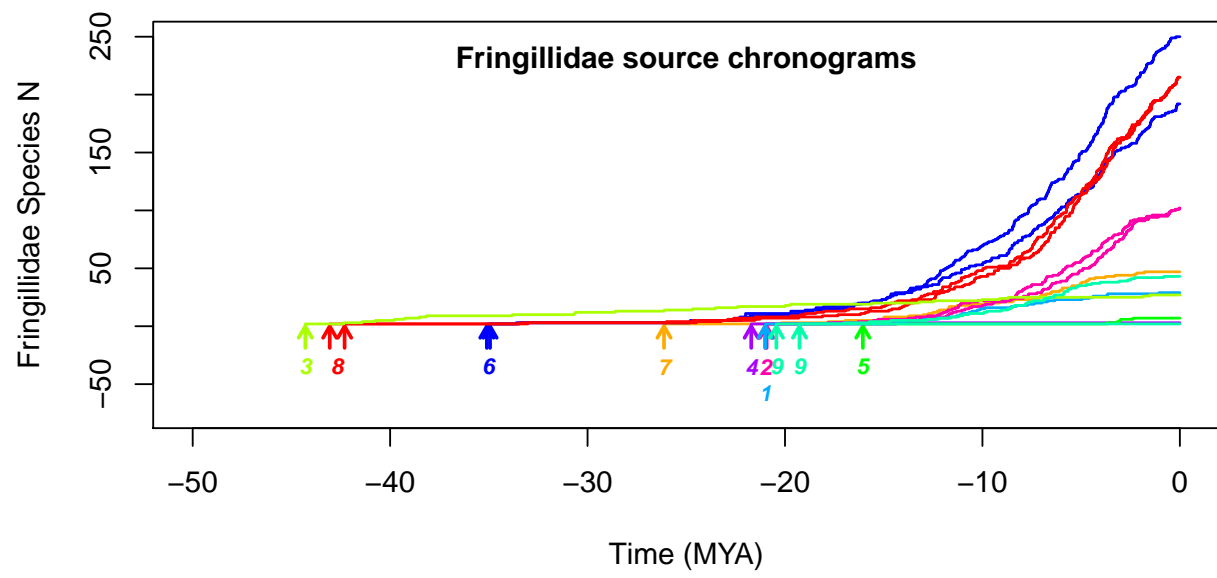


Figure 3:

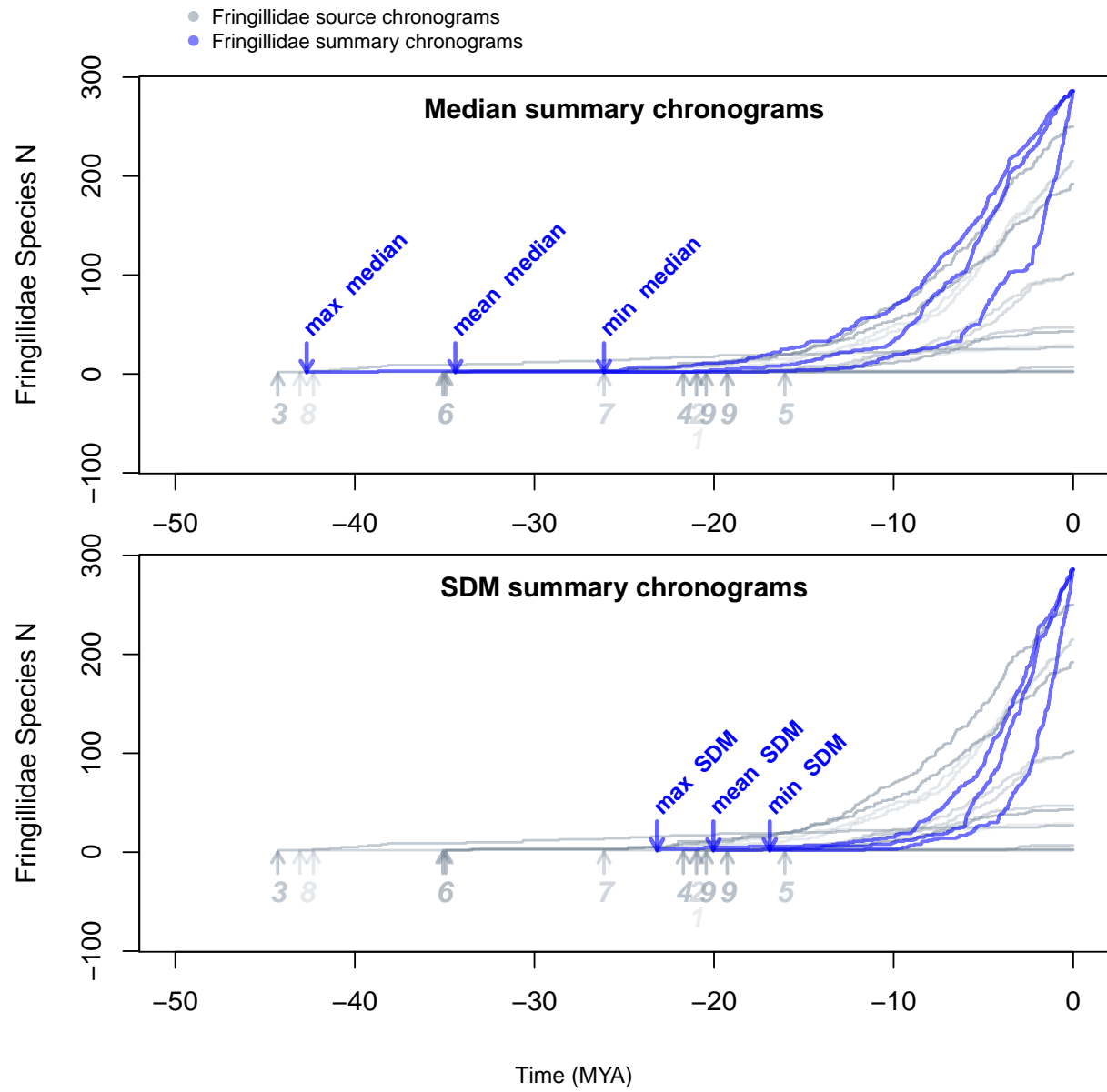


Figure 4:

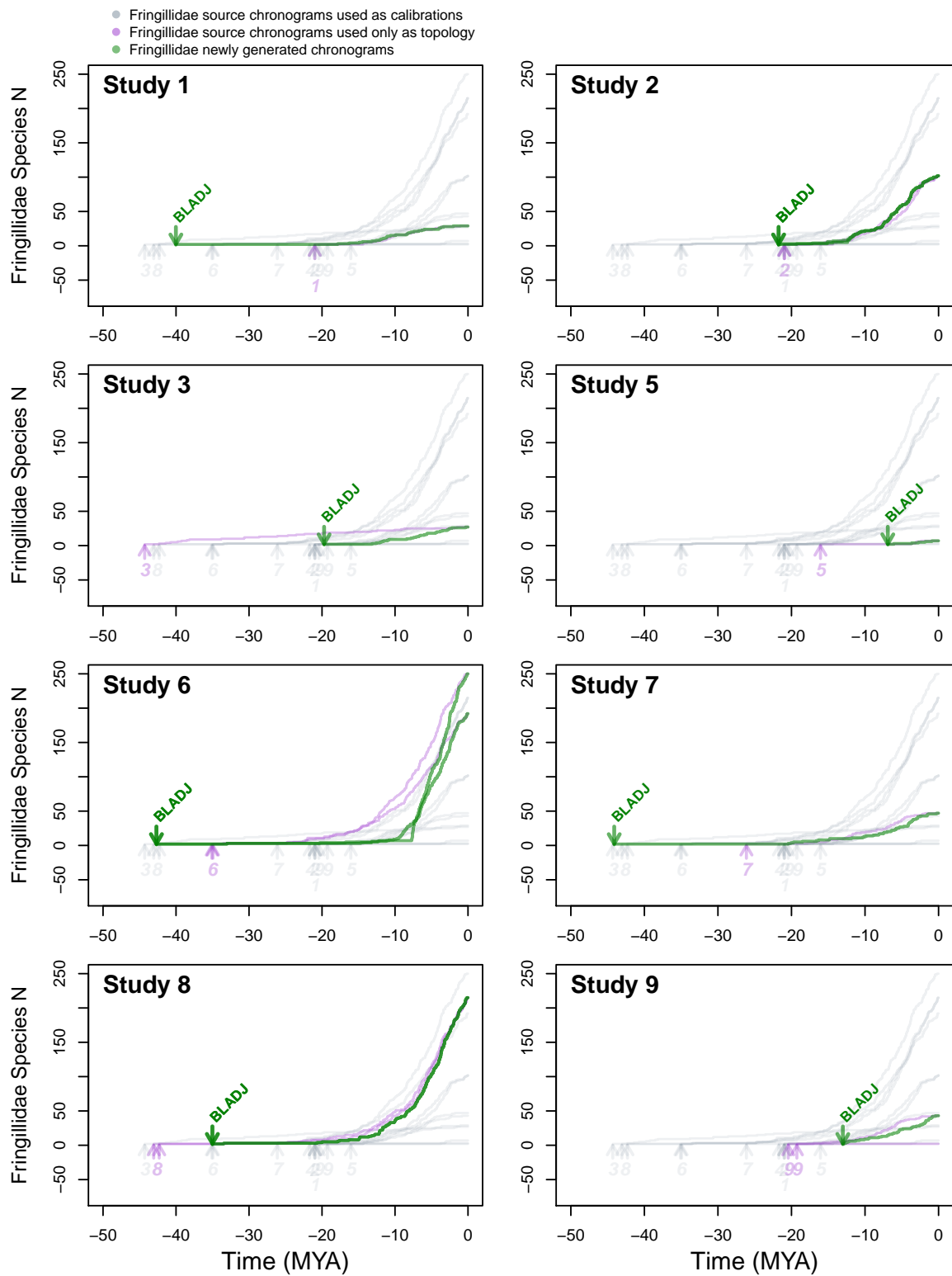


Figure 5:

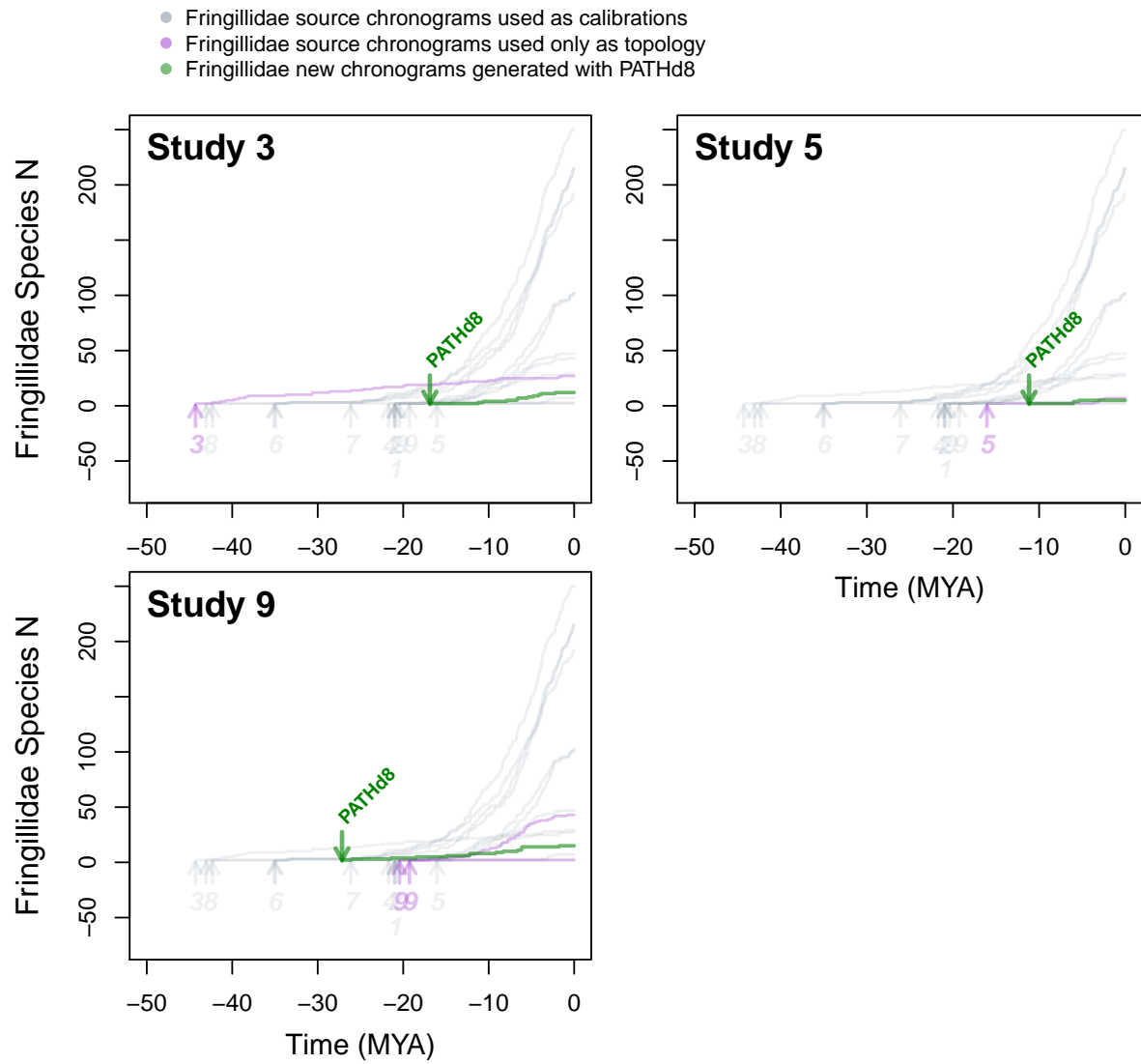


Figure 6: