

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

² Luna L. Sánchez Reyes^{1,2}, Emily Jane McTavish¹, & Brian O'Meara²

³ ¹ University of California, Merced, USA

⁴ ² University of Tennessee, Knoxville, USA

⁵ Author Note

⁶ School of Natural Sciences, University of California, Merced, 258 Science and
⁷ Engineering Building 1, Merced, CA 95340, USA.

⁸ Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville,
⁹ 446 Hesler Biology Building, Knoxville, TN 37996, USA.

¹⁰ The authors made the following contributions. Luna L. Sánchez Reyes: Data curation,
¹¹ Investigation, Software, Visualization, Validation, Writing - Original Draft Preparation,
¹² Writing - Review & Editing; Emily Jane McTavish: Resources, Software, Writing - Review &
¹³ Editing; Brian O'Meara: Conceptualization, Funding acquisition, Methodology, Resources,
¹⁴ Software, Supervision, Writing - Review & Editing.

¹⁵ Correspondence concerning this article should be addressed to Luna L. Sánchez Reyes, .
¹⁶ E-mail: sanchez.reyes.luna@gmail.com

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

18 **Abstract**

19 Chronograms –phylogenies with branch lengths proportional to time– represent key
20 data on evolutionary time frame for the study of natural processes in many areas of biological
21 research. Chronograms also provide valuable information that can be used for education,
22 science communication, and policy decisions. Yet, achieving a high-quality reconstruction of
23 a chronogram is a difficult and resource-consuming task. Here we present DateLife, a service
24 implemented as an R package and an R Shiny web application available at www.datelife.org,
25 that provides services for efficient and easy discovery, summary, reuse, and reanalysis of node
26 age data mined from a curated database of expert, peer-reviewed, and openly available
27 chronograms. The main DateLife workflow starts with one or more scientific taxon names
28 provided by a user. Names are processed and standardized to a unified taxonomy. Then,
29 DateLife runs a name match across its local chronogram database that is curated from Open
30 Tree of Life’s phylogenetic repository, and extracts all chronograms that contain at least two
31 queried taxon names, along with their metadata. Finally, node ages from matching
32 chronograms are mapped upon a chosen tree topology using the congruification algorithm.
33 Congruified node ages are used as secondary calibrations to date the chosen topology, with or
34 without initial branch lengths, using different phylogenetic dating methods such as BLADJ,
35 treePL, PATHd8 and MrBayes. We performed a cross-validation test to compare node ages
36 resulting from a DateLife analysis (i.e, phylogenetic dating using secondary calibrations) to
37 those from the original chronograms (i.e, obtained with primary calibrations), and found
38 that DateLife’s node age estimates appear generally correlated to those from the original
39 chronograms, with the largest variation in ages occurring around topologically deeper nodes.
40 Software such as DateLife can help increase awareness of the existing variation in alternative
41 hypothesis of evolutionary time frame for the same organisms, and can support exploration
42 of the effect of alternative chronogram hypotheses on downstream analyses, providing a

⁴³ framework for a more informed interpretation of evolutionary results.

⁴⁴ *Keywords:* Tree; Phylogeny; Scaling; Dating; Ages; Divergence times; Open Science;

⁴⁵ Congruification; Supertree; Calibrations; Secondary calibrations.

⁴⁶ Word count: 6111

47 Chronograms—phylogenies with branch lengths proportional to time—provide key data
48 on evolutionary time frame for the study of natural processes in many areas of biological
49 research, such as comparative analysis (Freckleton, Harvey, & Pagel, 2002; Harvey, Pagel, &
50 others, 1991), developmental biology (Delsuc et al., 2018; Laubichler & Maienschein, 2009),
51 conservation biology and ecology (Felsenstein, 1985; Webb, 2000), historical biogeography
52 (Posadas, Crisci, & Katinas, 2006), and species diversification (Magallon & Sanderson, 2001;
53 Morlon, 2014).

Building a chronogram is not an easy task. It requires obtaining and curating a homology hypothesis to construct a phylogeny, selecting and placing appropriate calibrations on the phylogeny using independent age data points from the fossil record or other dated events, and inferring a full dated tree. All of this entails specialized biological training, taxonomic domain knowledge, and a non-negligible amount of research time, computational resources and funding.

60 Here we present the DateLife project which has the main goal of extracting and
61 exposing age data from published chronograms, making age data readily accessible to the
62 wider community for reuse and reanalysis, research, teaching, and science communication
63 and policy. DateLife’s core software application is available as an R package (Sanchez-Reyes
64 et al., 2022), and as an online Rshiny interactive website at www.datelife.org. It features key
65 elements for scientific reproducibility, such as a curated, versioned, open and fully public
66 chronogram database (McTavish et al., 2015) that stores data in a computer-readable format
67 (Vos et al., 2012); automated and programmatic ways of accessing and downloading the data
68 in a computer-readable format also (Stoltzfus et al., 2013); and methods to summarize and
69 compare the data.

DESCRIPTION

⁷¹ DateLife's core software applications are implemented in the R package `datelife`. Its

72 current stable version – v0.6.6, is available from The Comprehensive R Archive Network
73 (CRAN) repository (Sanchez-Reyes et al., 2022), and relies on functionalities from other
74 biological R packages: ape (Paradis, Claude, & Strimmer, 2004), bold (Chamberlain, 2018),
75 geiger (Pennell et al., 2014), msa (Bodenhofer, Bonatesta, Horejš-Kainrath, & Hochreiter,
76 2015), paleotree (Bapst, 2012), phyloch (Heibl, 2008), phylocomr (Ooms & Chamberlain,
77 2018), phytools (Revell, 2012), rotl (Michonneau, Brown, & Winter, 2016), and taxize
78 (Chamberlain, 2018; Chamberlain & Szöcs, 2013). Figure 1 provides a graphical summary of
79 the three main steps of the DateLife workflow: creating a search query, searching a database,
80 and summarizing results from the search.

81 *Creating a Search Query*

82 DateLife starts by processing an input consisting of at least one taxon scientific name.
83 Multiple input names can be provided as a comma separated character string or as tip labels
84 on a tree. If the input is a tree, it can be provided as a classic newick character string
85 (Archie et al., 1986), or as a “phylo” R object (Paradis et al., 2004). The input tree is not
86 required to have branch lengths, and its topology is used in the summary steps described in
87 the next section.

88 DateLife processes input scientific names using a Taxonomic Name Resolution Service
89 (TNRS), which increases the probability of correctly finding the queried taxon names in the
90 chronogram database. TNRS detects, corrects and standardizes name misspellings and typos,
91 variant spellings and authorities, and nomenclatural synonyms to a single taxonomic
92 standard (Boyle et al., 2013). TNRS also allows to correctly choose between homonyms, by
93 considering other taxa provided as input to infer the taxonomic context of the homonym.
94 DateLife implements TNRS using the Open Tree of Life (OpenTree) unified Taxonomy
95 (OTT, Open Tree Of Life et al., 2016; Rees & Cranston, 2017) as standard, storing
96 taxonomic identification numbers (OTT ids) for further processing and analysis. Other
97 taxonomies currently supported by DateLife are the National Center of Biotechnology

98 Information (NCBI) taxonomic database (Schoch et al., 2020), the Global Biodiversity
99 Information Facility (GBIF) taxonomic backbone (GBIF Secretariat, 2022), and the Interim
100 Register of Marine and Non-marine Genera (IRMNG) database (Rees et al., 2017).

101 Besides binomial species names, DateLife accepts scientific names from any inclusive
102 taxonomic group (e.g., genus, family, tribe), as well as subspecific taxonomic variants (e.g.,
103 subspecies, variants, strains). If a taxon name belongs to an inclusive taxonomic group,
104 DateLife has two alternative behaviors defined by the “get species from taxon” flag. If the
105 flag is active, DateLife retrieves all species names within the taxonomic group from the
106 standard taxonomy of choice, and adds them to the search query. In this case, subspecific
107 variants are excluded. If the flag is inactive, DateLife excludes any taxon names above the
108 species level from the search query. Species and subspecific variant names are processed and
109 searched as provided by the user. The processed taxon names are saved as an R object of a
110 newly defined class, `datelifeQuery`, that is used in the following steps. This object contains
111 the input names standardized to a taxonomy of choice (OTT by default), the corresponding
112 OTT id numbers, and the topology of an input tree, if one was provided.

113 *Searching a Chronogram Database*

114 At the time of writing of this manuscript (Feb 21, 2023), DateLife’s chronogram
115 database latest version consist of 253 chronograms published in 187 different studies. It is
116 curated from OpenTree’s phylogenetic database, the Phylesystem, which constitutes an open
117 source of expert and peer-reviewed phylogenetic knowledge with rich metadata (McTavish et
118 al., 2015), which allows automatic and reproducible assembly of our chronogram database.
119 Datelife’s chronogram database is navigable as an R data object within the `datelife` R
120 package.

121 A unique feature of the Phylesystem is that any user can add new published,
122 state-of-the-art chronograms any time, through their curator application

123 (<https://tree.opentreeoflife.org/curator>). As chronograms are added to Phylesystem, they
124 can be incorporated into the chronogram database of the **datelife** R package, which is
125 currently manually updated as new chronogram data is added to Phylesystem. The updated
126 database is assigned a new version number, followed by a package release on CRAN. Users
127 can directly implement **datelife** functions to trigger an update of the local chronogram
128 database, to incorporate any new chronograms to their DateLife analysis before an official
129 database update is released on CRAN.

130 A DateLife search is implemented by matching processed taxon names provided by the
131 user to tip labels in the chronogram database. Chronograms with at least two matching
132 taxon names on their tip labels are identified and pruned down to preserve only the matched
133 taxa. These matching pruned chronograms are referred to as source chronograms. Total
134 distance (in units of millions of years) between taxon pairs within each source chronogram
135 are stored as a patristic distance matrix (Figure 1). The matrix format speeds up extraction
136 of pairwise taxon ages of any queried taxa, as opposed to searching the ancestor node of a
137 pair of taxa in a “phylo” object or newick string. Finally, the patristic matrices are
138 associated to the study citation where the original chronogram was published, and stored as
139 an R object of the newly defined class **datelifeResult**.

140 *Summarizing Search Results*

141 Summary information is extracted from the **datelifeResult** object to inform
142 decisions for subsequent steps in the analysis workflow. Basic summary information available
143 to the user is:

- 144 1. The matching pruned chronograms as newick strings or “phylo” objects.
145 2. The ages of the root of all source chronograms. These ages can correspond to the age
146 of the most recent common ancestor (mrca) of the user’s group of interest if the source
147 chronograms have all taxa belonging to the group. If not, the root corresponds to the

148 mrca of a subgroup withing the group of interest.

149 3. Study citations where original chronograms were published.

150 4. A report of input taxon names matches across source chronograms.

151 5. The source chronogram(s) with the most input taxon names.

152 6. Various single summary chronograms resulting from summarizing age data, generated

153 using the methodology described next.

154 *Choosing a Topology*

155 DateLife requires a tree topology to summarize age data upon. We recommend that

156 users provide as input a tree topology from the literature, or one of their own making. If no

157 topology is provided, DateLife automatically extracts one from the OpenTree synthetic tree,

158 a phylogeny currently encompassing 2.3 million taxa across all life, assembled from 1, 239

159 published phylogenetic trees and OpenTree’s unified Taxonomy, OTT (Open Tree Of Life et

160 al., 2019). Alternatively, DateLife can combine topologies from source chronograms using a

161 supertree approach (Criscuolo, Berry, Douzery, & Gascuel, 2006). To do this, DateLife first

162 identifies the source chronograms that form a grove, roughly, a sufficiently overlapping set of

163 taxa between trees, by implementing definition 2.8 for n-overlap from Ané et al. (2009). If

164 the source chronograms do not form a grove, the supertree reconstruction will fail. In rare

165 cases, a group of trees can have multiple groves. By default, DateLife chooses the grove with

166 the most taxa, however, the “criterion = trees” flag allows the user to choose the grove with

167 the most trees instead. The result is a single summary (or supertree) topology, that

168 combines topologies from source chronograms in a grove.

169 *Applying Secondary Calibrations*

170 Once a topology is chosen, DateLife applies the congruification method (Eastman,

171 Harmon, & Tank, 2013) that find nodes belonging to the same clade across source

172 chronograms, and then extracts the corresponding node ages from patristic distance matrices

173 stored as a `datelifeResult` object. Note that by definition, these matrices store total
174 distance (time from tip to tip), assuming that the terminal taxa are coeval and occur at the
175 present. Hence, node ages correspond to half the values stored in the `datelifeResult`
176 matrices. A table of congruified node ages that can be used as calibrations for a dating
177 analysis is stored as a `congruifiedCalibrations` object.

178 For each congruent node, the pairwise distances that traverse that node are summarized
179 into a single summary matrix using classic summary statistics (i.e., mean, median, minimum
180 and maximum ages), and the Supermatrix Distance Method (SDM; Criscuolo et al., 2006),
181 which deforms patristic distance matrices by minimizing variance and then averaging them.
182 These single summary taxon pair age matrices are stored as summarized calibrations that
183 can be used as secondary calibrations to date a tree topology - with or without initial branch
184 lengths, using phylogenetic dating methods currently supported within DateLife: BLADJ
185 (Webb, Ackerly, & Kembel, 2008; Webb & Donoghue, 2005), MrBayes (Huelsenbeck &
186 Ronquist, 2001; Ronquist & Huelsenbeck, 2003), PATHd8 (Britton, Anderson, Jacquet,
187 Lundqvist, & Bremer, 2007), and treePL (Smith & O'Meara, 2012).

188 *Dating a Tree Topology*

189 **With no branch lengths.**— When producing or obtaining a tree with branch lengths
190 for a group of interest is not possible, DateLife can date a topology without branch lengths,
191 obtained from OpenTree or by implementing the supertree approach described above, by
192 implementing the Branch Length Adjuster (BLADJ; Webb et al., 2008; Webb & Donoghue,
193 2005) algorithm, which requires no initial branch lengths. The algorithm starts by fixing
194 ages for nodes with calibration data upon the given topology. Then, it distributes time for
195 nodes with no data evenly between calibrated nodes, minimizing age variance in the
196 resulting chronogram (Webb et al., 2008), which has proven useful for ecological analyses
197 that require age data (Webb et al., 2008). When there is conflict between ages of calibrated
198 nodes, BLADJ ignores node ages that are older than the age of a parent node. The BLADJ

199 algorithm requires a root age estimate to run. If there is no information on the age of the
200 root in the chronogram database, users can provide an estimate from the literature. If none
201 is provided, DateLife will not return a dated topology and will provide a warning message
202 along with suggestions on how the user can provide an age for the root so that a fully dated
203 tree can be returned.

204 In the absence of genetic data, we consider that BLADJ is the most agnostic way to
205 assign ages to nodes with no available data, as it does not require any assumptions on the
206 underlying model of branch length distribution. It is however common practice in the
207 literature to use a birth-death model to assign ages to nodes with no genetic data (Jetz,
208 Thomas, Joy, Hartmann, & Mooers, 2012; Rabosky et al., 2018; Smith & Brown, 2018). To
209 do so, DateLife implements MrBayes (Huelsenbeck & Ronquist, 2001; Ronquist &
210 Huelsenbeck, 2003), using nodes with published age data as calibration priors on a fixed
211 topology, a simple birth-death model with parameters that can be determined by the user,
212 and no genetic data.

213 **Dating a tree with branch lengths.**— Relative branch lengths provide key
214 information for phylogenetic dating, specifically for nodes without any calibration data
215 available. While, using initial branch lengths data is the golden standard for phylogenetic
216 dating analyses, producing such requires assembling and curating a homology hypothesis,
217 and choosing and implementing a method for phylogenetic inference. DateLife implements a
218 workflow to streamline this process by leveraging on genetic data from the Barcode of Life
219 Data System, BOLD (Ratnasingham & Hebert, 2007) to obtain markers for input taxa. By
220 default, mined genetic sequences are aligned with MUSCLE (Edgar, 2004) using functions
221 from the msa R package (Bodenhofer et al., 2015). Alternatively, sequences can be aligned
222 with MAFFT (Katoh, Asimenos, & Toh, 2009), using functions from the ape R package
223 (Paradis et al., 2004). The BOLD sequence alignment is then used to obtain initial branch
224 lengths with the accelerated transformation (ACCTRAN) parsimony algorithm, which

resolves ambiguous character optimization by assigning changes along branches of the tree as close to the root as possible (Agnarsson & Miller, 2008), resulting in older nodes (Forest et al., 2005). The parsimony branch lengths are then optimized using Maximum Likelihood, given the alignment, the topology and a simple Jukes-Cantor model, producing a BOLD tree with branch lengths proportional to expected number of substitutions per site. Both parsimony and ML optimizations are done with functions from the `phangorn` package (Schliep, 2011). The BOLD workflow is currently only available on DateLife’s R package and not on the web application.

Phylogenetic dating methods supported in DateLife that incorporate branch length information from the input topology in combination with the secondary calibrations include: PATHd8, a non-clock, rate-smoothing method to date trees (Britton et al., 2007); treePL (Smith & O’Meara, 2012), a semi-parametric, rate-smoothing, penalized likelihood dating method (Sanderson, 2002); and MrBayes (Huelsenbeck & Ronquist, 2001; Ronquist & Huelsenbeck, 2003), a Bayesian inference program implementing Markov chain Monte Carlo (MCMC) methods to estimate a posterior distribution of model parameters.

Visualizing Results

Finally, users can save all source and summary chronograms in formats allowing for reuse and reanalysis, such as newick and the R “phylo” format. Input and summary chronograms can be visualized and compared graphically, and users can construct their own graphs using DateLife’s chronogram plot generation functions available from the R package `datelifeplot` (Sanchez-Reyes & O’Meara, 2022).

BENCHMARK

R package `datelife` 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 time

increases roughly linearly with number of input taxon names, and increases considerably if Taxonomic Name Resolution Service (TNRS) is activated. Up to ten thousand names can be processed and searched in less than 30 minutes with the most time consuming settings. Once names have been processed as described in methods, a name search through the chronogram database 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) that were run both locally with the `devtools` package (R Core Team, 2018), and on a public server using the continuous integration tool of GitHub actions (<https://docs.github.com/en/actions>). At present, unit tests cover more than 40% of `datelife`'s code (<https://codecov.io/gh/phylotastic/datelife>). Unit testing helps identify potential issues as code is updated or, more critically, as services code relies upon may change.

263 CASE STUDIES

We illustrate the DateLife workflow using a family within the passeriform birds encompassing the true finches, Fringillidae, as case study. On a small example, we analysed 6 bird species, and results from each step of the workflow are shown in Fig. 3. As a second example, we analysed 289 bird species in the family Fringillidae that are included in the NCBI taxonomy. The resulting summary chronogram is shown in Fig. 5, and results from previous steps of the workflow are available as Supplementary Figures.

270 A Small Example

271 Creating a search query.— We chose 6 bird species within the Passeriformes. The sample includes two species of cardinals: the black-thighed grosbeak – *Pheucticus tibialis* and the crimson-collared grosbeak – *Rhodothraupis celaeno*; three species of buntings: the yellowhammer – *Emberiza citrinella*, the pine bunting – *Emberiza leucocephalos* and the

275 yellow-throated bunting – *Emberiza elegans*; and one species of tanager, the vegetarian finch –
276 *Platyspiza crassirostris*. Processing of input names found that *Emberiza elegans* is synonym
277 for *Schoeniclus elegans* in the default reference taxonomy (OTT v3.3, June 1, 2021). For a
278 detailed discussion on the state of the synonym, refer to Avibase (Avibase, 2022; Lepage,
279 2004; Lepage, Vaidya, & Guralnick, 2014). Discovering this synonym allowed assigning five
280 age data points for the parent node of *Emberiza elegans*, shown as *Schoeniclus elegans* in
281 figure 3A, which would not have had any data otherwise.

282 **Searching the database.**— DateLife used the processed input names to search the
283 local chronogram database and found 9 matching chronograms in 6 different studies (Fig.
284 3B). Three studies matched five input names (Barker, Burns, Klicka, Lanyon, & Lovette,
285 2015; Hedges, Marin, Suleski, Paymer, & Kumar, 2015; Jetz et al., 2012), one study matched
286 four input names (Hooper & Price, 2017) and two studies matched two input names (Barker,
287 Burns, Klicka, Lanyon, & Lovette, 2013; Burns et al., 2014). No studies matched all input
288 names. Together, source chronograms provide 28 unique age data points, covering all nodes
289 on our chosen tree topology to date (Table 1).

290 **Summarizing search results.**— DateLife obtained OpenTree’s synthetic tree
291 topology for these taxa (Fig. 3C), and congruified and mapped age data to nodes in this
292 chosen topology (Table 1). The name processing step allowed including five data points for
293 node “n4” (parent of *Schoeniclus elegans*; Fig. 3A) that would not have had any data
294 otherwise due to name mismatch. Age summary statistics per node were calculated (Table 2)
295 and used as calibrations to date the tree topology using the BLADJ algorithm. As expected,
296 more inclusive nodes (e.g., node “n1”) have more variance in age data than less inclusive
297 nodes (e.g., node “n5”). Summary age data for node “n2” were excluded as final calibration
298 because they are older than age data of the more inclusive node, “n1” (Fig. 3C4).

299

An Example with the Family of True Finches

300

Creating a query.— To obtain ages for all species within the family of true finches,

301 Fringillidae, we ran a DateLife query using the “get species from taxon” flag, which gets all
302 recognized species names within a named group from a taxonomy of choice. Following the
303 NCBI taxonomy, our DateLife query has 289 Fringillidae species names. This
304 taxon-constrained approach implies that the full DateLife analysis will be performed using a
305 tree topology and ages available for species names from a given taxonomic group, which do
306 not necessarily correspond to a monophyletic group. Users can change this behavior by
307 providing all species names corresponding to a monophyletic group as input for a DateLife
308 search, or a monophyletic tree to construct a DateLife summary.

309

Searching the database.— Next, we used the processed species names in our

310 DateLife query to identify chronograms with at least two Fringillidae species as tip taxa.
311 The DateLife search identified 19 chronograms matching this criteria, published in 13
312 different studies (Barker et al., 2013, 2015; Burns et al., 2014; Claramunt & Cracraft, 2015;
313 Gibb et al., 2015; Hedges et al., 2015; Hooper & Price, 2017; Jetz et al., 2012; Kimball et al.,
314 2019; Oliveros et al., 2019; Price et al., 2014; Roquet, Lavergne, & Thuiller, 2014; Uyeda,
315 Pennell, Miller, Maia, & McClain, 2017). Once identified, DateLife pruned these matching
316 chronograms to remove tips that do not belong to the queried taxon names, and transformed
317 these pruned chronograms to pairwise distance matrices, revealing 1,206 different age data
318 points available for species within the Fringillidae (Supplementary Table S1).

319

Summarizing search results.— The final step entailed congruifying and

320 summarizing the age data available for the Fringillidae species into two single summary
321 chronograms, using two different types of summary ages, median and SDM. As explained in
322 the “Description” section, a tree topology to summarize age data upon is required. By
323 default, DateLife uses the topology from OpenTree’s synthetic tree that contains all taxa
324 from the search query. According to OpenTree’s synthetic tree, species belonging to the

325 family Fringillidae do not form a monophyletic group (Fig. 4). Hence, a topology containing
326 only the 289 species from the original query was extracted from Open Tree of Life's synthetic
327 tree v12.3 (Open Tree Of Life et al., 2019).

328 Age data from source chronograms was congruified to OpenTree's topology (Fig. 4B),
329 reducing the age data set to 818 different data points (Supplementary Table S2). For each
330 congruent node, age summary statistics were calculated and used as fixed secondary
331 calibrations over the chosen tree topology, to obtain a fully dated phylogeny with the
332 program BLADJ (Fig. 5).

333 **CROSS-VALIDATION TEST**

334 We performed a cross validation test of a DateLife analysis using the Fringillidae
335 source chronograms obtained above (Supplementary Figs. S2-S20). We used as inputs for a
336 DateLife analysis all individual tree topologies from each of the 19 source chronograms from
337 13 studies, treating their node ages as unknown. We congruified node ages extracted from
338 chronograms from all other studies upon the individual topologies, effectively excluding
339 original ages from each topology. Average node ages per node were applied as secondary
340 calibrations, smoothed with the BLADJ algorithm. We found that node ages from the
341 original studies, and ages estimated using all other age data available are largely correlated
342 (Fig. 6). For five studies, DateLife tended to underestimate ages for topologically deeper
343 nodes (those with many descendant taxa, aka “closer to the root”) relative to the original
344 estimate, and overestimate ages for nodes closer to the tips. Accordingly, root ages are
345 generally older in the original study than estimated using cross-validated ages
346 (Supplementary Fig. S1). In general, topologically deeper nodes display the largest age
347 variation between node ages from the original chronograms and ages summarized with
348 DateLife.

349 **DISCUSSION**

350 DateLife makes state-of-the-art data on evolutionary time frame easily accessible for
351 comparison, reuse, and reanalysis, to researchers in all areas of science and with all levels of
352 expertise in the matter. It is an open service that does not require any expert biological
353 knowledge from users –besides the names of the species or group they want to work with, for
354 any of its functionality.

355 A total of 99,474 unique terminal taxa are represented in DateLife’s database.

356 Incorporation of more chronograms into the database will continue to improve DateLife’s
357 services. One option to increase the number of chronograms in the DateLife database is the
358 Dryad data repository. Methods to automatically mine chronograms from Dryad could be
359 designed and implemented. However, Dryad’s metadata system has no information to
360 automatically detect branch length units, and those would still need to be determined
361 manually by a human curator. We would like to emphasize on the importance of sharing
362 chronogram data, including systematically curated metadata, into open repositories, such as
363 OpenTree’s Phylesystem (McTavish et al., 2015) for the benefit of the scientific community
364 as a whole.

365 While it is known that incorporating at least some data on lineage divergence times
366 represents a relevant improvement for testing alternative hypothesis using phylogenetic
367 distance in ecological and conservation biology studies (Webb et al., 2008). Next, we discuss
368 some particularities of using summarized node age data from a DateLife analysis that users
369 should be aware of.

370 *Age Variation in Source Chronograms*

371 Conflict in estimated ages among alternative studies is common in the literature. See,
372 for example, the robust ongoing debate about crown group age of angiosperms
373 (Barba-Montoya, Reis, Schneider, Donoghue, & Yang, 2018; Magallón, Gómez-Acevedo,
374 Sánchez-Reyes, & Hernández-Hernández, 2015; Ramshaw et al., 1972; Sanderson & Doyle,

375 2001; Sauquet, Ramírez-Barahona, & Magallón, 2021). Source chronograms available for the
376 same organisms have potentially been estimated implementing calibrations very differently.
377 In the Fringillidae example above, the chronograms from Burns et al. (2014) were inferred
378 using molecular substitution rate estimates across birds (Weir & Schluter, 2008), and have
379 much older age estimates for the same nodes than chronograms that were inferred using
380 fossils as calibrations (Figs. 5, 6; Supplementary Figs. S1, S5).

381 Different calibration implementations might also imply fundamentally distinct
382 evolutionary hypotheses (Antonelli et al., 2017). For example, two independent researchers
383 working on the same clade should both carefully select and justify their choices of fossil
384 calibration placement. Yet, if one researcher concludes that a fossil should calibrate the
385 ingroup of a clade, while another researcher concludes that the same fossil should calibrate
386 the outgroup of the clade, the resulting age estimates will differ, as the placement of
387 calibrations as stem or crown group has been proven to significantly affect time of lineage
388 divergence estimates (Sauquet, 2013).

389 *Primary vs Secondary Calibrations*

390 While most chronograms in DateLife's database are constructed using primary
391 calibrations (molecular substitution rates or ages obtained from the fossil record or geological
392 events), DateLife summarizes chronograms using secondary calibrations (ages coming from
393 other chronograms). Graur and Martin (2004) cautioned on the increased error and
394 uncertainty in estimated ages when using secondary calibrations in dating analyses. Schenk
395 (2016) showed that, in simulations, divergence times inferred using secondary calibrations are
396 significantly younger than those inferred with primary calibrations, when obtained with
397 Bayesian inference methods, and when priors are implemented in similar ways in both
398 analyses. Accordingly, the scientific community seems to have more confidence in
399 chronograms obtained from a single analysis, using fossil data as primary sources of
400 calibrations (Schenk, 2016), and using fossils that have been widely discussed and curated as

401 calibrations to date other trees, making sure that all data reflect a coherent evolutionary
402 history (Sauquet, 2013), as for example done by Antonelli et al. (2017). There have been
403 attempts to create fossil calibration databases (Ksepka et al., 2015), though these still have
404 room to grow.

405 It seems that using several (as opposed to just a few) secondary calibrations can
406 provide sufficient information to alleviate or even neutralize potential biases (Sauquet, 2013).
407 Certainly, further studies are required to fully understand the effect of secondary calibrations
408 on outputs from different tree dating methods, and on downstream analyses. It is possible
409 that secondary calibrations can be safely used with dating methods that do not require
410 setting priors, such as penalized likelihood (Sanderson, 2003), with methods that do not
411 make any assumptions on the ages and fix them to a node on a tree topology, such as
412 BLADJ (Webb et al., 2008; Webb & Donoghue, 2005), or methods that summarize age data
413 unto a tree topology.

414 Our cross validation analysis might provide some insight in this regard. When ages are
415 estimated with secondary calibrations, nodes closer to the root do tend to be slightly
416 younger than ages estimated with primary calibrations. However, nodes closer to the tip
417 tend to be older when estimated using secondary calibrations with a dating method that
418 does not make any prior assumptions on the nature of the calibrations themselves
419 (Supplementary Figures S2-S20). The only exception to this was observed on cross validation
420 results of the Burns et al. (2014) chronogram, which displays much younger node ages when
421 estimated using secondary calibrations (Supplementary Figs. S1, S5).

422 *Summarizing Chronograms*

423 By default, DateLife currently summarizes all source chronograms that overlap with at
424 least two species names. Users can exclude source chronograms if they have reasons to do so.
425 Strictly speaking, a good chronogram should reflect the real time of lineage divergence

426 accurately and precisely. To our knowledge, there are no tested measures to determine
427 independently when a chronogram is better than another. Yet, several characteristics of the
428 data used for dating analyses, as well as from the output chronogram itself, could be used to
429 score the quality of source chronograms.

430 Some measures that have been proposed are the proportion of lineage sampling and the
431 number of calibrations used (Magallón, 2010; Magallón et al., 2015). Some characteristics
432 that are often cited in published studies as a measure of improved age estimates as compared
433 to previously published estimates are: quality of alignment (missing data, GC content),
434 lineage sampling (strategy and proportion), phylogenetic and dating inference method,
435 number of fossils used as calibrations, support for nodes and ages, and magnitude of
436 confidence intervals.

437 DateLife provides an opportunity to capture concordance and conflict among date
438 estimates, which can also be used as a metric for chronogram reliability. Its open database of
439 chronograms allows other researchers to do such analyses themselves reproducibly, and
440 without needing permission. Though, of course, they should follow proper citation practices,
441 especially for the source chronogram studies.

442 The exercise of summarizing age data from across multiple studies provides the
443 opportunity to work with a chronogram that reflects a unified evolutionary history for a
444 lineage, by putting together evidence from different hypotheses. The largest, and
445 taxonomically broadest chronogram currently available from OpenTree was constructed
446 summarizing age data from 2,274 published chronograms using NCBI's taxonomic tree as
447 backbone (Hedges et al., 2015). A summarizing exercise may also amplify the effect of
448 uncertainty and errors in source data, and blur parts of the evolutionary history of a lineage
449 that might only be reflected in source chronograms and lost on the summary chronogram
450 (Sauquet et al., 2021).

451 *Effects of Phylogenetic Sampling on Downstream Analyses*

452 Analysis of species diversification using simulated and empirical phylogenies suggest
453 that using a more completely sampled phylogeny provides estimates that are closer to the
454 true diversification history than when analysing incompletely sampled phylogenies (Chang,
455 Rabosky, & Alfaro, 2020; Cusimano, Stadler, & Renner, 2012; Sun et al., 2020). Ideally,
456 phylogenies should be completed using genetic data, but this is a difficult task to achieve for
457 many biological groups. Hence, DateLife's workflow features different ways of assigning node
458 ages in the absence of calibrations and branch length information for certain taxa.

459 Completing a phylogeny using a stochastic birth-death polotomy resolver and a backbone
460 taxonomy is a common practice in scientific publications: Jetz et al. (2012), created a
461 chronogram of all 9, 993 bird species, where 67% had molecular data and the rest was
462 simulated; Rabosky et al. (2018) created a chronogram of 31, 536 ray-finned fishes, of which
463 only 37% had molecular data; Smith and Brown (2018) constructed a chronogram of 353, 185
464 seed plants where only 23% had molecular data. These stochastically resolved chronograms
465 provide less biased diversification rates estimated with methods that account for sampling
466 fractions (Chang et al., 2020; Cusimano et al., 2012), but can also introduce spurious
467 patterns of early bursts of diversification (Cusimano & Renner, 2010; Sun et al., 2020).

468 Taxonomy based polotomy resolvers also introduce topological differences. The study of
469 macroevolutionary processes largely depends on an understanding of the timing of species
470 diversification events, and different phylogenetic and chronogram hypothesis can provide
471 very different overviews of the macroevolutionary history of a biological group. For example,
472 alternative topologies in chronograms from the same biological group can infer very different
473 species diversification patterns (Rabosky, 2015; Title & Rabosky, 2016). Similarly there are
474 worries that patterns of morphological evolution cannot be accurately inferred with these
475 types of phylogenies, as any patterns would be erased by randomization (Rabosky, 2015).
476 We note that the same applies for geographical and morphological dependent diversification

477 analysis. Hence, we suggest that these types of randomly resolved phylogenies can be used as
478 null or neutral models, representing the case of diversification independent of traits and
479 geographical scenario.

480 While stochastic politomy resolvers have been useful to advance research, notably, risks
481 come with this practice. Taken to the extreme, one could generate a fully resolved,
482 calibrated tree of all modern and extinct taxa using a single taxonomy, a single calibration,
483 and assigning branch lengths following a birth-death diversification model. Clearly, this can
484 lead to a misrepresentation of the true evolutionary history. We urge DateLife users to follow
485 the example of the large tree papers cited above, by carefully considering the statistical
486 assumptions being made, potential biases, and assessing the consistency of the results with
487 prior work.

488 CONCLUSIONS

489 Knowledge of the evolutionary time frame of organisms is key to many research areas:
490 trait evolution, species diversification, biogeography, macroecology and more. It is also
491 crucial for education, science communication and policy, but generating chronograms is
492 difficult, especially for those who want to use phylogenies but who are not systematists, or
493 do not have the time to acquire and develop the necessary knowledge and skills to construct
494 them on their own. Importantly, years of primarily publicly funded research have resulted in
495 vast amounts of chronograms that are already available on scientific publications, but hidden
496 to the public and scientific community for reuse.

497 The DateLife project allows for easy and fast summary of public and state-of-the-art
498 data on time of lineage divergence. It provides a straightforward way to get an informed idea
499 on the state of knowledge of the time frame of evolution of different regions of the tree of life,
500 and allows identifying regions that require more research, or that have conflicting
501 information. It is available as an R package, and as a web-based R shiny application at

502 www.datelife.org Both summary and newly generated trees are useful to evaluate
503 evolutionary hypotheses in different areas of research. The DateLife project should improve
504 awareness of the existing variation in expert time of divergence data, and foster exploration
505 of the effect of alternative divergence time hypothesis on the results of analyses, nurturing a
506 culture of more cautious interpretation of evolutionary results.

507 **AVAILABILITY**

508 The DateLife software is free and open source. It can be used online through its R
509 shiny web application hosted at <http://www.datelife.org>, and locally through the **datelife**
510 R package, downloadable from Zenodo (<https://doi.org/10.5281/zenodo.593938>) and the
511 CRAN repository (Sanchez-Reyes et al., 2022). DateLife's web application is maintained
512 using RStudio's shiny server and the shiny package open infrastructure, as well as Docker
513 and OpenTree's infrastructure (dates.opentreeoflife.org/datelife). **datelife**'s stable version
514 is available for installation from CRAN repository using the command
515 `install.packages(pkgs = "datelife")` from within R. Development versions are
516 available from DateLife's GitHub repository (<https://github.com/phylotastic/datelife>) and
517 can be installed using the command
518 `devtools::install_github("phylotastic/datelife")`.

519 **SUPPLEMENTARY MATERIAL**

520 Supplementary material, including code, biological examples, benchmark results, data
521 files and online-only appendices, can be found in the Dryad data repository
522 (<https://doi.org/10.5061/dryad.cnp5hqc6w>), as well as in the Zenodo stable repositories that
523 host the reproducible manuscript (<https://doi.org/10.5281/zenodo.7435094>), the biological
524 examples (<https://doi.org/10.5281/zenodo.7435101>), and the software benchmark
525 (<https://doi.org/10.5281/zenodo.7435106>). The corresponding development versions hosted
526 on GitHub can be found at <https://github.com/LunaSare/datelifeMS1#readme>,

- 527 https://github.com/LunaSare/datelife_examples#readme, and
528 https://github.com/LunaSare/datelife_benchmark#readme.

529

FUNDING

530 Funding was provided by the US National Science Foundation (NSF) grants
531 ABI-1458603 to the DateLife project; DBI-0905606 to the National Evolutionary Synthesis
532 Center (NESCent); ABI-1458572 to the Phylotastic project; and ABI-1759846 to the Open
533 Tree of Life project.

534

ACKNOWLEDGEMENTS

535 The DateLife project was born as a prototype tool aiming to provide the services
536 describe in this paper, and was initially developed over a series of hackathons at the National
537 Evolutionary Synthesis Center, NC, USA (Stoltzfus et al., 2013). We thank colleagues from
538 the O'Meara Lab at the University of Tennessee Knoxville for suggestions, discussions and
539 software testing. The late National Evolutionary Synthesis Center (NESCent), which
540 sponsored hackathons that led to initial work on this project. The team that assembled
541 DateLife's first proof of concept: Tracy Heath, Jonathan Eastman, Peter Midford, Joseph
542 Brown, Matt Pennell, Mike Alfaro, and Luke Harmon. The Open Tree of Life project that
543 provides the open, metadata rich repository of trees used to construct DateLife's chronogram
544 database. The many scientists who publish their chronograms in an open, reusable form, and
545 the scientists who curate them for deposition in the Open Tree of Life repository. The NSF
546 for funding nearly all the above, in addition to the ABI grant that funded this project itself.

References

- 547
- 548 Agnarsson, I., & Miller, J. A. (2008). Is acctran better than deltran? *Cladistics*, 24(6),
549 1032–1038.
- 550 Alström, P., Hooper, D. M., Liu, Y., Olsson, U., Mohan, D., Gelang, M., … Price, T. D.
551 (2014). Discovery of a relict lineage and monotypic family of passerine birds. *Biology
552 Letters*, 10(3), 20131067.
- 553 Ané, C., Eulenstein, O., Piaggio-Talice, R., & Sanderson, M. J. (2009). Groves of
554 phylogenetic trees. *Annals of Combinatorics*, 13(2), 139–167.
- 555 Antonelli, A., Hettling, H., Condamine, F. L., Vos, K., Nilsson, R. H., Sanderson, M. J., …
556 Vos, R. A. (2017). Toward a self-updating platform for estimating rates of speciation
557 and migration, ages, and relationships of Taxa. *Systematic Biology*, 66(2), 153–166.
558 <https://doi.org/10.1093/sysbio/syw066>
- 559 Archie, J., Day, W. H., Felsenstein, J., Maddison, W., Meacham, C., Rohlf, F. J., &
560 Swofford, D. (1986). The Newick tree format. Retrieved from
561 {<https://evolution.genetics.washington.edu/phylip/newicktree.html>}
- 562 Avibase. (2022). Yellow-throated Bunting. *Avibase - the World Bird Database*, (Online
563 Resource). Retrieved from
564 {<https://avibase.bsc-eoc.org/species.jsp?lang=EN&avibaseid=82D1EE0049D8D927>}
- 565 Bapst, D. W. (2012). Paleotree: An R package for paleontological and phylogenetic analyses
566 of evolution. *Methods in Ecology and Evolution*, 3(5), 803–807.
567 <https://doi.org/10.1111/j.2041-210X.2012.00223.x>
- 568 Barba-Montoya, J., Reis, M. dos, Schneider, H., Donoghue, P. C., & Yang, Z. (2018).
569 Constraining uncertainty in the timescale of angiosperm evolution and the veracity of

- 570 a cretaceous terrestrial revolution. *New Phytologist*, 218(2), 819–834.
- 571 Barker, F. K. (2014). Mitogenomic data resolve basal relationships among passeriform and
572 passeridan birds. *Molecular Phylogenetics and Evolution*, 79, 313–324.
- 573 Barker, F. K., Burns, K. J., Klicka, J., Lanyon, S. M., & Lovette, I. J. (2013). Going to
574 extremes: Contrasting rates of diversification in a recent radiation of new world
575 passerine birds. *Systematic Biology*, 62(2), 298–320.
- 576 Barker, F. K., Burns, K. J., Klicka, J., Lanyon, S. M., & Lovette, I. J. (2015). New insights
577 into new world biogeography: An integrated view from the phylogeny of blackbirds,
578 cardinals, sparrows, tanagers, warblers, and allies. *The Auk: Ornithological Advances*,
579 132(2), 333–348.
- 580 Barker, F. K., Cibois, A., Schikler, P., Feinstein, J., & Cracraft, J. (2004). Phylogeny and
581 diversification of the largest avian radiation. *Proceedings of the National Academy of
582 Sciences*, 101(30), 11040–11045.
- 583 Beresford, P., Barker, F., Ryan, P., & Crowe, T. (2005). African endemics span the tree of
584 songbirds (passeri): Molecular systematics of several evolutionary “enigmas”.
585 *Proceedings of the Royal Society B: Biological Sciences*, 272(1565), 849–858.
- 586 Bodenhofer, U., Bonatesta, E., Horejš-Kainrath, C., & Hochreiter, S. (2015). Msa: An r
587 package for multiple sequence alignment. *Bioinformatics*, 31(24), 3997–3999.
- 588 Boyle, B., Hopkins, N., Lu, Z., Raygoza Garay, J. A., Mozzherin, D., Rees, T., ... Enquist,
589 B. J. (2013). The taxonomic name resolution service: An online tool for automated
590 standardization of plant names. *BMC Bioinformatics*, 14(1).
591 <https://doi.org/10.1186/1471-2105-14-16>
- 592 Britton, T., Anderson, C. L., Jacquet, D., Lundqvist, S., & Bremer, K. (2007). Estimating

- 593 Divergence Times in Large Phylogenetic Trees. *Systematic Biology*, 56(788777878),
594 741–752. <https://doi.org/10.1080/10635150701613783>
- 595 Bryson Jr, R. W., Chaves, J., Smith, B. T., Miller, M. J., Winker, K., Pérez-Emán, J. L., &
596 Klicka, J. (2014). Diversification across the new world within the ‘blue’cardinalids
597 (aves: Cardinalidae). *Journal of Biogeography*, 41(3), 587–599.
- 598 Burleigh, J. G., Kimball, R. T., & Braun, E. L. (2015). Building the avian tree of life using a
599 large-scale, sparse supermatrix. *Molecular Phylogenetics and Evolution*, 84, 53–63.
- 600 Burns, K. J., Shultz, A. J., Title, P. O., Mason, N. A., Barker, F. K., Klicka, J., ... Lovette,
601 I. J. (2014). Phylogenetics and diversification of tanagers (passeriformes:
602 Thraupidae), the largest radiation of neotropical songbirds. *Molecular Phylogenetics
603 and Evolution*, 75, 41–77.
- 604 Chamberlain, S. (2018). *bold: Interface to Bold Systems API*. Retrieved from
605 <https://CRAN.R-project.org/package=bold>
- 606 Chamberlain, S. A., & Szöcs, E. (2013). taxize : taxonomic search and retrieval in R [version
607 2; referees: 3 approved]. *F1000Research*, 2(191), 1–29.
608 <https://doi.org/10.12688/f1000research.2-191.v2>
- 609 Chang, J., Rabosky, D. L., & Alfaro, M. E. (2020). Estimating diversification rates on
610 incompletely sampled phylogenies: Theoretical concerns and practical solutions.
611 *Systematic Biology*, 69(3), 602–611.
- 612 Chaves, J. A., Hidalgo, J. R., & Klicka, J. (2013). Biogeography and evolutionary history of
613 the neotropical genus Saltator (aves: Thraupini). *Journal of Biogeography*, 40(11),
614 2180–2190.
- 615 Claramunt, S., & Cracraft, J. (2015). A new time tree reveals earth history’s imprint on the

- 616 evolution of modern birds. *Science Advances*, 1(11), e1501005.
- 617 Criscuolo, A., Berry, V., Douzery, E. J., & Gascuel, O. (2006). SDM: A fast distance-based
618 approach for (super)tree building in phylogenomics. *Systematic Biology*, 55(5),
619 740–755. <https://doi.org/10.1080/10635150600969872>
- 620 Cusimano, N., & Renner, S. S. (2010). Slowdowns in diversification rates from real
621 phylogenies may not be real. *Systematic Biology*, 59(4), 458–464.
- 622 Cusimano, N., Stadler, T., & Renner, S. S. (2012). A new method for handling missing
623 species in diversification analysis applicable to randomly or nonrandomly sampled
624 phylogenies. *Systematic Biology*, 61(5), 785–792.
- 625 Delsuc, F., Philippe, H., Tsagkogeorga, G., Simion, P., Tilak, M.-K., Turon, X., ... Douzery,
626 E. J. (2018). A phylogenomic framework and timescale for comparative studies of
627 tunicates. *BMC Biology*, 16(1), 1–14.
- 628 Eastman, J. M., Harmon, L. J., & Tank, D. C. (2013). Congruification: Support for time
629 scaling large phylogenetic trees. *Methods in Ecology and Evolution*, 4(7), 688–691.
630 <https://doi.org/10.1111/2041-210X.12051>
- 631 Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high
632 throughput. *Nucleic Acids Research*, 32(5), 1792–1797.
- 633 Felsenstein, J. (1985). Phylogenies and the Comparative Method. *The American Naturalist*,
634 125(1), 1–15. Retrieved from <http://www.jstor.org/stable/2461605>
- 635 Forest, F., Savolainen, V., Chase, M. W., Lupia, R., Bruneau, A., & Crane, P. R. (2005).
636 Teasing apart molecular-versus fossil-based error estimates when dating phylogenetic
637 trees: A case study in the birch family (betulaceae). *Systematic Botany*, 30(1),
638 118–133.

- 639 Freckleton, R. P., Harvey, P. H., & Pagel, M. (2002). Phylogenetic analysis and comparative
640 data: A test and review of evidence. *The American Naturalist*.
- 641 GBIF Secretariat. (2022). GBIF Backbone Taxonomy. *Checklist dataset*, (Online Resource
642 accessed via GBIF.org). Retrieved from {<https://doi.org/10.15468/39omei> }
- 643 Gibb, G. C., England, R., Hartig, G., McLenachan, P. A., Taylor Smith, B. L., McComish,
644 B. J., ... Penny, D. (2015). New zealand passerines help clarify the diversification of
645 major songbird lineages during the oligocene. *Genome Biology and Evolution*, 7(11),
646 2983–2995.
- 647 Graur, D., & Martin, W. (2004). Reading the entrails of chickens: Molecular timescales of
648 evolution and the illusion of precision. *TRENDS in Genetics*, 20(2), 80–86.
- 649 Hackett, S. J., Kimball, R. T., Reddy, S., Bowie, R. C., Braun, E. L., Braun, M. J., ...
650 others. (2008). A phylogenomic study of birds reveals their evolutionary history.
651 *Science*, 320(5884), 1763–1768.
- 652 Harvey, P. H., Pagel, M. D., & others. (1991). *The comparative method in evolutionary
653 biology* (Vol. 239). Oxford university press Oxford.
- 654 Hedges, S. B., Marin, J., Suleski, M., Paymer, M., & Kumar, S. (2015). Tree of life reveals
655 clock-like speciation and diversification. *Molecular Biology and Evolution*, 32(4),
656 835–845. <https://doi.org/10.1093/molbev/msv037>
- 657 Heibl, C. (2008). *PHYLOCH: R language tree plotting tools and interfaces to diverse
658 phylogenetic software packages*. Retrieved from
659 <http://www.christophheibl.de/Rpackages.html>
- 660 Hooper, D. M., & Price, T. D. (2017). Chromosomal inversion differences correlate with
661 range overlap in passerine birds. *Nature Ecology & Evolution*, 1(10), 1526.

- 662 Huelsenbeck, J. P., & Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic
663 trees. *Bioinformatics*, 17(8), 754–755.
- 664 <https://doi.org/10.1093/bioinformatics/17.8.754>
- 665 Jetz, W., Thomas, G., Joy, J. J., Hartmann, K., & Mooers, A. (2012). The global diversity
666 of birds in space and time. *Nature*, 491(7424), 444–448.
667 <https://doi.org/10.1038/nature11631>
- 668 Johansson, U. S., Fjeldså, J., & Bowie, R. C. (2008). Phylogenetic relationships within
669 passerida (aves: Passeriformes): A review and a new molecular phylogeny based on
670 three nuclear intron markers. *Molecular Phylogenetics and Evolution*, 48(3), 858–876.
- 671 Katoh, K., Asimenos, G., & Toh, H. (2009). Multiple alignment of dna sequences with mafft.
672 In *Bioinformatics for dna sequence analysis* (pp. 39–64). Springer.
- 673 Kimball, R. T., Oliveros, C. H., Wang, N., White, N. D., Barker, F. K., Field, D. J., ...
674 others. (2019). A phylogenomic supertree of birds. *Diversity*, 11(7), 109.
- 675 Klicka, J., Barker, F. K., Burns, K. J., Lanyon, S. M., Lovette, I. J., Chaves, J. A., & Bryson
676 Jr, R. W. (2014). A comprehensive multilocus assessment of sparrow (aves:
677 Passerellidae) relationships. *Molecular Phylogenetics and Evolution*, 77, 177–182.
- 678 Ksepka, D. T., Parham, J. F., Allman, J. F., Benton, M. J., Carrano, M. T., Cranston, K.
679 A., ... others. (2015). The fossil calibration database—a new resource for divergence
680 dating. *Systematic Biology*, 64(5), 853–859.
- 681 Lamichhaney, S., Berglund, J., Almén, M. S., Maqbool, K., Grabherr, M., Martinez-Barrio,
682 A., ... others. (2015). Evolution of darwin's finches and their beaks revealed by
683 genome sequencing. *Nature*, 518(7539), 371–375.
- 684 Laubichler, M. D., & Maienschein, J. (2009). *Form and function in developmental evolution*.

- 685 Cambridge University Press.
- 686 Lepage, D. (2004). *Avibase: The world bird database*. Bird Studies Canada.
- 687 Lepage, D., Vaidya, G., & Guralnick, R. (2014). Avibase—a database system for managing
688 and organizing taxonomic concepts. *ZooKeys*, (420), 117.
- 689 Lerner, H. R., Meyer, M., James, H. F., Hofreiter, M., & Fleischer, R. C. (2011). Multilocus
690 resolution of phylogeny and timescale in the extant adaptive radiation of hawaiian
691 honeycreepers. *Current Biology*, 21(21), 1838–1844.
- 692 Lovette, I. J., Pérez-Emán, J. L., Sullivan, J. P., Banks, R. C., Fiorentino, I.,
693 Córdoba-Córdoba, S., ... others. (2010). A comprehensive multilocus phylogeny for
694 the wood-warblers and a revised classification of the parulidae (aves). *Molecular
695 Phylogenetics and Evolution*, 57(2), 753–770.
- 696 Magallon, S., & Sanderson, M. (2001). Absolute diversification rates in angiosperm clades.
697 *Evolution*, 55(9), 1762–1780.
- 698 Magallón, S. (2010). Using fossils to break long branches in molecular dating: A comparison
699 of relaxed clocks applied to the origin of angiosperms. *Systematic Biology*, 59(4),
700 384–399.
- 701 Magallón, S., Gómez-Acevedo, S., Sánchez-Reyes, L. L., & Hernández-Hernández, T. (2015).
702 A metacalibrated time-tree documents the early rise of flowering plant phylogenetic
703 diversity. *New Phytologist*, 207(2), 437–453.
- 704 McTavish, E. J., Hinchliff, C. E., Allman, J. F., Brown, J. W., Cranston, K. A., Holder, M.
705 T., ... Smith, S. (2015). Phylesystem: A git-based data store for community-curated
706 phylogenetic estimates. *Bioinformatics*, 31(17), 2794–2800.
- 707 Michonneau, F., Brown, J. W., & Winter, D. J. (2016). rotl: an R package to interact with

- 708 the Open Tree of Life data. *Methods in Ecology and Evolution*, 7(12), 1476–1481.
- 709 <https://doi.org/10.1111/2041-210X.12593>
- 710 Morlon, H. (2014). Phylogenetic approaches for studying diversification. *Ecology Letters*,
- 711 17(4), 508–525. <https://doi.org/10.1111/ele.12251>
- 712 Moyle, R. G., Oliveros, C. H., Andersen, M. J., Hosner, P. A., Benz, B. W., Manthey, J. D.,
- 713 ... Faircloth, B. C. (2016). Tectonic collision and uplift of Wallacea triggered the
- 714 global songbird radiation. *Nature Communications*, 7(1), 1–7.
- 715 Oliveros, C. H., Field, D. J., Ksepka, D. T., Barker, F. K., Aleixo, A., Andersen, M. J., ...
- 716 others. (2019). Earth history and the passerine superradiation. *Proceedings of the*
- 717 *National Academy of Sciences*, 116(16), 7916–7925.
- 718 Ooms, J., & Chamberlain, S. (2018). *Phylocomr: Interface to 'phylocom'*. Retrieved from
- 719 <https://CRAN.R-project.org/package=phylocomr>
- 720 Open Tree Of Life, Redelings, B., Cranston, K. A., Allman, J., Holder, M. T., & McTavish,
- 721 E. J. (2016). Open Tree of Life APIs v3.0. *Open Tree of Life Project*, (Online
- 722 Resources). Retrieved from
- 723 {<https://github.com/OpenTreeOfLife/germinator/wiki/Open-Tree-of-Life-Web->
- 724 APIs}
- 725 Open Tree Of Life, Redelings, B., Sánchez Reyes, L. L., Cranston, K. A., Allman, J., Holder,
- 726 M. T., & McTavish, E. J. (2019). Open tree of life synthetic tree v12.3. *Zenodo*.
- 727 Retrieved from <https://doi.org/10.5281/zenodo.3937742>
- 728 Ödeen, A., Håstad, O., & Alström, P. (2011). Evolution of ultraviolet vision in the largest
- 729 avian radiation—the passerines. *BMC Evolutionary Biology*, 11(1), 1–8.
- 730 Paradis, E., Claude, J., & Strimmer, K. (2004). APE: analyses of phylogenetics and

- 731 evolution in R language. *Bioinformatics*, 20(2), 289–290.
- 732 Parchman, T. L., Benkman, C. W., & Mezquida, E. T. (2007). Coevolution between
733 hispaniolan crossbills and pine: Does more time allow for greater phenotypic
734 escalation at lower latitude? *Evolution*, 61(9), 2142–2153.
- 735 Päckert, M., Martens, J., Sun, Y.-H., Severinghaus, L. L., Nazarenko, A. A., Ting, J., ...
736 Tietze, D. T. (2012). Horizontal and elevational phylogeographic patterns of
737 himalayan and southeast asian forest passerines (aves: Passeriformes). *Journal of
738 Biogeography*, 39(3), 556–573.
- 739 Pennell, M. W., Eastman, J. M., Slater, G. J., Brown, J. W., Uyeda, J. C., FitzJohn, R. G.,
740 ... Harmon, L. J. (2014). Geiger v2. 0: An expanded suite of methods for fitting
741 macroevolutionary models to phylogenetic trees. *Bioinformatics*, 30(15), 2216–2218.
- 742 Posadas, P., Crisci, J. V., & Katinas, L. (2006). Historical biogeography: A review of its
743 basic concepts and critical issues. *Journal of Arid Environments*, 66(3), 389–403.
- 744 Powell, A. F., Barker, F. K., Lanyon, S. M., Burns, K. J., Klicka, J., & Lovette, I. J. (2014).
745 A comprehensive species-level molecular phylogeny of the new world blackbirds
746 (icteridae). *Molecular Phylogenetics and Evolution*, 71, 94–112.
- 747 Price, T. D., Hooper, D. M., Buchanan, C. D., Johansson, U. S., Tietze, D. T., Alström, P.,
748 ... others. (2014). Niche filling slows the diversification of himalayan songbirds.
749 *Nature*, 509(7499), 222.
- 750 Pulgarín-R, P. C., Smith, B. T., Bryson Jr, R. W., Spellman, G. M., & Klicka, J. (2013).
751 Multilocus phylogeny and biogeography of the new world pheucticus grosbeaks (aves:
752 Cardinalidae). *Molecular Phylogenetics and Evolution*, 69(3), 1222–1227.
- 753 Rabosky, D. L. (2015). No substitute for real data: A cautionary note on the use of

- 754 phylogenies from birth–death polytomy resolvers for downstream comparative
755 analyses. *Evolution*, 69(12), 3207–3216.
- 756 Rabosky, D. L., Chang, J., Title, P. O., Cowman, P. F., Sallan, L., Friedman, M., ... others.
757 (2018). An inverse latitudinal gradient in speciation rate for marine fishes. *Nature*,
758 559(7714), 392.
- 759 Ramshaw, J., Richardson, D., Meatyard, B., Brown, R., Richardson, M., Thompson, E., &
760 Boulter, D. (1972). The time of origin of the flowering plants determined by using
761 amino acid sequence data of cytochrome c. *New Phytologist*, 71(5), 773–779.
- 762 Ratnasingham, S., & Hebert, P. D. (2007). BOLD: The barcode of life data system
763 (<http://www.Barcodinglife.Org>). *Molecular Ecology Notes*, 7(3), 355–364.
- 764 R Core Team. (2018). *R: a language and environment for statistical computing*. Vienna,
765 Austria: R Foundation for Statistical Computing.
- 766 Rees, & Cranston, K. (2017). Automated assembly of a reference taxonomy for phylogenetic
767 data synthesis. *Biodiversity Data Journal*, (5).
- 768 Rees, Vandepitte, L., Decock, W., & Vanhoorne, B. (2017). IRMNG 2006–2016: 10 Years of
769 a Global Taxonomic Database. *Biodiversity Informatics*, 12.
- 770 Revell, L. J. (2012). Phytools: An r package for phylogenetic comparative biology (and other
771 things). *Methods in Ecology and Evolution*, 3, 217–223.
- 772 Ronquist, F., & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference
773 under mixed models. *Bioinformatics*, 19(12), 1572–1574.
774 <https://doi.org/10.1093/bioinformatics/btg180>
- 775 Roquet, C., Lavergne, S., & Thuiller, W. (2014). One tree to link them all: A phylogenetic
776 dataset for the european tetrapoda. *PLoS Currents*, 6.

- 777 Sanchez-Reyes, L. L., & O'Meara, B. (2022). *datelifeplot*: Methods to plot chronograms
778 and outputs of the *datelife* package. *R Package Release V0.2.2*. Retrieved from
779 <https://zenodo.org/badge/latestdoi/381501451>
- 780 Sanchez-Reyes, L. L., O'Meara, B., Eastman, J., Heath, T., Wright, A., Schliep, K., ...
781 Alfaro, M. (2022). *datelife*: Scientific Data on Time of Lineage Divergence for Your
782 Taxa. In *R package version 0.6.6*. Retrieved from
783 <https://CRAN.R-project.org/package=datelife> and
784 <https://doi.org/10.5281/zenodo.593938>
- 785 Sanderson, M. (2002). Estimating Absolute Rates of Molecular Evolution and Divergence
786 Times: A Penalized Likelihood Approach. *Molecular Biology and Evolution*, 19(1),
787 101–109. <https://doi.org/10.1093/oxfordjournals.molbev.a003974>
- 788 Sanderson, M. (2003). r8s: Inferring Absolute Rates of Molecular Evolution and Divergence
789 Times in the Absence of a Molecular Clock. *Bioinformatics*, 19(2), 301–302.
- 790 Sanderson, M., & Doyle, J. (2001). Sources of error and confidence intervals in estimating
791 the age of angiosperms from *rbcL* and 18S rDNA data. *American Journal of Botany*,
792 88(8), 1499–1516.
- 793 Sauquet, H. (2013). A practical guide to molecular dating. *Comptes Rendus Palevol*, 12(6),
794 355–367.
- 795 Sauquet, H., Ramírez-Barahona, S., & Magallón, S. (2021). *The age of flowering plants is
796 unknown*.
- 797 Schenk, J. J. (2016). Consequences of secondary calibrations on divergence time estimates.
798 *PLoS ONE*, 11(1). <https://doi.org/10.1371/journal.pone.0148228>
- 799 Schliep, K. P. (2011). Phangorn: Phylogenetic analysis in r. *Bioinformatics*, 27(4), 592–593.

- 800 Schoch, C. L., Ciufo, S., Domrachev, M., Hotton, C. L., Kannan, S., Khovanskaya, R., ...
801 others. (2020). NCBI Taxonomy: a Comprehensive Update on Curation, Resources
802 and Tools. *Database*, 2020.
- 803 Selvatti, A. P., Gonzaga, L. P., & Moraes Russo, C. A. de. (2015). A paleogene origin for
804 crown passerines and the diversification of the oscines in the new world. *Molecular
805 Phylogenetics and Evolution*, 88, 1–15.
- 806 Smith, S., & Brown, J. (2018). Constructing a broadly inclusive seed plant phylogeny.
807 *American Journal of Botany*, 105(3), 302–314.
- 808 Smith, S., & O'Meara, B. (2012). TreePL: Divergence time estimation using penalized
809 likelihood for large phylogenies. *Bioinformatics*, 28(20), 2689–2690.
810 <https://doi.org/10.1093/bioinformatics/bts492>
- 811 Stoltzfus, A., Lapp, H., Matasci, N., Deus, H., Sidlauskas, B., Zmasek, C. M., ... Jordan, G.
812 (2013). Phylotastic! Making tree-of-life knowledge accessible, reusable and convenient.
813 *BMC Bioinformatics*, 14. <https://doi.org/10.1186/1471-2105-14-158>
- 814 Sun, M., Folk, R. A., Gitzendanner, M. A., Soltis, P. S., Chen, Z., Soltis, D. E., & Guralnick,
815 R. P. (2020). Estimating rates and patterns of diversification with incomplete
816 sampling: A case study in the rosids. *American Journal of Botany*, 107(6), 895–909.
- 817 Tietze, D. T., Päckert, M., Martens, J., Lehmann, H., & Sun, Y.-H. (2013). Complete
818 phylogeny and historical biogeography of true rosefinches (aves: Carpodacus).
819 *Zoological Journal of the Linnean Society*, 169(1), 215–234.
- 820 Title, P. O., & Rabosky, D. L. (2016). Do Macrophylogenies Yield Stable Macroevolutionary
821 Inferences? An Example from Squamate Reptiles. *Systematic Biology*, syw102.
822 <https://doi.org/10.1093/sysbio/syw102>

- 823 Treplin, S., Siegert, R., Bleidorn, C., Thompson, H. S., Fotso, R., & Tiedemann, R. (2008).
824 Molecular phylogeny of songbirds (aves: Passeriformes) and the relative utility of
825 common nuclear marker loci. *Cladistics*, 24(3), 328–349.
- 826 Uyeda, J. C., Pennell, M. W., Miller, E. T., Maia, R., & McClain, C. R. (2017). The
827 evolution of energetic scaling across the vertebrate tree of life. *The American
828 Naturalist*, 190(2), 185–199.
- 829 Vos, R. A., Balhoff, J. P., Caravas, J. A., Holder, M. T., Lapp, H., Maddison, W. P., ...
830 others. (2012). NeXML: Rich, extensible, and verifiable representation of
831 comparative data and metadata. *Systematic Biology*, 61(4), 675–689.
832 <https://doi.org/10.1093/sysbio/sys025>
- 833 Webb, C. (2000). Exploring the Phylogenetic Structure of Ecological Communities : An
834 Example for Rain Forest Trees. *The American Naturalist*, 156(2), 145–155.
- 835 Webb, C., Ackerly, D., & Kembel, S. (2008). Phylocom: Software for the analysis of
836 phylogenetic community structure and trait evolution. *Bioinformatics*, 24(18),
837 2098–2100. <https://doi.org/10.1093/bioinformatics/btn358>
- 838 Webb, C., & Donoghue, M. (2005). Phylomatic: Tree assembly for applied phylogenetics.
839 *Molecular Ecology Notes*, 5(1), 181–183.
- 840 Weir, J., & Schluter, D. (2008). Calibrating the avian molecular clock. *Molecular Ecology*,
841 17(10), 2321–2328.
- 842 Zuccon, D., Prŷs-Jones, R., Rasmussen, P. C., & Ericson, P. G. (2012). The phylogenetic
843 relationships and generic limits of finches (fringillidae). *Molecular Phylogenetics and
844 Evolution*, 62(2), 581–596.

845 TABLE 1. Ages of congruified nodes. See Figure 3, step C2.

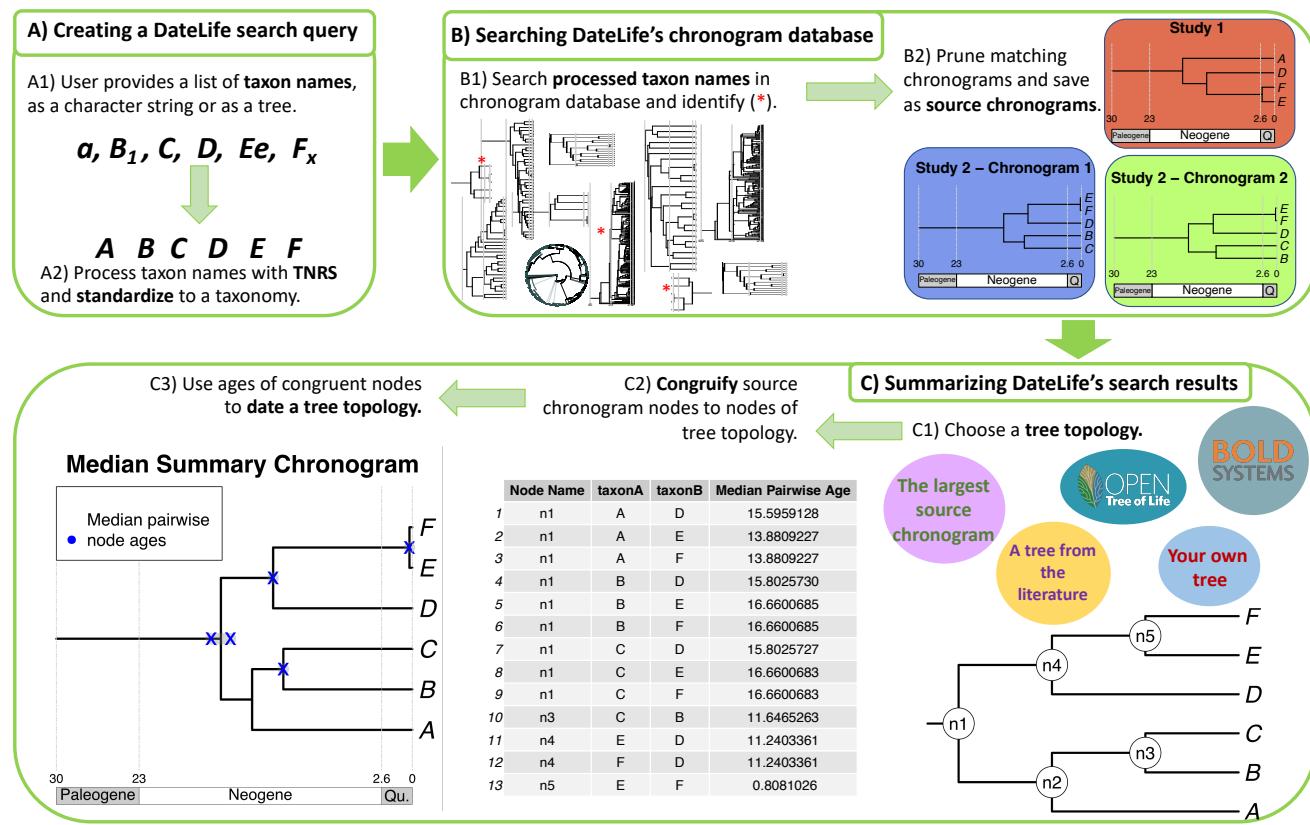
	Node Name	taxon A	taxon B	Node Age	Study chronogram
1	n1	<i>Emberiza citrinella</i>	<i>Platyspiza crassirostris</i>	9.6509881	Jetz et al. 2012 – chronogram 1
2	n1	<i>Pheucticus tibialis</i>	<i>Schoeniclus elegans</i>	14.3336520	Barker et al. 2015 – chronogram 1
3	n1	<i>Rhodothraupis celaeno</i>	<i>Schoeniclus elegans</i>	16.2984859	Barker et al. 2015 – chronogram 2
4	n1	<i>Emberiza citrinella</i>	<i>Platyspiza crassirostris</i>	16.9499615	Hooper et al. 2017
5	n1	<i>Rhodothraupis celaeno</i>	<i>Schoeniclus elegans</i>	21.5140867	Jetz et al. 2012 – chronogram 2
6	n1	<i>Emberiza leucocephalos</i>	<i>Platyspiza crassirostris</i>	22.0000000	Hedges et al. 2015 – chronogram 2
7	n1	<i>Emberiza citrinella</i>	<i>Platyspiza crassirostris</i>	22.3757277	Hedges et al. 2015 – chronogram 1
8	n2	<i>Platyspiza crassirostris</i>	<i>Rhodothraupis celaeno</i>	7.9691925	Jetz et al. 2012 – chronogram 1
9	n2	<i>Platyspiza crassirostris</i>	<i>Rhodothraupis celaeno</i>	19.7085830	Jetz et al. 2012 – chronogram 2
10	n2	<i>Platyspiza crassirostris</i>	<i>Rhodothraupis celaeno</i>	19.7085900	Hedges et al. 2015 – chronogram 2
11	n2	<i>Platyspiza crassirostris</i>	<i>Rhodothraupis celaeno</i>	19.7128363	Hedges et al. 2015 – chronogram 1
12	n2	<i>Pheucticus tibialis</i>	<i>Platyspiza crassirostris</i>	44.2958603	Burns et al. 2014
13	n3	<i>Pheucticus tibialis</i>	<i>Rhodothraupis celaeno</i>	10.5304440	Barker et al. 2015 – chronogram 1
14	n3	<i>Pheucticus tibialis</i>	<i>Rhodothraupis celaeno</i>	10.5379092	Barker et al. 2013
15	n3	<i>Pheucticus tibialis</i>	<i>Rhodothraupis celaeno</i>	11.2095375	Barker et al. 2015 – chronogram 2
16	n3	<i>Pheucticus tibialis</i>	<i>Rhodothraupis celaeno</i>	18.1570685	Hedges et al. 2015 – chronogram 1
17	n4	<i>Emberiza citrinella</i>	<i>Schoeniclus elegans</i>	5.3238969	Jetz et al. 2012 – chronogram 1
18	n4	<i>Emberiza leucocephalos</i>	<i>Schoeniclus elegans</i>	9.8622460	Barker et al. 2015 – chronogram 1
19	n4	<i>Emberiza leucocephalos</i>	<i>Schoeniclus elegans</i>	10.3391445	Jetz et al. 2012 – chronogram 2
20	n4	<i>Emberiza leucocephalos</i>	<i>Schoeniclus elegans</i>	11.7317630	Barker et al. 2015 – chronogram 2
21	n4	<i>Emberiza citrinella</i>	<i>Schoeniclus elegans</i>	12.5133870	Hooper et al. 2017
22	n5	<i>Emberiza citrinella</i>	<i>Emberiza leucocephalos</i>	0.1407015	Jetz et al. 2012 – chronogram 1
23	n5	<i>Emberiza citrinella</i>	<i>Emberiza leucocephalos</i>	0.1516230	Hedges et al. 2015 – chronogram 2
24	n5	<i>Emberiza citrinella</i>	<i>Emberiza leucocephalos</i>	0.2011990	Barker et al. 2015 – chronogram 1
25	n5	<i>Emberiza citrinella</i>	<i>Emberiza leucocephalos</i>	0.2409300	Barker et al. 2015 – chronogram 2
26	n5	<i>Emberiza citrinella</i>	<i>Emberiza leucocephalos</i>	0.2732460	Jetz et al. 2012 – chronogram 2
27	n5	<i>Emberiza citrinella</i>	<i>Emberiza leucocephalos</i>	0.5760260	Hedges et al. 2015 – chronogram 1
28	n5	<i>Emberiza citrinella</i>	<i>Emberiza leucocephalos</i>	2.2898230	Hooper et al. 2017

846

847 TABLE 2. Summary of congruified nodes ages. See Figure 3, step C3.

Node Name	Min Age	Q1	Median Age	Mean Age	Q3	Max Age	Variance	SD
n1	9.6509881	15.316069	16.94996	17.5889860	21.757043	22.375728	22.2431847	4.7162681
n2	7.9691925	19.708583	19.70859	22.2790124	19.712836	44.295860	177.3279940	13.3164558
n3	10.5304440	10.536043	10.87372	12.6087398	12.946420	18.157069	13.7831237	3.7125630
n4	5.3238969	9.862246	10.33914	9.9540875	11.731763	12.513387	7.8263782	2.7975665
n5	0.1407015	0.176411	0.24093	0.5533641	0.424636	2.289823	0.6079318	0.7796998

848



849 FIGURE 1. Main DateLife workflow. Analyses can be performed via DateLife's interactive
850 website at www.datelife.org, or using the `datelife` R package. Details on the R functions used to
851 perform the analyses are available from `datelife`'s R package vignettes at
852 <https://phylotastic.org/datelife>.

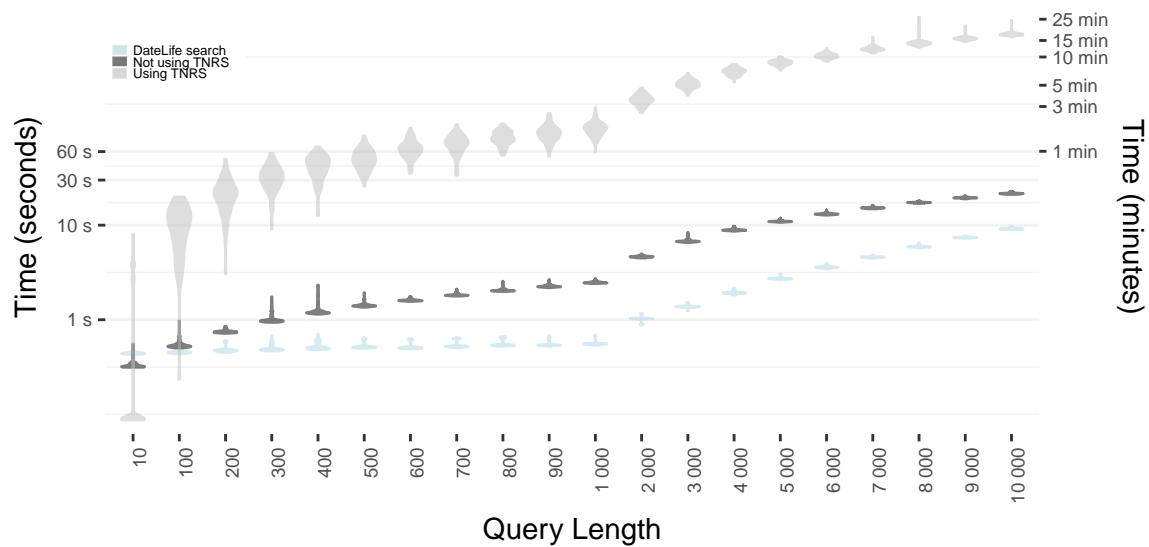
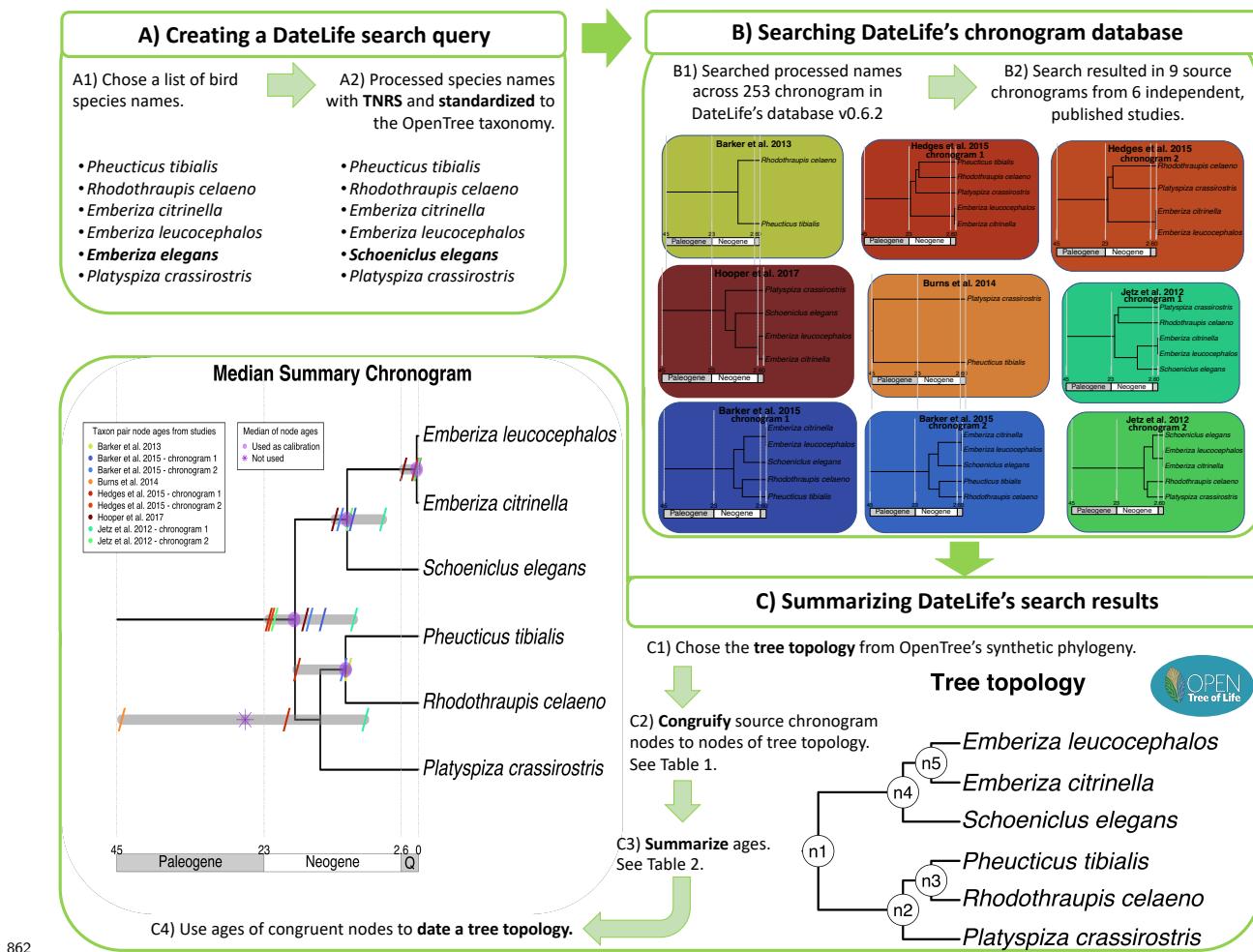


FIGURE 2. *DateLife's* benchmarking results. Computation time used to process a query and a search across *datelife's* chronogram database, relative to number of input taxon names. For each $N = \{10, 100, 200, \dots, 1\,000, \dots, 9\,000, 10\,000\}$, we sampled N species names from the class Aves a hundred times, and then performed a *datelife* search processing the input names with Taxon Names Resolution Service (TNRS; light gray), and without processing names (dark gray). For comparison, we performed a search using an input that had been pre-processed with TNRS (light blue).



863 FIGURE 3. *DateLife analysis results for a small sample of A) 6 bird species within the*
864 *Passeriformes. B) Processed species names were found across 9 chronograms within 6*
865 *independent studies (Barker et al. (2012), Barker et al. (2015), Burns et al. (2014), Hedges*
866 *et al. (2015), Hooper and Price (2017), Jetz et al. (2012).) C) This revealed 28 source age*
867 *data points for the queried species names. Summarized age data is used as secondary*
868 *calibrations to date a tree topology obtained from OpenTree's synthetic tree, resulting in a*
869 *summary chronogram of source ages.*

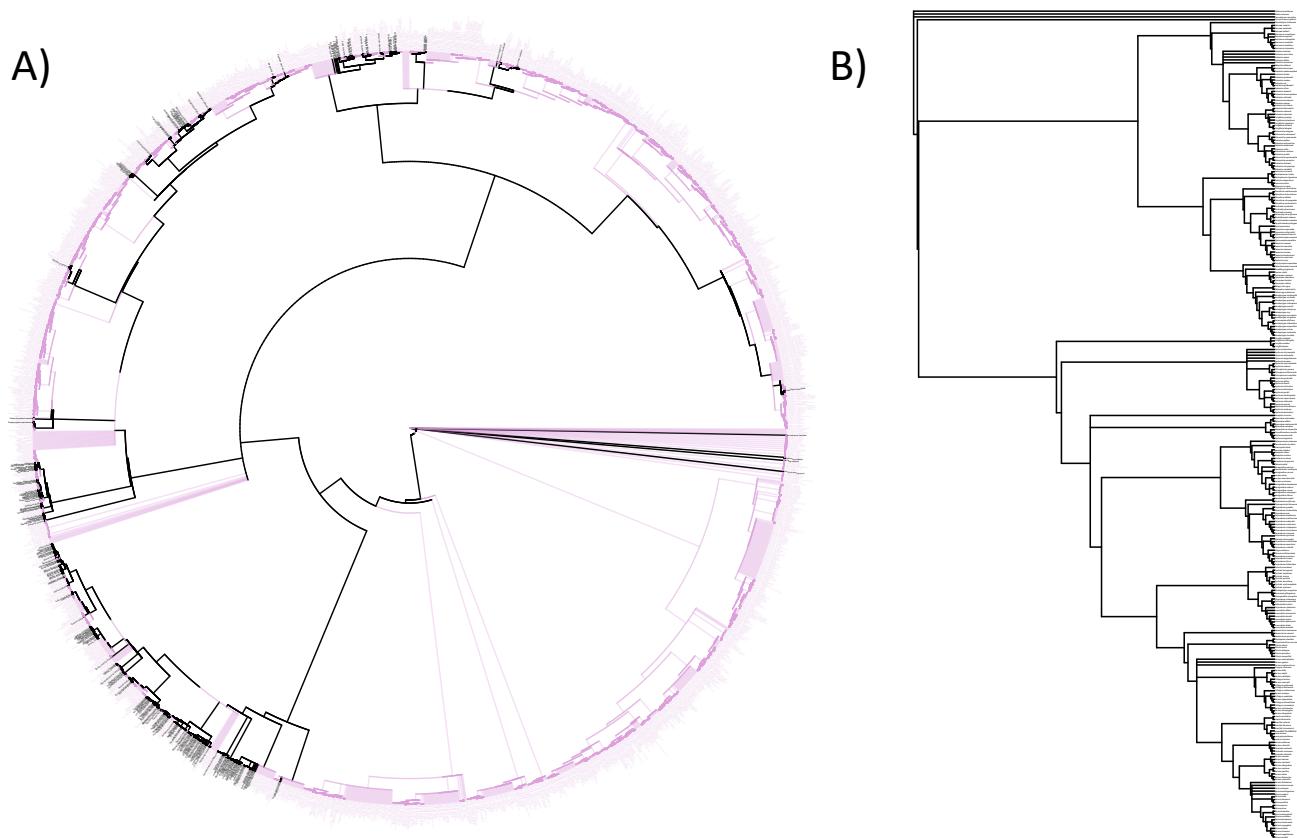
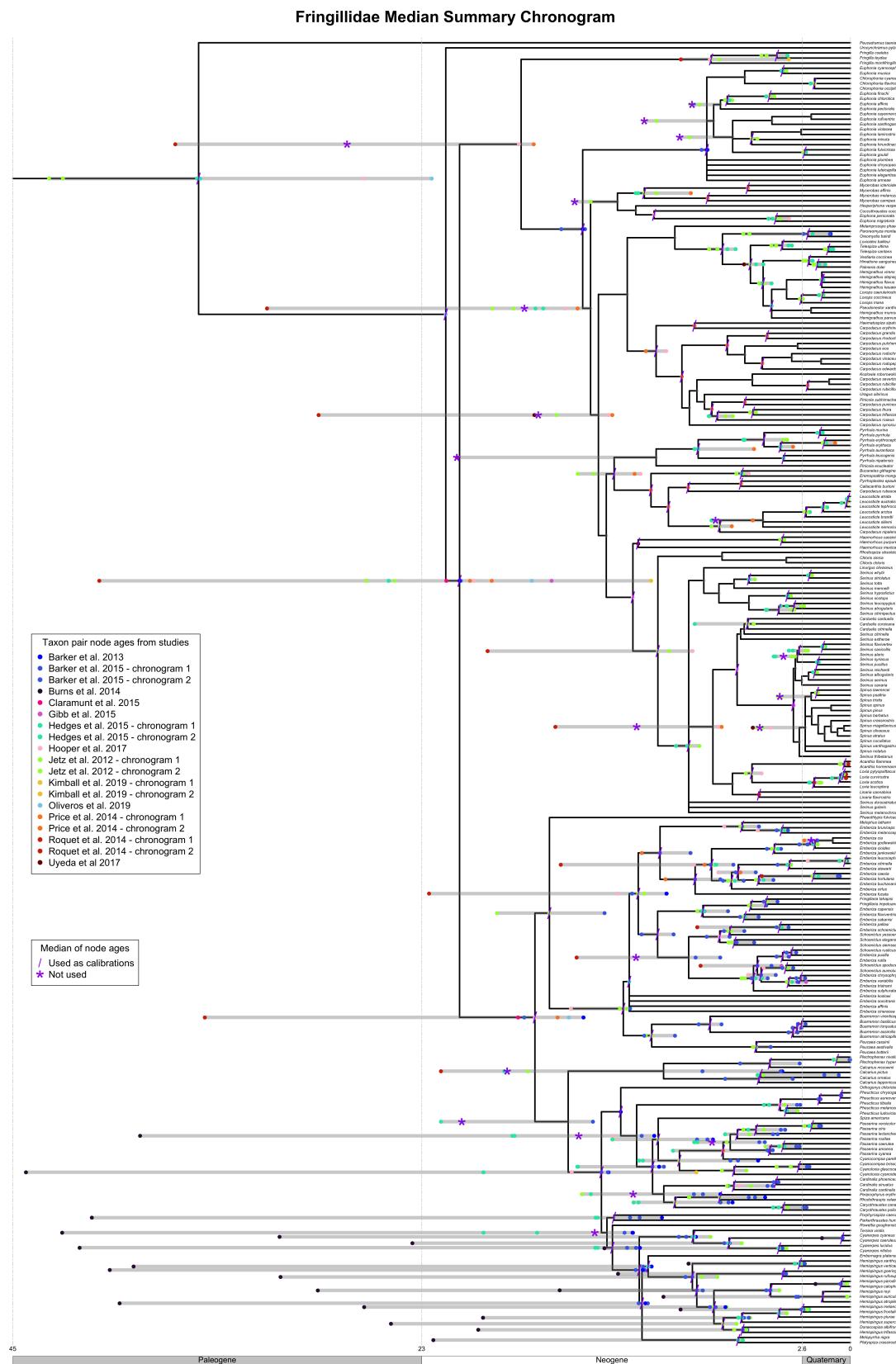
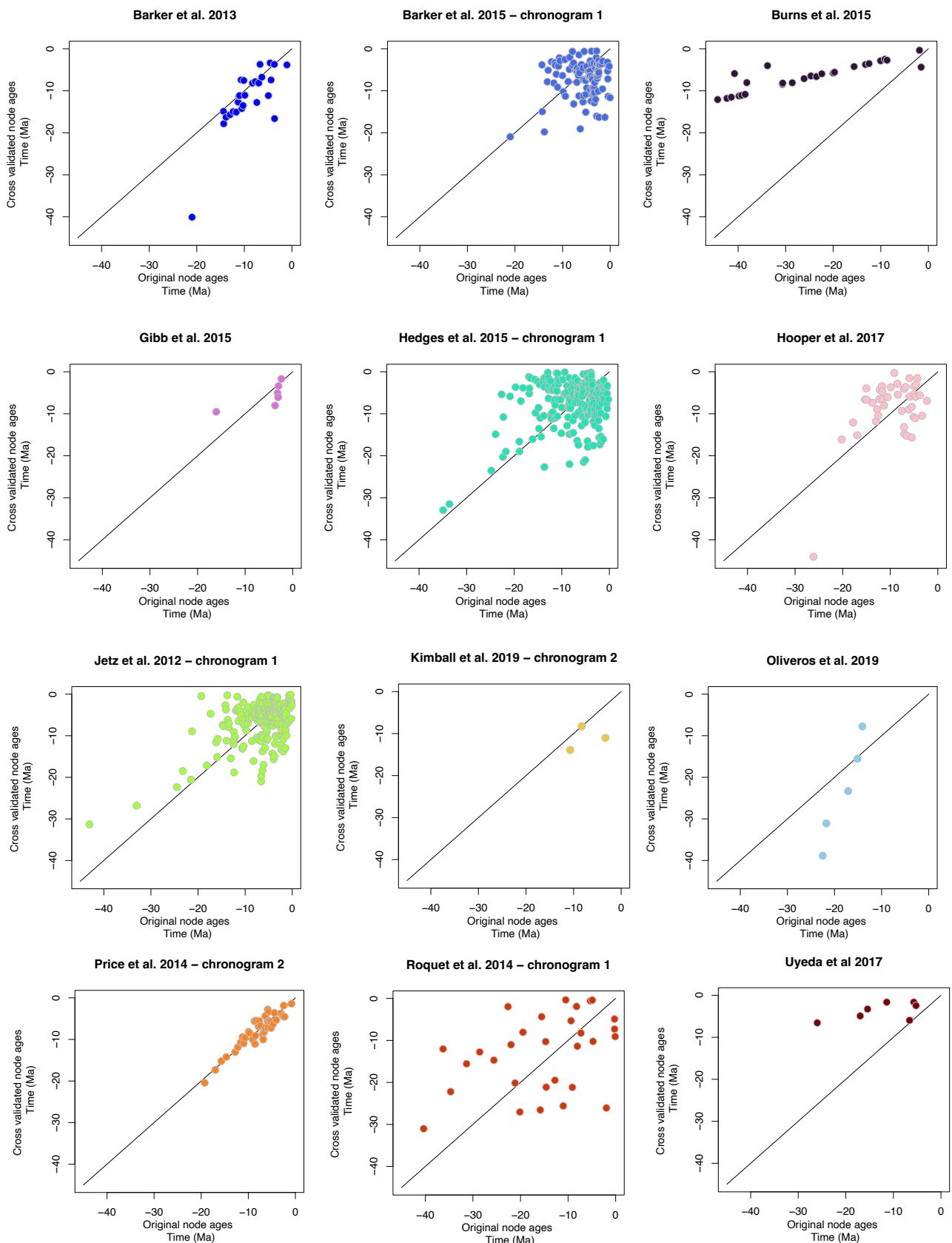


FIGURE 4. Tree topologies extracted from Open Tree of Life's (OpenTree) synthetic phylogenetic tree.
 870 A) Topology of 2,333 tips and 1,305 internal nodes, encompassing bird species within the family
 871 Fringillidae following the NCBI taxonomy (black), as well as all other bird species that share the same
 872 Most Recent Common Ancestor (MRCA) node in OpenTree's synthetic tree (purple). B) Subtree
 873 topology of 289 tips and 253 internal nodes, resulting from pruning species that do not belong to the
 874 family Fringillidae according to the NCBI taxonomy (purple branches in topology A). Bird species
 875 within the Fringillidae are paraphyletic (Alström et al. 2014, Barker, Cibois, Schikler, Feinstein, &
 876 Cracraft 2004, Barker et al. 2013, Barker 2014, Barker et al. 2015, Beresford, Barker, Ryan, &
 877 Crowe 2005, Bryson Jr et al. 2014, Burleigh, Kimball, & Braun 2015, Burns et al. 2014, Chaves,
 878 Hidalgo, & Klicka 2013, Claramunt & Cracraft 2015, Gibb et al. 2015, Hackett et al. 2008, Jetz et al.
 879 2012, Johansson, Fjeldså, & Bowi 200, Kimball et al. 2019, Klicka et al. 2014, Lamichhaney et al.
 880 2015, Lerner, Meyer, James, Hofreiter, & Fleischer 2011, Lovette et al. 2010, Moyle et al. 2016,
 881 Ödeen, Håstad, & Alström 2011, Oliveros et al. 2019, Päckert et al. 2012, Parchman, Benkman, &
 882 Mezquida 2007, Powell et al. 2014, Price et al. 2014, Pulgarín-R, Smith, Bryson Jr, Spellman, &
 883 Klicka 2013, Selvatti, Gonzaga, & Moraes Russo 2015, Tietze, Päckert, Martens, Lehmann, & Sun
 884 2013, Treplin et al. 2008, Zuccon, Prŷs-Jones, Rasmussen, & Ericson 2012).



892

FIGURE 5. *Fringillidae median summary chronogram generated with DateLife. It has 289 tips and 253 nodes, from which 212 have age data from at least one published chronogram. These nodes have been split into 10 different studies, from 12 different studies (including this), which*



893 FIGURE 6. Results from cross validation analysis. Each plot compares the original age estimate (x
 894 axis) with the age obtained with a DateLife analysis (y axis), per node.