1   **Molecular Ecology Resources**

2   **(submitted)**

3

4

5

6

7

8   **Molecular dating for phylogenies containing a mix of populations and species**

9

10   Beatriz Mello[1,2], Qiqing Tao[2,3] and Sudhir Kumar[2,3]

11

12

13

14

15

16   [1]Department of Genetics, Federal University of Rio de Janeiro, RJ, Brazil

17   [2]Institute for Genomics and Evolutionary Medicine, Temple University, Philadelphia,

18   PA, USA

19   [3]Department of Biology, Temple University, Philadelphia, PA, USA

20

21

22

23

24   **Running title**: Dating datasets with mixed sampling

25

26

27

28   *Correspondence to:*

29   Sudhir Kumar

30   Temple University

31   Philadelphia, PA 19122, USA

32   E-mail: s.kumar@temple.edu

Abstract

Concurrent molecular dating of population and species divergences is essential in many biological investigations, including phylogeography, phylodynamics, and species delimitation studies. Multiple sequence alignments used in these investigations frequently consist of both intra- and inter-species samples (mixed samples). As a result, the phylogenetic trees contain inter-species, inter-population, and within population divergences. To date these sequence divergences, Bayesian relaxed clock methods are often employed, but they assume the same tree prior for both inter- and intra-species branching processes and require specification of a clock model for branch rates (independent vs. autocorrelated rates models). We evaluated the impact of using the same tree prior on the Bayesian divergence time estimates by analyzing computer-simulated datasets. We also examined the effect of the assumption of independence of evolutionary rate variation among branches when the branch rates are autocorrelated. Bayesian approach with Skyline-coalescent tree priors generally produced excellent molecular dates, with some tree priors (e.g., Yule) performing the best when evolutionary rates were autocorrelated, and lineage sorting was incomplete. We compared the performance of the Bayesian approach with a non-Bayesian, the RelTime method, which does not require specification of a tree prior or selection of a clock model. We found that RelTime performed as well as the Bayesian approach, and when the clock model was mis-specified, RelTime performed slightly better. These results suggest that the computationally efficient RelTime approach is also suitable to analyze datasets containing both populations and species variation.

Keywords: tree prior, divergence times, RelTime, Bayesian, molecular dating

2

Introduction

63 Divergence times derived from molecular data have become critical for elucidating
64 earth's historical processes that have shaped the evolution of life (Hedges, Marin,
65 Suleski, Paymer, & Kumar, 2015; Ho, 2014; O'Reilly, dos Reis, & Donoghue, 2015).
66 Recent biological timescales of closely-related taxa are generally built in the context of
67 phylogeographic, phylodynamics and species delimitation studies (Esselstyn, Evans,
68 Sedlock, Anwarali Khan, & Heaney, 2012; McCluskey & Postlethwait, 2015; Mello,
69 Vilela, & Schrago, 2018; C. Wang, Shikano, Persat, & Merilä, 2015). Divergence times
70 are usually estimated from molecular data with multiple individuals per species
71 sampled. Consequently, both inter- and intra-species divergences (nodes) are present
72 in the same phylogenetic tree, such that both speciation and population processes are
73 at play. Such datasets are becoming increasingly common (Edwards, Shultz, &
74 Campbell-Staton, 2015; Lemmon & Lemmon, 2012; Manzanilla et al., 2018;
75 McCormack, Hird, Zellmer, Carstens, & Brumfield, 2013; Melville et al., 2017; Merwe,
76 McPherson, Siow, & Rossetto, 2014).

77 In order to accurately infer divergence times in datasets with a mixture of micro-
78 and macro-evolutionary events, a multispecies coalescent (MSC) approach would be
79 highly appropriate, because it explicitly accounts for conflicts between gene
80 genealogies and the species tree by modelling incomplete lineage sorting (ILS) across
81 lineages (Degnan & Salter, 2005; Edwards & Beerli, 2000; Heled & Drummond, 2010;
82 Maddison, 1997; Rannala & Yang, 2003; Takahata, 1989; Takahata & Nei, 1985; Yang
83 & Rannala, 2014). However, MSC is not often applied because computational times
84 required to estimate divergence times are still impractical for contemporary data sizes,
85 even though some recent advances have been made to reduce the computational
86 burden (Ogilvie, Bouckaert, & Drummond, 2017; Ogilvie, Heled, Xie, & Drummond,
87 2016; Rannala & Yang, 2017; Xu & Yang, 2016). Moreover, it currently requires *a priori*
88 delimitation of species limits and, for some implementations, the assumption of strict
89 molecular clocks (Heled & Drummond, 2010; Xu & Yang, 2016; Yang, 2015).

90 Instead, researchers frequently use the standard Bayesian framework, mainly
91 the BEAST software (Bouckaert et al., 2014; dos Reis & Yang, 2011). BEAST software
92 requires specification of priors, including models that describe the branching pattern
93 (tree prior) and that assume the absence of autocorrelated branch rates (ABR). The

3

94   commonly used tree priors in BEAST analyses are those that describe speciation

95   processes, e.g., Yule (Yule, 1924) and birth and death (BD) (Feller, 1939; Kendall,

96   1948), and within-species population processes, e.g., coalescent priors considering a

97   constant population size or an exponential growth (Kuhner, Yamato, & Felsenstein,

98   1995, 1998). Yule and BD priors model macro-evolutionary events, so their use to

99   describe intra-species divergences would improperly assume these divergences to be

100   like speciation events rather than due to intra-population coalescence. On the other

101   hand, the use of coalescent priors may bias time estimation of ancient inter-species

102   divergences, like the tMRCA (time to the most recent common ancestor) of a whole

103   genus or several extant species.

104   Currently, in the standard Bayesian molecular dating, there is no composite tree

105   prior that account for macro- and micro-evolutionary processes in estimating

106   divergence times. Therefore, it is critical to investigate whether the divergence times

107   produced by Bayesian methods are robust to the selection of tree prior and to the

108   application of only one tree prior for trees with mixed samples. Also, the relaxed clock

109   method in BEAST assumes that there is no autocorrelation of evolutionary rates

110   among branches. However, recent evidence suggests that autocorrelated branch rates

111   are likely to be the norm (Tao, Tamura, Battistuzzi, & Kumar, 2019), so the impact of

112   rate model misspecification in using BEAST needs to be evaluated.

113   Ritchie et al. (2017) have already conducted some simulations to evaluate the

114   impact of tree priors on time estimates in BEAST. In their study, they only considered

115   the case where data were simulated under an independent branch rate (IBR) model

116   and with incomplete lineage sorting (ILS) (**Table 1**). Therefore, we do not yet know the

117   impact of using a single tree prior on time estimates in BEAST under many other

118   scenarios, such as an ABR model and the absence of ILS. Furthermore, Ritchie et al.

119   (2017) did not evaluate the performance of any non-Bayesian method. For example,

120   RelTime has been shown to work well for trees consisting exclusively of interspecies

121   divergences (Mello, Tao, Tamura, & Kumar, 2017; Tamura, Tao, & Kumar, 2018), so

122   its absolute and comparative performance will inform about its suitability for use in

123   datasets containing both inter- and intra-species sampling. Thus, we conducted many

124   simulated analyses that present a significant advance beyond Ritchie et al. (2017)

4

125 (**Table 1**). They uniquely provide insights into many questions that are faced by

126 practitioners in molecular dating studies.

127

128 **Material and methods**

129 We analyzed simulated datasets to test the performance of a Bayesian approach

130 (BEAST) (Drummond, Ho, Phillips, & Rambaut, 2006) and RelTime method (Kumar,

131 Stecher, Li, Knyaz, & Tamura, 2018; Tamura et al., 2012). Simulations provide

132 valuable means to test the reliability of inferred node ages because the actual time is

133 known, and the impact of many biologically realistic conditions can be explored,

134 including the presence/absence of autocorrelation of branch rates across the tree

135 (ABR vs. IBR model) and incomplete lineage sorting (ILS vs. no-ILS) (**Fig. 1**).

136 *Computer Simulated Datasets*

137 We considered two distinct scenarios to simulate datasets with a mixed sampling of

138 intra- and inter-species sequences. First, we assumed that there was complete lineage

139 sorting (no-ILS), which represents the baseline scenario for understanding sources of

140 error. In no-ILS phylogenies, all species are reciprocally monophyletic and speciation

141 times are concordant with the divergence of sequences (e.g., **Fig. 1a**-**c**). Second,

142 genealogies were simulated within species phylogeny under the MSC approach, which

143 resulted in ILS. In this case, the sequence phylogeny is not fully concordant with the

144 species tree, and species are not automatically monophyletic (e.g., **Fig. 1d**-**f**). The ILS

145 scenario allowed us to assess the performance of methods under more complex

146 evolutionary histories.

147   The no-ILS phylogenies contained ten species and were simulated assuming a

148 Yule process with $\lambda = 0.11$ (speciation rate) and a root height of 10 million years (Ma)

149 in the TreeSim R package (Stadler, 2011). Then, the *ms* function from phyclust R

150 package, which generates coalescent trees under a modified Hudson (2002)'s neutral

151 model, was used to simulate coalescent trees under constant population size with ten

152 samples ($k = 10$) (Chen, 2011; Hudson, 2002). Assumed generation time was 20 years,

153 and effective population size (N_e) was 50,000. Coalescent trees were then "pasted" at

154 each tip (species) of the phylogeny while keeping the expected phylogeny ultrametric

155 because all the sequences were contemporaneous. For phylogenies with ILS, the

156 same procedure and parameters were used to estimate species phylogenies (Yule

157   process with ten species, λ = 0.11, and root height = 10 Ma). Then, single gene

158   genealogies were simulated in Phybase R package (Liu & Yu, 2010) with a constant

159   scaled population parameter (theta) of 0.014 and ten individual samples per species,

160   which is equivalent to the configuration used to simulate coalescent trees in the

161   scenario without ILS. An evolutionary rate of 3.5×10^{-9} substitutions/site/year (Endicott

162   & Ho, 2008) was used, and one hundred datasets were obtained for both ILS and no-

163   ILS analysis.

164   We simulated three different types of evolutionary rate variation among

165   branches: constant branch rates (CBR, **Fig. 1a** and **d**), independent branch rates (IBR,

166   **Fig. 1b** and **e**), and autocorrelated branch rates (ABR, **Fig. 1c** and **f**). In CBR, the

167   evolutionary rate of 3.5×10^{-9} substitutions/site/year was applied to all branches of the

168   phylogenies. In IBR, branch rates were allowed to randomly vary as much as fifty

169   percent of the mean (Tamura et al., 2012). In ABR, we used modified functions from

170   NELSI R package (Ho, Duchêne, & Duchêne, 2015) using 3.5×10^{-9}

171   substitutions/site/year as the initial rate and a correlation parameter (ν) of 0.01

172   (Kishino, Thorne, & Bruno, 2001). Sequences were simulated in seq-gen (Rambaut &

173   Grass, 1997) under HKY substitution model (Hasegawa, Kishino, & Yano, 1985) to

174   generate alignments of 10,000 sites with equilibrium frequencies and

175   transition/transversion ratios sampled from an empirical distribution (Rosenberg &

176   Kumar, 2003). Thus, we simulated six scenarios: CBR – no-ILS, IBR – no-ILS, ABR –

177   no-ILS, CRB – ILS, IBR – ILS, and ABR – ILS, each of which contains 100 simulated

178   datasets.

179   *Analysis of simulated data*

180   Bayesian divergence times were estimated by using BEAST v2.4.7 (Bouckaert et al.,

181   2014) and RelTime estimates were obtained using MEGA 7 (Kumar, Stecher, &

182   Tamura, 2016; Tamura et al., 2012). Correct substitution model (HKY) and topologies

183   were used in all the dating analyses to prevent confounding errors from phylogeny

184   inference and divergence time estimation.

185   In BEAST, the root height was calibrated with a normal prior density with mean

186   1.0 and standard deviation of 0.05. This was done to exclude the effects of multiple

187   calibrations and the interaction between them, which can impact the results

188   considerably (Battistuzzi, Billing-Ross, Murillo, Filipski, & Kumar, 2015; dos Reis et al.,

189 2018; dos Reis et al., 2015; Duchêne, Duchêne, Holmes, & Ho, 2015). Node ages
190 were inferred based on the Yule (pure birth, BEAST-Yule) and birth-death (BEAST-
191 BD) speciation processes in BEAST, both with default settings, in order to test the
192 performance of distinct Bayesian tree priors. BEAST was also run with coalescent
193 priors, namely the constant population size (BEAST-Constant) and the skyline
194 coalescent (BEAST-Skyline) with default settings (Drummond, Rambaut, Shapiro, &
195 Pybus, 2005). In all analyses, ESS (effective sample sizes) values were checked with
196 the function *effectiveSize* from coda R package (Plummer, Best, Cowles, & Vines,
197 2006), after discarding the burn-in period.

198   In RelTime, one does not need to specify a tree prior, a clock model for
199 evolutionary rates, or a statistical distribution to describe the heterogeneity of branch
200 rates (Tamura et al., 2012, 2018). RelTime computes relative times and lineage rates
201 directly from the branch lengths that are inferred from the molecular sequences
202 (Tamura et al., 2012, 2018). Calibrations are also not a prerequisite to estimating
203 divergence times, and so all RelTime inferred times are relative to the height of the
204 ingroup root node.

205 *Measures*
206 To compare estimated and true times, we report two primary metrics. One is the
207 normalized time difference (Δt), which was computed for each node in every phylogeny
208 analyzed. Δt is the difference between the estimated and the true divergence time
209 divided by the true time. For no-ILS phylogenies, we report Δt for within-population,
210 coalescent, and interspecies comparisons in order to evaluate performance for these
211 three distinct evolutionary levels. However, such a distinction is not possible for ILS
212 datasets. Therefore, we also conducted a linear regression between all the estimated
213 and true times in a phylogeny. The slope of the linear regression through the origin is
214 referred to as the time slope, which is reported alongside the standard R^2 statistic.

215

216 **Results**
217 We carried out a total of 2,400 BEAST analyses for 600 simulated datasets, as four
218 different tree priors were applied for each dataset. The RelTime analysis was
219 conducted only once for each dataset, as RelTime does not require specification of a
220 tree prior. RelTime calculations completed in less than 2 minutes for every dataset, but

221 BEAST analyses took orders of magnitude longer (e.g. Battistuzzi, Billing-Ross,

222 Paliwal, & Kumar, 2011). Also, several BEAST's MCMC calculations did not reach

223 acceptable ESS values (≥200) even with chain lengths that exceeded one hundred

224 million generations and took ~12 hours to run on an Intel Core i7® iMac @ 4GHz

225 machine. This problem was encountered frequently in BEAST-Yule analyses; they

226 failed for 15% and 32% of no-ILS datasets for IBR and ABR models, respectively. For

227 ILS datasets, 4%, 5% and 3% of ABR replicates in BEAST-Yule, BEAST-BD and

228 BEAST-Skyline did not reach ESS values greater than 200. We excluded these

229 datasets from further consideration and present the results based on the remaining

230 datasets for BEAST analyses.

231 *Performance for Constant Rate Phylogenies*

232 Biologically, no-ILS phylogenies in which sequence evolved with a constant rate (CBR

233 model) represent the simplest, baseline scenario. For these datasets, distributions of

234 normalized difference between true and estimated time (Δt) for within-population

235 divergences were centered around zero for BEAST (**Fig. 2a**). BEAST-BD and BEAST-

236 Skyline performed the best, and BEAST-Constant showed a slight tendency to

237 overestimate times (**Fig. 2a**). BEAST-Yule overestimated divergence times with much

238 higher median Δt than other tree priors (**Fig. 3a**). In contrast, RelTime showed a

239 tendency to underestimate within-population divergences (median $\Delta t < 0$; **Fig. 2a**),

240 because it assigns a time equal to 0 for nodes at which the tip sequences show a small

241 amount of difference.

242   In contrast, BEAST tree priors assign a non-zero time to all the branches (Marin

243 & Hedges, 2018). In our simulations, all the expected divergence times were non-zero,

244 so the difference between the small non-zero divergence times assigned by BEAST

245 and the expected values was smaller than the difference between the zero times

246 assigned by RelTime and the expected values. However, this strategy would result in

247 overestimation of times in BEAST if some of the sampled sequences were truly

248 identical or the sequence divergence was very small. For example, Δt for nodes < 0.1%

249 of the root (ingroup) height was always larger than 0 for BEAST, but it was smaller for

250 RelTime (**Fig. 4**). Consistent with this pattern is the observation that RelTime produced

251 fewer outlier Δt values as compared to BEAST (numbers in parentheses in **Fig. 2a** and

252 **3a**).

253   Patterns for coalescent time (tMRCA) estimates were similar to those for within-

254 population divergences for BEAST and RelTime (**Fig. 2b** and **3a**), except that the

255 median Δt were closer to zero for BEAST-Constant, BEAST-Yule, and RelTime (**Fig.**

256 **2b** and **3a**). RelTime performed better in estimating species tMRCAs than within-

257 population divergences because the effect of very short branches was much smaller.

258 All methods performed well for interspecies comparisons, and Δt values were generally

259 close to zero (**Fig. 2c** and **3a**).

260   Overall, the analyses of data simulated under a strict clock and without any ILS

261 produced similar patterns among divergence levels. Interestingly, such baseline

262 situations have not been analyzed previously. So, with this baseline, we can begin to

263 assess the impact of biological complexities on divergence time inference over

264 different time scales. For example, a major biological feature of the data utilized in

265 phylogeography studies is the presence of incomplete lineage sorting (ILS). We

266 examined whether the presence of ILS has any impact on node time estimates. For

267 this comparison, we used time slopes, because the presence of ILS does not allow

268 one to distinguish between within-population, coalescence, and inter-species

269 divergences clearly. The distributions of time slopes, derived from 100 simulated

270 datasets in each case, are shown for RelTime (gray), BEAST-Yule (red), BEAST-BD

271 (blue), BEAST-Constant (green), and BEAST-Skyline (yellow) (**Fig. 5**). Generally,

272 BEAST analyses under distinct tree priors and RelTime performed well, with time

273 slopes showing a sharp peak close to 1 for no-ILS (**Fig. 5a**) and ILS phylogenies (**Fig.**

274 **5d**). Interestingly, the dispersion of time slopes for ILS phylogenies was smaller than

275 that for the no-ILS phylogenies (**Fig. 5a** and **d**), which is likely because no-ILS

276 phylogenies contained a higher amount of short branches near the tips of the

277 phylogeny as compared to the ILS phylogenies (e.g., **Fig. 1**).

278 *Branch Rates Varying Independently (IBR)*

279 The presence of evolutionary rate variability across branches produced results similar

280 to those observed for the strict clock for most of the methods (**Fig. 2** and **4**). The main

281 problem was observed in BEAST-Yule analysis, in which the overestimation of times

282 became more acute (**Fig. 3b**) and time slopes showed clear departure away from 1

283 (red curves in **Fig. 5b** and **e**). These results agree with Ritchie et al. (2017)'s report

284 about the worse performance of BEAST-Yule for ILS phylogenies, but our analyses

9

285   showed that this problem becomes even more acute for no-ILS phylogenies (**Fig. 5b**,

286   red curve). RelTime performed well for both no-ILS and ILS phylogenies simulated

287   under the IBR model. For intra-population nodes, RelTime underestimated times

288   slightly (**Fig. 2d**), but in RelTime time slopes were closer to 1 than those from BEAST

289   analyses, although dispersion around the mean was slightly higher for RelTime (**Fig.**

290   **5b** and **e**).

291   As expected, the variability of evolutionary rates caused a greater dispersion

292   of times slopes for BEAST and RelTime for no-ILS datasets (**Fig. 5a** vs. **5b**) and ILS

293   datasets (**Fig. 5d** vs. **5e**). However, the difference in dispersion for no-ILS and ILS

294   datasets was rather small (**Fig. 5a** vs. **5d** and **Fig. 5b** vs. **5e**), which can be observed

295   through the similar standard deviation (SD) values of the distribution of slopes. Overall,

296   we found that all methods performed similarly for no-ILS and ILS phylogenies. The

297   exception was BEAST-Yule analyses, in which the mean of the distribution of slopes

298   was closer to one with SD values considerably lower in ILS simulations (SD=0.124 for

299   no-ILS datasets and SD=0.074 for ILS datasets) (**Fig. 5b** vs. **5e**).

300   *Branch rate variation with autocorrelation (ABR)*

301   In the analysis of CBR and IBR datasets, we were able to assume correct evolutionary

302   rate prior in the BEAST analyses. However, BEAST did not have a facility to select an

303   autocorrelated clock model, so we examined how the use of IBR model for ABR

304   datasets impacts performance. There was not much difference between the time

305   slopes for IBR and ABR models in no-ILS phylogenies, except for the BEAST-Yule

306   analyses (**Fig. 5b** vs. **5c**). This means that the violation of the clock model in BEAST

307   when analyzing ABR datasets, caused limited performance detriment. However, the

308   dispersion of time slopes was markedly higher for datasets with ILS (**Fig. 5e** vs. **5f**).

309   RelTime produced a distribution of error that was slightly narrower when compared to

310   BEAST under the BD and coalescent tree priors (**Fig. 5f**). Interestingly, BEAST-Yule's

311   performance for ABR datasets with ILS turns out to be better than other BEAST

312   analyses (**Fig. 5f**). This result differs from IBR analysis in our simulation and the

313   conclusion of Ritchie et al. (2017), which is made only based on BEAST-Yule's

314   performance under the IBR model.

315   **Discussion**

316   Bayesian methods are frequently used to estimate sequence divergence times in
317   inferring phylogeography of species and populations. However, there has been a
318   paucity of computer simulation studies assessing their accuracy, apart from Ritchie's
319   et al. (2017) study. We have conducted simulations under a variety of scenarios to
320   explore the accuracy of Bayesian time estimates in mixed sample phylogenies with the
321   presence of ILS, the violation of a strict molecular clock, and the misspecification of
322   branch rate model (**Table 1**). Also, we have compared the performance of Bayesian
323   methods with RelTime, a non-Bayesian approach. Thus, we were able to access the
324   impact of tree priors on divergence time estimation in BEAST and showed that BEAST
325   performs well for this kind of dataset. Besides this, we found that RelTime is a reliable
326   and   computationally   efficient   method   for   estimating   divergence   times   in
327   phylogeographic,   phylodynamics   and   species   delimitation   studies.   In   these
328   evaluations, we chose not to study the impact of calibrations on time estimation as
329   other molecular dating studies have done before (Duchêne, Lanfear, & Ho, 2014;
330   Warnock, Parham, Joyce, Lyson, & Donoghue, 2014; Warnock, Yang, & Donoghue,
331   2017) in order to avoid confounding the impact of calibrations with that of rate models
332   and other priors.

333   First, we found that RelTime performed well in inferring timetrees for datasets
334   with mixed sampling. RelTime's performance was comparable to the performance of
335   BEAST when birth-death, constant coalescent and skyline tree priors were used. We
336   found that RelTime's performance was the least affected by the range of simulated
337   scenarios. It is because RelTime does not require *a priori* specification of tree priors or
338   rate models and directly estimates relative lineage rates and times from branch
339   lengths. In RelTime, the relative rate between sister lineages is the ratio of the
340   evolutionary depths of the two lineages and a relative rate framework is used. This
341   approach contrasts with Bayesian methods, which require the specification of a branch
342   rate model. Therefore, the varying modes of the rate of evolution simulated as well as
343   the presence/absence of ILS did not impact RelTime performance as much as they did
344   impact BEAST under some tree priors. Importantly, RelTime outperformed BEAST in
345   estimating divergences when the branch rates were autocorrelated. As Tao et al.
346   (2019) have shown that the autocorrelation of rates is common in molecular
347   phylogenies, RelTime will perform better than BEAST in empirical data analysis of

11

348 datasets with a mix of inter and intra-species divergences. RelTime also performed
349 well for ILS phylogenies, which is essential because ILS is widespread for datasets
350 with mixed sampling in phylogeographic studies (Jennings & Edwards, 2005; K. Wang
351 et al., 2018).

352   We have also found that the performance of BEAST-Yule varied greatly
353 depending on the presence/absence of rate variation, ILS, and rate autocorrelation.
354 Setting the performance of BEAST-Yule aside, one can conclude that the presence of
355 rate variation introduced $\geq$3x more uncertainty in time estimation (for instance, the SD
356 in slopes' distribution was ~0.019 and ~0.103 for ILS datasets under CBR and ABR,
357 respectively). The difference was the highest for ILS datasets with the ABR model.
358 Interestingly, our simulation results suggest that the accuracy of dating with datasets
359 containing a mix of inter- and intra-species sampling may not be strongly impacted by
360 the actual shape of the phylogenetic trees, as ILS and no-ILS phylogenies resulted in
361 the similar performance of the dating methods. For BEAST-Yule, the performance was
362 significantly improved when ILS data was analyzed, mainly for ABR simulations. This
363 improvement is likely because of a trade-off between the fact that no-ILS phylogenies
364 contain many more closely-related sequences for which divergence dates are harder
365 to estimate. The number of short branches close to the tips of the phylogeny is much
366 higher in no-ILS than in ILS phylogenies (see **Fig. 1**). This pattern leads to a stronger
367 violation of the Yule process than in the case where ILS was allowed to occur, but it
368 does not appear to impact dating with BD and coalescent priors. This difference is
369 likely because the Yule prior has less flexibility when compared to the other ones, as
370 it considers all divergences to be speciation events and does not model extinction).

371   Here, for datasets under CBR evolution, the choice of tree priors had little
372 impact on BEAST results. In CBR phylogenies, the confounding effect of times and
373 rates is less problematic than the case in which rates vary among branches. It is,
374 however, important to note that BEAST analyses used a strict clock prior in such cases,
375 so we expect it to perform better than RelTime (which does not assume any *a priori*
376 rate model).

377   Our analyses confirmed conclusions reached by Ritchie et al. (2017) about the
378 poor performance of BEAST-Yule analyses as compared to BEAST-BD and BEAST-
379 Skyline. Besides this, our analysis of no-ILS phylogenies with IBR model extends this

380 conclusion, as we have found BEAST-Yule to perform even worse for no-ILS datasets
381 as compared to ILS datasets. However, we find that those conclusions do not apply
382 for ABR phylogenies. In the light of the knowledge that the ABR model likely applies
383 for molecular phylogenies and that ILS is likely to be present in mixed sample datasets,
384 BEAST-Yule's better performance is noteworthy. Overall, however, BEAST-Skyline
385 approach appears to perform the best across all rate and lineage sorting models, which
386 is likely due to the greater flexibility of the Skyline approach (Drummond et al., 2005)
387 that does not assume a constant population size and can accommodate increased
388 diversification in recent times.

389   Lastly, although equivalent in performance, RelTime generates time estimates
390 orders of magnitude quicker than Bayesian approaches. RelTime has increasingly
391 been used to estimate divergence times for large trees and phylogenomic data (*e.g.*
392 Bond et al. 2014; Mahony et al. 2017; Irisarri et al. 2018; Shin et al. 2018) and was
393 already reported as a robust method for dating large datasets (Filipski, Murillo,
394 Freydenzon, Tamura, & Kumar, 2014; Mello et al., 2017; Tamura et al., 2012). Here,
395 we extend its applicability to datasets that have a mixture of intra- and inter-species
396 sampling, as frequently found in phylogeographic, phylodynamics and species
397 delimitation studies, without assuming *a priori* any model for the branching process on
398 the tree or the branch rate variation.

399   In conclusion, our simulations covered a wide range of modes of evolution that
400 considered distinct scenarios of rate evolution (CBR, IBR, and ABR) and
401 presence/absence of ILS, many of which was never carried out before for datasets
402 with a mixed sampling of intra- and inter-species sequences. Therefore, our study
403 provides important insights about the impact of tree priors on divergence time
404 estimation in BEAST. Comparisons between BEAST and RelTime's performance on
405 mixed sampling data also showed that RelTime could be used as a reliable and
406 computationally efficient alternative for estimating divergence times in
407 phylogeographic, phylodynamics and species delimitation studies.

408

Acknowledgments

13

415

Author contributions

417 SK and BM conceived the idea, and BM and SK designed research. BM performed

418 analyses; BM, QT, and SK discussed results and wrote the manuscript.

419

Data accessibility statement

421 All the simulated datasets will be made available from FigShare, a publicly accessible

422 repository.

423

References

425 Battistuzzi, F. U., Billing-Ross, P., Murillo, O., Filipski, A., & Kumar, S. (2015). A
426   Protocol for Diagnosing the Effect of Calibration Priors on Posterior Time
427   Estimates: A Case Study for the Cambrian Explosion of Animal Phyla. *Molecular
428   Biology and Evolution*, *32*(7), 1907–1912. doi:10.1093/molbev/msv075
429 Battistuzzi, F. U., Billing-Ross, P., Paliwal, A., & Kumar, S. (2011). Fast and Slow
430   Implementations of Relaxed-Clock Methods Show Similar Patterns of Accuracy
431   in Estimating Divergence Times. *Molecular Biology and Evolution*, *28*(9), 2439–
432   2442. doi:10.1093/molbev/msr100
433 Bond, J. E., Garrison, N. L., Hamilton, C. A., Godwin, R. L., Hedin, M., & Agnarsson,
434   I. (2014). Phylogenomics Resolves a Spider Backbone Phylogeny and Rejects
435   a Prevailing Paradigm for Orb Web Evolution. *Current Biology*, *24*(15), 1765–
436   1771. doi:10.1016/j.cub.2014.06.034
437 Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., … Drummond,
438   A. J. (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis.
439   *PLOS Computational Biology*, *10*(4), e1003537.
440   doi:10.1371/journal.pcbi.1003537
441 Chen, W.-C. (2011). Overlapping Codon Model, Phylogenetic Clustering, and
442   Alternative Partial Expectation Conditional Maximization Algorithm. *Ph.D. Diss.,
443   Iowa Stat University*.
444 Degnan, J. H., & Salter, L. A. (2005). Gene tree distributions under the coalescent
445   process. *Evolution; International Journal of Organic Evolution*, *59*(1), 24–37.
446 dos Reis, M., Gunnell, G. F., Barba-Montoya, J., Wilkins, A., Yang, Z., & Yoder, A. D.
447   (2018). Using Phylogenomic Data to Explore the Effects of Relaxed Clocks and
448   Calibration Strategies on Divergence Time Estimation: Primates as a Test
449   Case. *Systematic Biology*, *67*(4), 594–615. doi:10.1093/sysbio/syy001
450 dos Reis, M., & Yang, Z. (2011). Approximate likelihood calculation on a phylogeny for
451   Bayesian estimation of divergence times. *Molecular Biology and Evolution*,
452   *28*(7), 2161–2172. doi:10.1093/molbev/msr045

dos Reis, M., Thawornwattana, Y., Angelis, K., Telford, M. J., Donoghue, P. C. J., & Yang, Z. (2015). Uncertainty in the Timing of Origin of Animals and the Limits of Precision in Molecular Timescales. *Current Biology*, *25*(22), 2939–2950. doi:10.1016/j.cub.2015.09.066

Drummond, A. J., Ho, S. Y. W., Phillips, M. J., & Rambaut, A. (2006). Relaxed Phylogenetics and Dating with Confidence. *PLoS Biology*, *4*(5), e88. doi:10.1371/journal.pbio.0040088

Drummond, A. J., Rambaut, A., Shapiro, B., & Pybus, O. G. (2005). Bayesian Coalescent Inference of Past Population Dynamics from Molecular Sequences. *Molecular Biology and Evolution*, *22*(5), 1185–1192. doi:10.1093/molbev/msi103

Duchêne, D. A., Duchêne, S., Holmes, E. C., & Ho, S. Y. W. (2015). Evaluating the Adequacy of Molecular Clock Models Using Posterior Predictive Simulations. *Molecular Biology and Evolution*, *32*(11), 2986–2995. doi:10.1093/molbev/msv154

Duchêne, S., Lanfear, R., & Ho, S. Y. W. (2014). The impact of calibration and clock-model choice on molecular estimates of divergence times. *Molecular Phylogenetics and Evolution*, *78*, 277–289. doi:10.1016/j.ympev.2014.05.032

Edwards, S. V., & Beerli, P. (2000). Gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution; International Journal of Organic Evolution*, *54*(6), 1839–1854.

Edwards, S. V., Shultz, A. J., & Campbell-Staton, S. C. (2015). Next-generation sequencing and the expanding domain of phylogeography. *Folia Zoologica*, *64*(3), 187–206. doi:10.25225/fozo.v64.i3.a2.2015

Endicott, P., & Ho, S. Y. W. (2008). A Bayesian evaluation of human mitochondrial substitution rates. *American Journal of Human Genetics*, *82*(4), 895–902. doi:10.1016/j.ajhg.2008.01.019

Esselstyn, J. A., Evans, B. J., Sedlock, J. L., Anwarali Khan, F. A., & Heaney, L. R. (2012). Single-locus species delimitation: a test of the mixed Yule-coalescent model, with an empirical application to Philippine round-leaf bats. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1743), 3678–3686. doi:10.1098/rspb.2012.0705

Feller, W. (1939). Die Grundlagen der Volteraschen Theorie Des Kampfes Ums Dasein in Wahrscheinlichkeitstheoretischer Behandlung. *Acta Biotheoretica*, *5*(1), 11–40.

Filipski, A., Murillo, O., Freydenzon, A., Tamura, K., & Kumar, S. (2014). Prospects for Building Large Timetrees Using Molecular Data with Incomplete Gene Coverage among Species. *Molecular Biology and Evolution*, *31*(9), 2542–2550. doi:10.1093/molbev/msu200

Hasegawa, M., Kishino, H., & Yano, T. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, *22*(2), 160–174.

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

Heled, J., & Drummond, A. J. (2010). Bayesian Inference of Species Trees from Multilocus Data. *Molecular Biology and Evolution*, *27*(3), 570–580. doi:10.1093/molbev/msp274

15

Ho, S. Y. W. (2014). The changing face of the molecular evolutionary clock. *Trends in Ecology and Evolution*, *29*(9), 496–503. doi:10.1016/j.tree.2014.07.004

Ho, S. Y. W., Duchêne, S., & Duchêne, D. A. (2015). Simulating and detecting autocorrelation of molecular evolutionary rates among lineages. *Molecular Ecology Resources*, *15*(4), 688–696. doi:10.1111/1755-0998.12320

Hudson, R. R. (2002). Generating samples under a Wright-Fisher neutral model of genetic variation. *Bioinformatics*, *18*(2), 337–338. doi:10.1093/bioinformatics/18.2.337

Irisarri, I., Singh, P., Koblmüller, S., Torres-Dowdall, J., Henning, F., Franchini, P., … Meyer, A. (2018). Phylogenomics uncovers early hybridization and adaptive loci shaping the radiation of Lake Tanganyika cichlid fishes. *Nature Communications*, *9*(1), 3159. doi:10.1038/s41467-018-05479-9

Jennings, W. B., & Edwards, S. V. (2005). Speciational history of Australian grass finches (Poephila) inferred from thirty gene trees. *Evolution; International Journal of Organic Evolution*, *59*(9), 2033–2047.

Kendall, D. G. (1948). On the Generalized "Birth-and-Death" Process. *The Annals of Mathematical Statistics*, *19*(1), 1–15.

Kishino, H., Thorne, J. L., & Bruno, W. J. (2001). Performance of a divergence time estimation method under a probabilistic model of rate evolution. *Mol Biol Evol*, *18*(3), 352–361. doi:10.1093/oxfordjournals.molbev.a003811

Kuhner, M. K., Yamato, J., & Felsenstein, J. (1995). Estimating effective population size and mutation rate from sequence data using Metropolis-Hastings sampling. *Genetics*, *140*(4), 1421–1430.

Kuhner, M. K., Yamato, J., & Felsenstein, J. (1998). Maximum Likelihood Estimation of Population Growth Rates Based on the Coalescent. *Genetics*, *149*(1), 429–434.

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution*, *35*(6), 1547–1549. doi:10.1093/molbev/msy096

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*, *33*(7), 1870–1874. doi:10.1093/molbev/msw054

Lemmon, A. R., & Lemmon, E. M. (2012). High-Throughput Identification of Informative Nuclear Loci for Shallow-Scale Phylogenetics and Phylogeography. *Systematic Biology*, *61*(5), 745–761. doi:10.1093/sysbio/sys051

Liu, L., & Yu, L. (2010). Phybase: An R package for species tree analysis. *Bioinformatics*, *26*(7), 962–963. doi:10.1093/bioinformatics/btq062

Maddison, W. P. (1997). Gene Trees in Species Trees. *Systematic Biology*, *46*(3), 523–536. doi:10.1093/sysbio/46.3.523

Mahony, S., Foley, N. M., Biju, S. D., & Teeling, E. C. (2017). Evolutionary History of the Asian Horned Frogs (Megophryinae): Integrative Approaches to Timetree Dating in the Absence of a Fossil Record. *Molecular Biology and Evolution*, *34*(3), 744–771. doi:10.1093/molbev/msw267

Manzanilla, V., Kool, A., Nguyen Nhat, L., Nong Van, H., Le Thi Thu, H., & de Boer, H. J. (2018). Phylogenomics and barcoding of Panax: toward the identification of ginseng species. *BMC Evolutionary Biology*, *18*, 44. doi:10.1186/s12862-018-1160-y

548   Marin, J., & Hedges, S. B. (2018). Undersampling Genomes has Biased Time and
549   Rate Estimates Throughout the Tree of Life. *Molecular Biology and Evolution*,
550   *35*(8), 2077–2084. doi:10.1093/molbev/msy103
551   McCluskey, B. M., & Postlethwait, J. H. (2015). Phylogeny of Zebrafish, a "Model
552   Species," within Danio, a "Model Genus." *Molecular Biology and Evolution*,
553   *32*(3), 635–652. doi:10.1093/molbev/msu325
554   McCormack, J. E., Hird, S. M., Zellmer, A. J., Carstens, B. C., & Brumfield, R. T. (2013).
555   Applications of next-generation sequencing to phylogeography and
556   phylogenetics. *Molecular Phylogenetics and Evolution*, *66*(2), 526–538.
557   doi:10.1016/j.ympev.2011.12.007
558   Mello, B., Tao, Q., Tamura, K., & Kumar, S. (2017). Fast and Accurate Estimates of
559   Divergence Times from Big Data. *Molecular Biology and Evolution*, *34*(1), 45–
560   50. doi:10.1093/molbev/msw247
561   Mello, B., Vilela, J. F., & Schrago, C. G. (2018). Conservation phylogenetics and
562   computational species delimitation of Neotropical primates. *Biological
563   Conservation*, *217*, 397–406. doi:10.1016/j.biocon.2017.11.017
564   Melville, J., Haines, M. L., Boysen, K., Hodkinson, L., Kilian, A., Smith Date, K. L., …
565   Parris, K. M. (2017). Identifying hybridization and admixture using SNPs:
566   application of the DArTseq platform in phylogeographic research on
567   vertebrates. *Royal Society Open Science*, *4*(7), 161061.
568   doi:10.1098/rsos.161061
569   Merwe, M., McPherson, H., Siow, J., & Rossetto, M. (2014). Next-Gen phylogeography
570   of rainforest trees: exploring landscape-level cpDNA variation from whole-
571   genome sequencing. *Molecular Ecology Resources*, *14*(1), 199–208.
572   doi:10.1111/1755-0998.12176
573   Ogilvie, H. A., Bouckaert, R. R., & Drummond, A. J. (2017). StarBEAST2 Brings Faster
574   Species Tree Inference and Accurate Estimates of Substitution Rates.
575   *Molecular Biology and Evolution*, *34*(8), 2101–2114.
576   doi:10.1093/molbev/msx126
577   Ogilvie, H. A., Heled, J., Xie, D., & Drummond, A. J. (2016). Computational
578   Performance and Statistical Accuracy of *BEAST and Comparisons with Other
579   Methods. *Systematic Biology*, *65*(3), 381–396. doi:10.1093/sysbio/syv118
580   O'Reilly, J. E., dos Reis, M., & Donoghue, P. C. J. (2015). Dating Tips for Divergence-
581   Time Estimation. *Trends in Genetics*, *31*(11), 637–650.
582   doi:10.1016/j.tig.2015.08.001
583   Plummer, M., Best, N., Cowles, K., & Vines, K. (2006). CODA: convergence diagnosis
584   and output analysis for MCMC. *R News*, *6*(1), 7–11.
585   Rambaut, A., & Grass, N. C. (1997). Seq-Gen: an application for the Monte Carlo
586   simulation of DNA sequence evolution along phylogenetic trees. *Comput Appl
587   Biosci*, *13*(3), 235–238. doi:10.1093/bioinformatics/13.3.235
588   Rannala, B., & Yang, Z. (2003). Bayes estimation of species divergence times and
589   ancestral population sizes using DNA sequences from multiple loci. *Genetics*,
590   *164*(4), 1645–1656.
591   Rannala, B., & Yang, Z. (2017). Efficient Bayesian Species Tree Inference under the
592   Multispecies Coalescent. *Systematic Biology*, *66*(5), 823–842.
593   doi:10.1093/sysbio/syw119

Rosenberg, M. S., & Kumar, S. (2003). Heterogeneity of nucleotide frequencies among evolutionary lineages and phylogenetic inference. *Molecular Biology and Evolution*, *20*(4), 610–621. doi:10.1093/molbev/msg067

Shin, S., Clarke, D. J., Lemmon, A. R., Moriarty Lemmon, E., Aitken, A. L., Haddad, S., … McKenna, D. D. (2018). Phylogenomic Data Yield New and Robust Insights into the Phylogeny and Evolution of Weevils. *Molecular Biology and Evolution*, *35*(4), 823–836. doi:10.1093/molbev/msx324

Stadler, T. (2011). Simulating trees with a fixed number of extant species. *Systematic Biology*, *60*(5), 676–684. doi:10.1093/sysbio/syr029

Takahata, N. (1989). Gene genealogy in three related populations: consistency probability between gene and population trees. *Genetics*, *122*(4), 957–966.

Takahata, N., & Nei, M. (1985). Gene genealogy and variance of interpopulational nucleotide differences. *Genetics*, *110*(2), 325–344.

Tamura, K., Battistuzzi, F. U., Billing-Ross, P., Murillo, O., Filipski, A., & Kumar, S. (2012). Estimating divergence times in large molecular phylogenies. *Proceedings of the National Academy of Sciences*, *109*(47), 19333–19338. doi:10.1073/pnas.1213199109

Tamura, K., Tao, Q., & Kumar, S. (2018). Theoretical Foundation of the RelTime Method for Estimating Divergence Times from Variable Evolutionary Rates. *Molecular Biology and Evolution*, *35*(7), 1770–1782. doi:10.1093/molbev/msy044

Tao, Q., Tamura, K., Battistuzzi, F., & Kumar, S. (2019). A new method for detecting autocorrelation of evolutionary rates in large phylogenies. *Molecular Biology and Evolution*. doi:10.1093/molbev/msz014

Wang, C., Shikano, T., Persat, H., & Merilä, J. (2015). Mitochondrial phylogeography and cryptic divergence in the stickleback genus Pungitius. *Journal of Biogeography*, *42*(12), 2334–2348. doi:10.1111/jbi.12591

Wang, K., Lenstra, J. A., Liu, L., Hu, Q., Ma, T., Qiu, Q., & Liu, J. (2018). Incomplete lineage sorting rather than hybridization explains the inconsistent phylogeny of the wisent. *Communications Biology*, *1*, 169. doi:10.1038/s42003-018-0176-6

Warnock, R. C. M., Parham, J. F., Joyce, W. G., Lyson, T. R., & Donoghue, P. C. J. (2014). Calibration uncertainty in molecular dating analyses: there is no substitute for the prior evaluation of time priors. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1798), 20141013–20141013. doi:10.1098/rspb.2014.1013

Warnock, R. C. M., Yang, Z., & Donoghue, P. C. J. (2017). Testing the molecular clock using mechanistic models of fossil preservation and molecular evolution. *Proc. R. Soc. B*, *284*(1857), 20170227. doi:10.1098/rspb.2017.0227

Xu, B., & Yang, Z. (2016). Challenges in Species Tree Estimation Under the Multispecies Coalescent Model. *Genetics*, *204*(4), 1353–1368. doi:10.1534/genetics.116.190173

Yang, Z. (2015). The BPP program for species tree estimation and species delimitation. *Current Zoology*, *61*(5), 854–865. doi:10.1093/czoolo/61.5.854

Yang, Z., & Rannala, B. (2014). Unguided Species Delimitation Using DNA Sequence Data from Multiple Loci. *Molecular Biology and Evolution*, *31*(12), 3125–3135. doi:10.1093/molbev/msu279

640 Yule, G. U. (1924). A mathematical theory of evolution, based on the conclusions of
641   Dr. J. C. Willis. *Phil. Trans. R. Soc. Lond. B*, *213*, 21–87.
642   doi:10.1098/rstb.1925.0002
643

644

645

646 **Table 1**. Computer simulation scenarios considered in this article.

Branch rates	BEAST	BEAST	RelTime	RelTime
	ILS	No ILS	ILS	No ILS
Constant (CBR)	New	New	New	New
Independent (IBR)	Ritchie et al.	New	New	New
Autocorrelated (ABR)	New	New	New	new

647

648 ILS, the presence of incomplete lineage sorting. No-ILS, the absence of incomplete

649 lineage sorting.

650

651
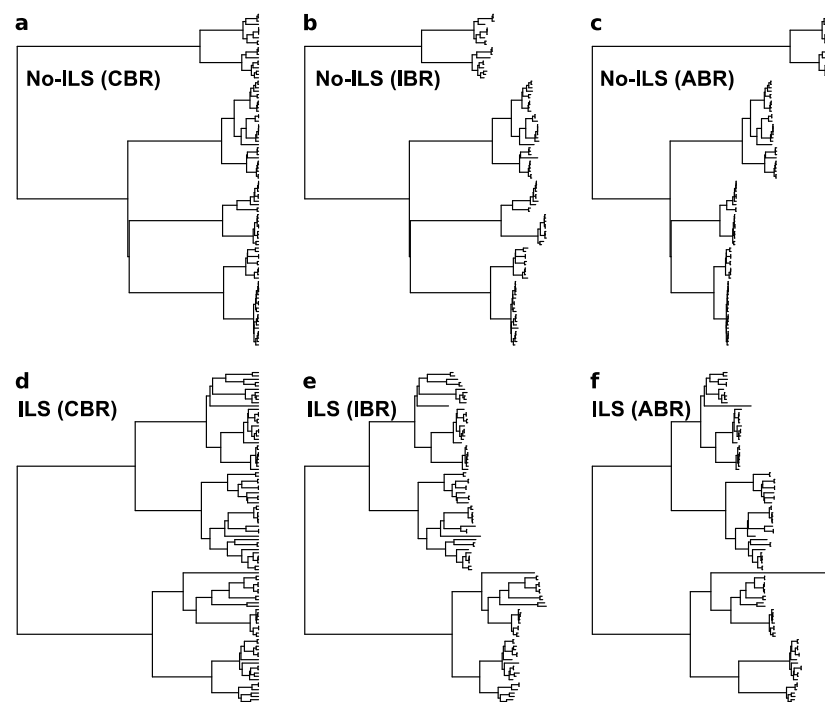
652   **Figure 1**. An example of the diversity of trees used in this study. Panels **a-c**, trees

653   simulated without lineage sorting (no-ILS); **d-e**, trees simulated with incomplete

654   lineage sorting (ILS). **a** and **d** were simulated under a constant branch rate (CBR)

655   model; **b** and **e** were simulated under an independent branch rate (IBR) model; and **c**

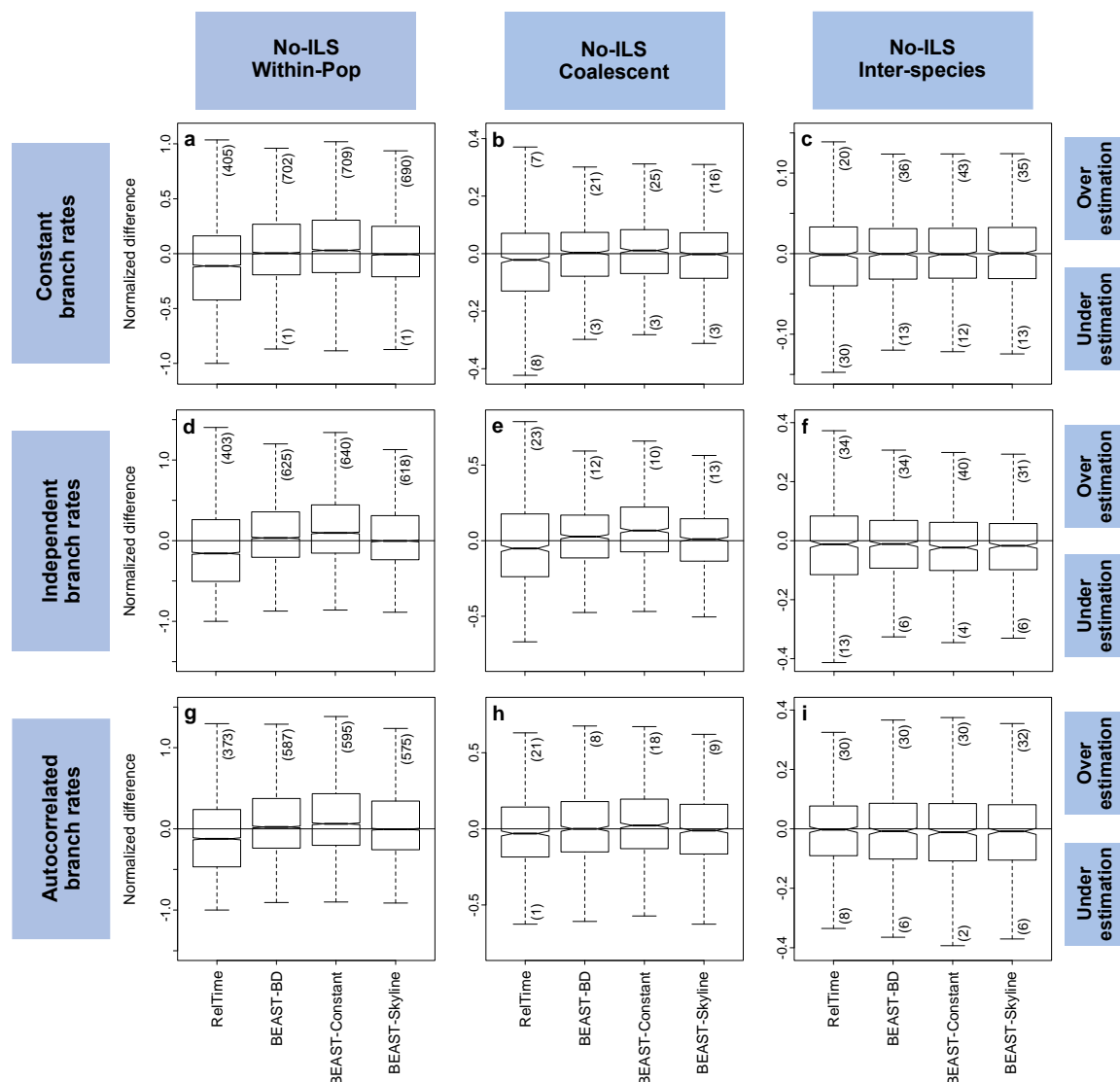656   and **f** were simulated under an autocorrelated branch rate (ABR) model.

21

Figure 2. Boxplots of the differences between true and estimated times normalized by the true times (simulations under no-ILS). Constant branch rates in the top panels (**a-c**), independent branch rates on the middle panels (**d-f**), and correlated branch rates on the bottom panels (**g-i**). Left-hand panels **a**, **d** and **g** display normalized difference values for population intra-divergences; center panels **b**, **e** and **h** for species coalescent times; and right-hand panels **c**, **f** and **i** for inter-species divergences. The number inside parentheses indicate the number of outliers, which are not displayed in the figure.
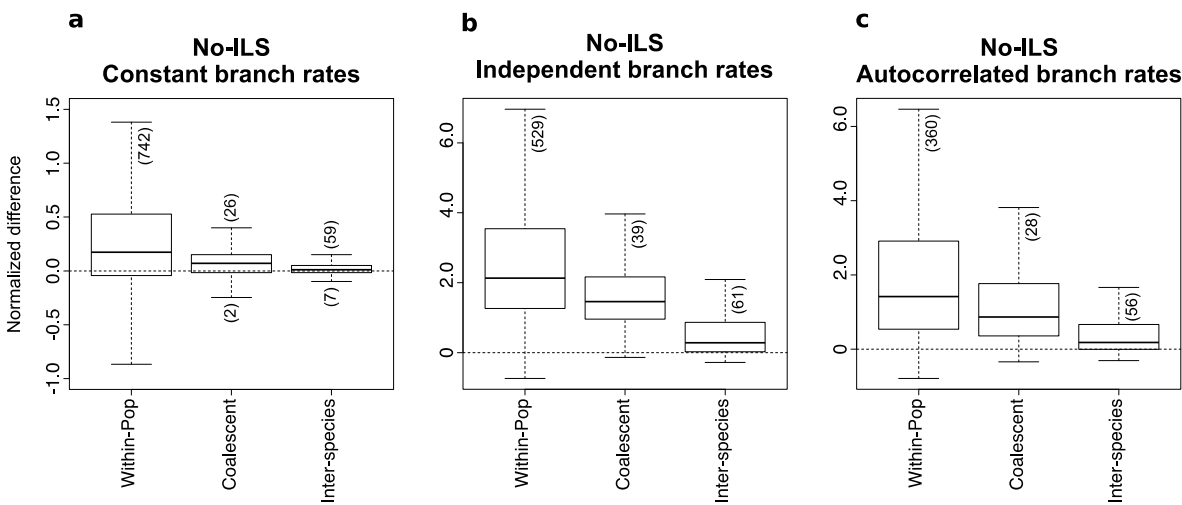
666

667

668

669 **Figure 3**. Performance of the Yule prior to estimating divergence times in BEAST
670 under the absence of lineage sorting (simulations under no-ILS). **a-c**, normalized
671 differences between actual and estimated times for intra-population divergences,
672 species tMRCAs, and inter-species divergences. The number inside parentheses
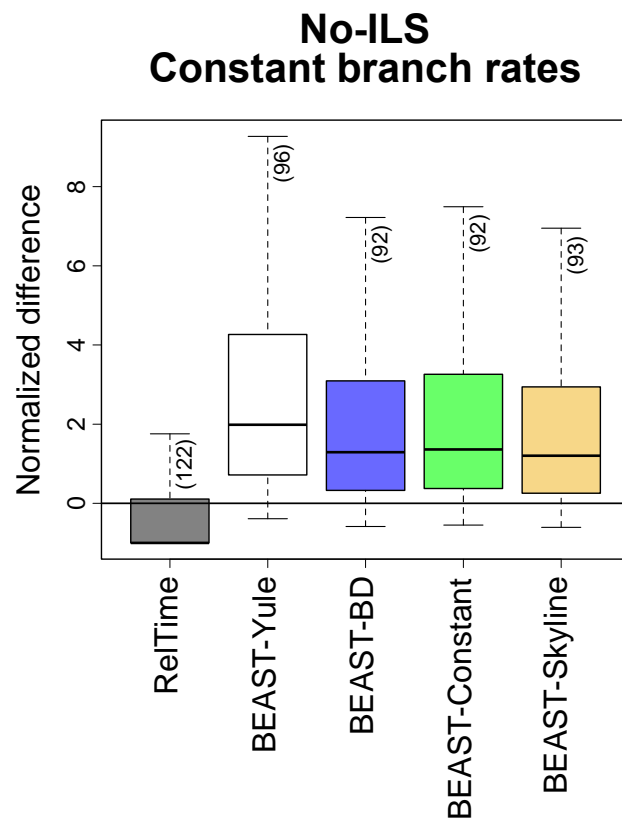673 indicate the number of outliers, which are not displayed in the figure.

Figure 4: Boxplots of the differences between actual and estimated times normalized by the true times for nodes < 0.1% of the root (ingroup) height in simulations under CBR and no-ILS (100 replicates). The number inside parentheses indicate the number of outliers, which are not displayed in the figure.
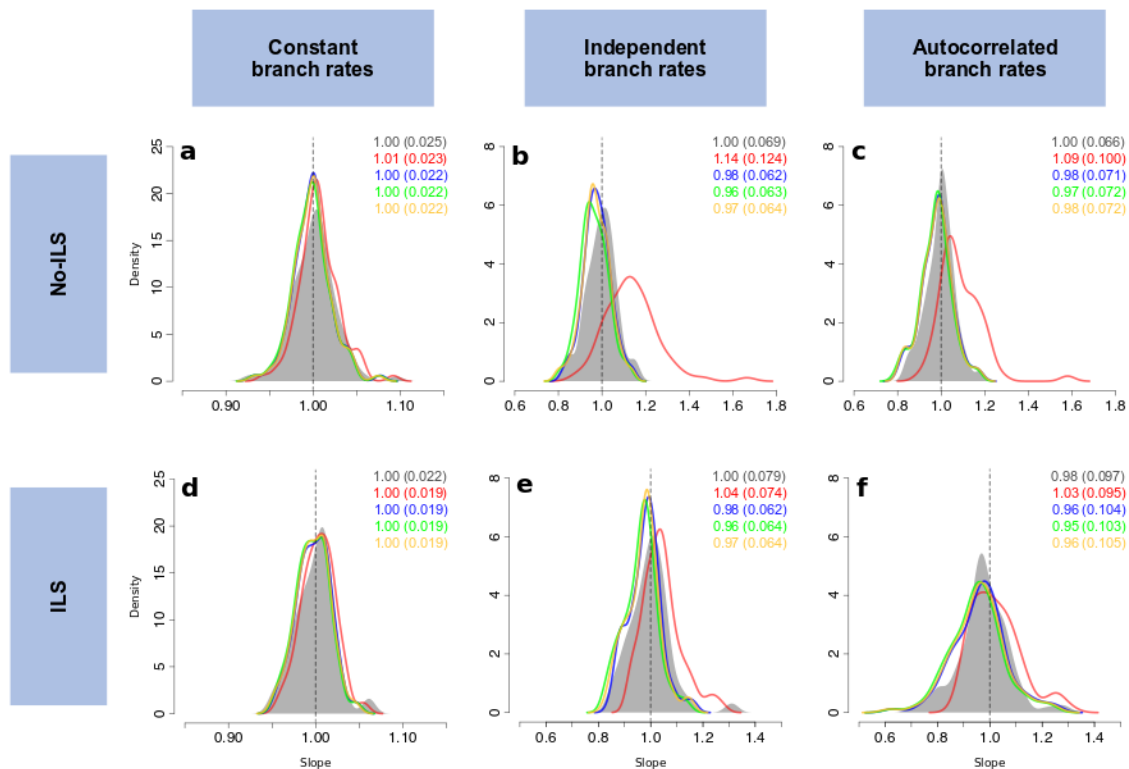
Figure 5. Histogram of slope values resulting from the regression lines between true times and inferred times at RelTime (gray) and BEAST under distinct tree priors: Yule (red), birth-death (blue), constant-size coalescent (green) and skyline coalescent (yellow) (for 100 simulations). Mean and standard deviation (in parentheses) values are displayed inside each panel colored according to the method. Top panels (**a-c**) display results under the absence of lineage sorting (no-ILS), while bottom panels (**d-f**) under the presence of incomplete lineage sorting (ILS). Constant branch rates on the left (**a** and **d**), independent branch rates on the center (**b** and **e**) and correlated branch rates on the right (**c** and **f**).

25