1

2   PASCAL TITLE (Orcid ID : 0000-0002-6316-0736)

3

4

6

7

9   Tip rates, phylogenies, and diversification: what are we estimating, and how good are the

10   estimates?

11

12   Pascal O. Title[1,2,3], Daniel L. Rabosky[1]

13

14   [1]Department of Ecology and Evolutionary Biology and Museum of Zoology, University of

15   Michigan, Ann Arbor, MI 48109, USA

16   [2]Environmental Resilience Institute, Indiana University, Bloomington, IN 47408, USA

17   [3]ptitle@iu.edu

18

19   <u>Abstract</u>

20   1. Species-specific diversification rates, or "tip rates", can be computed quickly from

21   phylogenies and are widely used to study diversification rate variation in relation to

22   geography, ecology, and phenotypes. These tip rates provide a number of theoretical and

23   practical advantages, such as the relaxation of assumptions of rate homogeneity in trait-

24   dependent diversification studies. However, there is substantial confusion in the literature

25   regarding whether these metrics estimate speciation or net diversification rates.

26   Additionally, no study has yet compared the relative performance and accuracy of tip rate

27   metrics across simulated diversification scenarios.

28   2. We compared the statistical performance of three model-free rate metrics (inverse

29   terminal branch lengths; node density metric; DR statistic) and a model-based approach

30   (BAMM). We applied each method to a large set of simulated phylogenies that had been

31   generated under different diversification processes. We summarized performance in

32   relation to the type of rate variation, the magnitude of rate heterogeneity and rate regime

33   size. We also compared the ability of the metrics to estimate both speciation and net

34   diversification rates.

35   3. We show decisively that model-free tip rate metrics provide a better estimate of the rate

36   of speciation than of net diversification. Error in net diversification rate estimates

37   increases as a function of the relative extinction rate. In contrast, error in speciation rate

38   estimates is low and relatively insensitive to extinction. Overall, and in particular when

39   relative extinction was high, BAMM inferred the most accurate tip rates and exhibited

40   lower error than non-model-based approaches. DR was highly correlated with true

41   speciation rates but exhibited high error variance, and was the best metric for very small

42   rate regimes.

43   4. We found that, of the metrics tested, DR and BAMM are the most useful metrics for

44   studying speciation rate dynamics and trait-dependent diversification. Although BAMM

45   was more accurate than DR overall, the two approaches have complementary strengths.

46   Because tip rate metrics are more reliable estimators of speciation rate, we recommend

47   that empirical studies using these metrics exercise caution when drawing biological

48   interpretations in any situation where the distinction between speciation and net

49   diversification is important.

50

51 Keywords: tip rates, diversification, terminal branch length, node density, DR, BAMM, trait-

52 dependent diversification

53

54

55 <u>Introduction</u>

56   Rates of speciation and extinction vary through time and among lineages (Nee, Mooers &

57 Harvey 1992; Sanderson & Donoghue 1996; Etienne & Haegeman 2012; Jetz et al. 2012; Moen

58 & Morlon 2014; Alfaro et al. 2018), contributing to dramatic heterogeneity in species richness

59 across the tree of life (Alfaro et al. 2009; Jetz et al. 2012; Barker et al. 2013). By characterizing

60 variation in rates of speciation and extinction, we can better understand the dynamics of

61 biological diversity through time, across geographic and environmental gradients (Zink, Klicka

62 & Barber 2004; Ricklefs 2006; Mittelbach et al. 2007; Silvestro, Schnitzler & Zizka 2011;

63 Rabosky, Title & Huang 2015), and in relation to traits and key innovations (FitzJohn, Maddison

64 & Otto 2009; Near et al. 2012; Beaulieu & O'Meara 2016). Consequently, there has been great

65 interest in statistical methods for inferring rates of speciation and extinction from molecular

66 phylogenies.

67   Although rates of diversification have traditionally been quantified for clades, there has

68 been a growing interest in estimating species-specific rates of diversification, which we refer to

69 here as "tip rates". Tip rates are increasingly used to describe patterns of geographic and trait-

70 associated variation in diversification (Freckleton, Phillimore & Pagel 2008; Jetz et al. 2012;

71 Kennedy et al. 2016; Harvey & Rabosky 2017; Quintero & Jetz 2018; Rabosky et al. 2018). It

72 may seem strange to view evolutionary rates as a property of individual lineages, but such rates

73 emerge naturally from the birth-death model we typically use to conceptualize the diversification

74 process (Nee, Mooers & Harvey 1992; Nee, May & Harvey 1994). Under the birth-death

75 process, individuals (species) are characterized by per-lineage rates of species origination

76 (speciation, λ) and extinction (μ). For the purposes of inference, these rates are typically

77 assumed to be constant among contemporaneous members of a focal clade. However, tip rates

78 can be viewed as our best estimate of the present-day rate of speciation or extinction for an

79 individual lineage, conditional on past (usually recent) evolutionary history. As such, they

80 provide information about the expected amount of time that will elapse before a lineage splits or

81 becomes extinct.

82   A number of approaches have been used to estimate tip rates, including both model-based

83 and non-model-based approaches (i.e., models that are parameterized with speciation and

84 extinction rates, vs metrics that simply rely on branch lengths and splitting events). These

85 approaches vary in terms of how much information they derive from a focal species (i.e., a

86 terminal branch) relative to the amount of information they incorporate from other regions of the

87 phylogeny. On one end of the spectrum, tree-wide estimates (i.e., one rate for the entire

88 phylogeny) of speciation and extinction rates under a constant-rate birth-death (CRBD) model

89 provide tip rates that are maximally auto-correlated (identical) across species in the clade; such

90 rates for any given species are not independent of rates for any other species in the group of

91 interest. On the other end of the spectrum, terminal branch lengths can be used to derive a

92   censored estimate of the rate of speciation that is minimally autocorrelated with rates for other

93   species in the focal clade. Terminal branch lengths are largely unique to each species (rates

94   might be identical only for sister taxa), but provide a noisy measure of speciation, due to the

95   stochasticity inherent in the diversification process (Nee, May & Harvey 1994), and they have

96   been employed as a summary statistic in assessing model adequacy (Bromham, Hua & Cardillo

97   2016; Gomes, Sorenson & Cardoso 2016). In contrast to single (terminal) branch estimates, tree-

98   wide estimates should be less susceptible to stochastic noise, because they incorporate

99   information from the entirety of the tree (e.g., multiple branches are used in the estimates). Of

100 course, the tree-wide estimate necessarily assumes that all tips share a common underlying

101 diversification process. Other tip rate metrics fall somewhere between these two extremes,

102 incorporating some tree-wide information but relaxing the assumption of homogeneous rates

103 across all lineages (node density metric: Freckleton, Phillimore & Pagel 2008; DR: Jetz et al.

104 2012). The estimation of tip-specific rates thus entails a tradeoff between the precision of

105 individual estimates and the stochastic error associated with those estimates.

106   BAMM (Bayesian Analysis of Macroevolutionary Mixtures, Rabosky 2014) is a model-

107 based approach that can accommodate heterogeneity in the rate of diversification through time

108 and among lineages. BAMM simulates a posterior distribution of macroevolutionary rate shift

109 configurations given a phylogeny of interest; marginal rates of speciation and extinction for

110 individual taxa can then be extracted from this distribution. In this framework, the correlation in

111 rates between any pair of species is a function of the posterior probability that they share a

112 common macroevolutionary rate regime (Rabosky et al. 2014). If the tree-wide posterior

113 probability of rate variation is low, the marginal rates estimates for individual species will be

114 similar across the entire tree, as under a CRBD model. Likewise, any pair of taxa that are

115 consistently assigned to the same macroevolutionary rate regime will necessarily have identical

116 tip rates.

117   Tip rates are best suited to a host of questions and hypotheses where the diversification

118 dynamics over the evolutionary history of a group are either less relevant, or no more relevant,

119 than the rates of diversification closer to the present day. For example, many hypotheses

120 involving trait-dependent diversification implicitly assume a time-homogeneous, or constant

121 through time, effect of the trait on diversification rate (Coyne & Orr 2004; Kay et al. 2006;

122 Jablonski 2008; FitzJohn 2010; Claramunt et al. 2011). Harvey & Rabosky (2017) found that the

123   use of tip rates for assessing correlations between continuous traits and diversification has good

124   performance across a range of diversification scenarios. Furthermore, hypotheses pertaining to

125   non-historical geographic patterns of diversity are also better addressed with recent rates of

126   diversification. For example, many hypotheses for the latitudinal diversity gradient propose time-

127   homogeneous effects of particular environmental factors (temperature, energy, geographic area)

128   on rates of diversification (Mittelbach et al. 2007; Kennedy et al. 2014; Rabosky, Title & Huang

129   2015; Schluter 2016; Rabosky et al. 2018). Put simply, if such time-homogeneous processes

130   have shaped the latitudinal diversity gradient (e.g., correlation between speciation and

131   temperature: Rohde 1992), then the effect should be manifest in the distribution of present-day

132   evolutionary rates.

133   At present, there is substantial confusion in the literature over what quantity various tip

134   rate metrics actually measure. The DR statistic (Jetz et al. 2012) was originally described as a

135   measure of the "species-level lineage diversification rate". While supplemental analyses and

136   subsequent work suggested that DR was a better measure of speciation rate than net

137   diversification (Jetz et al. 2012, Belmaker & Jetz 2015, Quintero & Jetz 2018), many studies

138   have nonetheless continued to describe DR as an estimate of the lineage-level net diversification

139   rate (Marin & Hedges 2016; Oliveira et al. 2016; Cai et al. 2017; and many others).The node

140   density metric of Freckleton, Phillimore & Pagel (2008) has also been described as a measure of

141   net diversification. Whether these metrics more accurately measure speciation or net

142   diversification is critically important for interpreting biodiversity patterns (e.g., two regions

143   might differ dramatically in speciation rate, but net diversification rates in each might

144   nonetheless be zero). An objective of our study is thus to compare the ability of DR, node

145   density, and other metrics to estimate speciation and net diversification rates.

146   Despite the potential utility of tip rates in geographic and trait-based analyses of

147   speciation rate heterogeneity (Jetz et al. 2012; Belmaker & Jetz 2015; Oliveira et al. 2016;

148   Quintero & Jetz 2018), there has yet been no comprehensive comparative assessment of the

149   accuracy and precision of the estimates, save for supplemental analyses in Jetz et al. (2012) and

150   Quintero & Jetz (2018). BAMM has low power to infer small rate regimes (Rabosky, Mitchell &

151   Chang 2017; Meyer & Wiens 2017), leading to the possibility that other approaches might

152   perform better for smaller phylogenies or when the variation in rates among clades is subtle.

153   However, DR and related methods will always identify variation in tip rates, even when none

154   exists, provided there is stochastic variation in branch lengths. A goal of this study is therefore to

155   evaluate the trade-off between the stochastic noise inherent in non-model-based approaches, and

156   the conservative but less noisy estimates from model-based metrics. We compare the

157   performance of these metrics across a range of simulation scenarios, which include both discrete

158   and continuous variation in rates.

159

160   **<u>Methods</u>**

161   <u>Tip rate metrics</u>

162   We assessed the accuracy of four tip rate metrics in this study at quantifying rates of

163   speciation. As we demonstrate below (see also Supplementary Figure 5 in Jetz et al. 2012;

164   extended Figure 5 in Quintero & Jetz 2018; Belmaker & Jetz 2015), these metrics are estimators

165   of speciation rate and not net diversification rate, and we refer to them as such throughout. The

166   first metric is the inverse of the equal splits measure (Redding and Mooers 2006), also called the

167   *DR statistic* (Jetz et al. 2012), *DivRate* (Belmaker & Jetz 2015; Oliveira et al. 2016), *ES* (Harvey

168   & Rabosky 2017) or *tip DR* (Quintero & Jetz 2018), which we denote in this study as λ_{DR}. This

169   species-specific measure incorporates the number of splitting events and the internode distances

170   along the root-to-tip path of a phylogeny, while giving greater weight to branches closer to the

171   present (Redding & Mooers 2006; Jetz et al. 2012). λ_{DR} is computed as:

$$\lambda_{DR_i} = \sum_{j=1}^{N_i} b_j \frac{1}{2^{j-1}}$$

172   where λ_{DR_i} is the tip rate for species *i*, N_i is the number of branches between species *i* and the

173   root, b_j is the length of branch *j*, starting at the terminal branch ($j = 1$) and ending with the root.

174   Jetz et al. (2012) demonstrated that, for trees deriving from a Yule process, and with mild

175   extinction, the mean λ_{DR} across tips converges on the true speciation rate.

176   We also considered a simpler metric, node density (Freckleton, Phillimore & Pagel 2008;

177   denoted by λ_{ND}). This is simply the number of splitting events along the path between the root

178   and tip of a phylogeny, divided by the age of the phylogeny. While λ_{DR} down-weights the

179   contribution of branch lengths that are closer to the root, λ_{ND} equally weights the contributions

180   of all branches along a particular root-to-tip path, regardless of where they occur in time. Under a

181   pure-birth model ($\mu = 0$), both λ_{DR} and λ_{ND} should yield unbiased estimates of the rate of

182   speciation.

183   The third measure we considered is the inverse of the terminal branch lengths (λ_{TB}).

184   Rapid speciation rates near the present should be associated with proportionately shorter terminal

185   branches; smaller values of λ_{TB} should thus characterize species with faster rates of speciation.

186   This measure has recently been used as a summary statistic to assess model adequacy in trait-

187   dependent diversification studies (Bromham, Hua & Cardillo 2016; Gomes, Sorenson & Cardoso

188   2016; Harvey & Rabosky 2017). Following Steel & Mooers (2010), we note that the terminal

189   branch lengths can be used to derive an estimate of the speciation rate; this follows from the fact

190   that interior and terminal branches have the same expected value under the Yule process (Steel &

191   Mooers 2010). The corresponding estimator for the *i*'th tip, λ_{TB} is approximately $1 / 2b$ where *b*

192   is the length of a given terminal branch (Steel & Mooers 2010). To our knowledge, λ_{TB} has not

193   been used to explicitly estimate tip rates as we do here, but given its utility as a summary statistic

194   and general theoretical properties (Steel & Mooers 2010), we see value in comparing the

195   performance of this metric to others currently in use.

196   Finally, we considered a Bayesian, model-based approach to estimating tip rates. BAMM

197   (Rabosky 2014) assumes that phylogenies are generated by set of discrete diversification

198   regimes. Using MCMC, the program simulates a posterior distribution of rate shift regimes, from

199   which marginal posterior rate distributions can be extracted for each tip in the phylogeny. Priors

200   for BAMM analyses were set using default settings from the setBAMMpriors function from

201   BAMMtools (Rabosky et al. 2014). The prior parameterizations specified by this function ensure

202   that the prior density on relative rate changes across the tree is invariant to the scale of the tree

203   (e.g., multiplying branch lengths by 10^6 will not change inferences about relative rates across the

204   tree). We denote BAMM tip speciation rates (mean of the marginal posterior) as λ_{BAMM}. As

205   BAMM also estimates extinction rates for each regime, we also calculated tip-specific net

206   diversification rate as $\lambda_{BAMM} - \mu_{BAMM}$, denoted as r_{BAMM}.

207

208   <u>Tip rate metrics estimate speciation, not net diversification</u>

209   As suggested previously (Belmaker & Jetz 2015; supplemental analyses in Jetz et al.

210   2012), DR and presumably other tip-based measurements, more accurately estimate the rate of

211   speciation than the rate of net diversification. However, numerous studies continue to refer to DR

212   as a measure of net diversification (Marin & Hedges 2016; Oliveira et al. 2016; Cai et al. 2017;

213   Quintero & Jetz 2018; and many others). This is incorrect and it is straightforward to

214   demonstrate that λ_{TB}, λ_{ND} and λ_{DR} are more reliable measures of speciation rates and not net

215   diversification rates, at least when extinction is moderate to high.

216   To illustrate this property of the metrics, we applied all approaches to constant-rate birth-

217   death phylogenies simulated across a range of extinction fractions ($\varepsilon = \mu / \lambda$) , including pure-

218   birth trees ($\varepsilon = 0$) as well as trees exhibiting very high turnover ($\varepsilon = 1$). To evaluate accuracy of

219   speciation estimates as a function of ε, we generated 1000 phylogenies with 100 tips each, where

220   λ and ε were drawn from uniform distributions (λ: [0.05, 0.3]; ε: [0, 1]). Importantly, when λ is

221   sampled uniformly with respect to ε, the distribution of r is not uniform: the mean, range and

222   variance in r decrease dramatically as ε increases. To evaluate the accuracy of r as a function of

223   ε, we thus generated a second set of trees by sampling r and ε from uniform distributions (r:

224   [0.05, 0.3], ε [0, 1]). As a result, λ has constant mean and variance with respect to ε in the first

225   set of simulations, and the same is true for r in the second set of simulations (Figure S1). All

226   phylogeny simulations were conducted with the TreeSim package in R (Stadler 2011).

227   We compared tip rate metrics to true speciation rates λ_{TRUE} (with the first simulation set)

228   and to true net diversification rates r_{TRUE} (with the second simulation set). We evaluated mean

229   per-tip accuracy of the tip rate metrics with three measures of error:

230    $$\text{mean absolute error} = \sum_{i=1}^{N_i} |\lambda_i - \lambda_{TRUE_i}| / N$$

231    $$\text{RMSE} = \sqrt{\sum_{i=1}^{N_i} (\lambda_i - \lambda_{TRUE_i})^2 / N}$$

232    $$\text{mean proportional error} = \sum_{i=1}^{N_i} \frac{\lambda_i - \lambda_{TRUE_i}}{\lambda_{TRUE_i}} / N$$

233   where λ_i is the estimated tip rate for species i out of N total species, λ_{TRUE} is the true tip rate.

234   Mean absolute error and root mean square error capture the magnitude in error in tip rates, and

235   mean proportional error quantifies the bias in tip rates, as a function of the true tip rates

236   (Rabosky et al. 2014). In analyses below, all error summaries yield generally congruent results;

237   results for mean absolute error are presented in the main text, and others in supplement.

238

239   <u>Assessment of tip rate metrics</u>

240   We tested the performance of the metrics by compiling publicly-available datasets from a

241   number of simulation-based studies (Table 1). By focusing on simulations from previously-

242   published work, we thus ensured that the simulation process itself was effectively blinded to the

243   objectives of the present study. We further note that our trial datasets included several studies

244   that were critical of BAMM (Moore et al. 2016, Meyer & Wiens 2017). These simulated trees

245   include rate heterogeneity in time and across lineages. Together, these phylogenies present a

246   wide range of tree sizes and diversification rate shifts, providing an ideal comparative dataset for

247   our purposes. To more easily distinguish between these tree types in the text, we refer to the

248   BAMM-type, multi-regime time-constant phylogenies simply as "multi-regime", and the multi-

249   regime diversity-dependent phylogenies simply as "diversity-dependent", even though discrete

250   rate shifts are present in both types of trees. In addition to discrete-shift scenarios (e.g., BAMM-

251   type process), we simulated phylogenies under an "evolving rates" model of diversification

252   (Rabosky 2010; as corrected in Beaulieu & O'Meara 2015) to explore performance of tip rate

253   metrics when diversification rates change continuously and independently along branches, as

254   might occur if diversification rates are correlated with an underlying continuous trait (FitzJohn

255   2010). In these simulations, we allowed the logarithm of λ to evolve across the tree under a

256   Brownian motion process, while holding ε constant. The magnitude of rate heterogeneity among

257   branches is controlled by the diffusion parameter σ, where greater values lead to greater

258   heterogeneity in speciation rates. Although published phylogenies with rate data were

259   unavailable for this simulation scenario, we used simulation code and parameters taken directly

260   from Beaulieu & O'Meara (2015) to generate trees with similar statistical properties to those in

261   their study. Simulations were performed with the following parameters: $\lambda = 0.078, 0.103, 0.145,$

262   0.249 and $\varepsilon = 0.0, 0.25, 0.50, 0.75$. We simulated 100 phylogenies for each (λ , ε) pair, and for

263   three values of σ ($\sigma = 0.03, 0.06, 0.12$). We evaluated tip rate accuracy by comparing estimated

264   to true tip rates, using the absolute and proportional error metrics described above. We also

265   examined the correlation between true and estimated tip rates, combining tip rates from all

266   phylogenies generated under the same class of diversification process, and visualizing these data

267   as density scatterplots, generated with the LSD package in R (Schwalb et al. 2018), where colors

268   indicate the density of points.

269   Size of diversification rate regimes might be an important factor in a tip rate metric's

270   ability to accurately estimate rates. For example, BAMM's statistical power in detecting a shift

271 to a new rate regime is a function of the number of taxa in that rate regime, and tip rates for taxa

272 from small regimes will more likely be parameterized according to the larger parent regime or

273 the tree-wide average rate (Rabosky, Mitchell & Chang 2017); this is the expected behavior

274 when BAMM fails to identify a rate shift. However, non-model-based approaches such as those

275 examined in this study might be more accurate for small regimes. To explore how rate regime

276 size influences the accuracy of tip rate metrics, we calculated the mean tip rate for each true rate

277 regime from all multi-regime phylogenies (simulation datasets from Moore et al. 2016; Rabosky,

278 Mitchell & Chang 2017; Meyer & Wiens 2017; Mitchell, Etienne & Rabosky 2018). We then

279 calculated the Pearson correlation coefficient and the slope of a linear model between true and

280 estimated mean regime rates. We explored the performance of all metrics with respect to regime

281 sample size, as in Rabosky, Mitchell & Chang (2017: Figure 13). For comparison, we repeated

282 all performance summaries on tip rates estimated by applying a simple constant-rate birth-death

283 (CRBD) process to each simulated phylogeny. This exercise is an important control, because it

284 indicates how much error we would expect for each simulated phylogeny under the simplifying

285 (incorrect) assumption that rates are constant among lineages and through time for each dataset.

286

287 **__Results__**

288 <u>Speciation or net diversification?</u>

289   As expected, the tip rate metrics examined in this study are more accurate estimators of

290 the rate of speciation (λ) and not the net rate of species diversification (r). Mean absolute error

291 increased exponentially with respect to the extinction fraction ε (Figure 1). However, mean

292 absolute error in speciation rate was largely invariant with respect to ε (0.95 quantile of r-based

293 and λ-based mean absolute error for λ_{DR}: 2.28 and 0.17, respectively). Nearly identical patterns

294 were found with RMSE (Figure S2). Note that r and λ for these simulations were drawn from

295 identical uniform distributions, and absolute error in the rates is thus comparable. Proportional

296 error generally exhibited the same pattern, and in terms of λ versus r, differences in speciation-

297 based error varied across ε (Figure S3). There was a weak but significant trend towards

298 progressively greater underestimation of speciation rates with increasing values of relative

299 extinction (linear model slopes: -0.08, -0.014, -0.011 for λ_{ND}, λ_{DR} and λ_{BAMM}, respectively).

300 Overall, error was highest for λ_{TB} by an order of magnitude (Figure S4), and decreased

301 progressively with λ_{ND} and λ_{DR}, with the lowest overall error in λ_{BAMM}. BAMM estimates of net

302 diversification rate were relatively accurate, except at the highest values of ε (Figures 1, S2, S3).

303

304 <u>Tip rate accuracy across rate-variable phylogenies</u>

305   Tip rates estimated with BAMM were consistently more accurate than those obtained

306 using the other methods across all diversification scenarios considered, including multi-regime,

307 diversity-dependent and evolving rates trees (Figure 2). λ_{DR} was the second-most accurate

308 metric, although its relationship with true rates was substantially weaker than λ_{BAMM}. λ_{ND} and

309 λ_{TB} were correlated with true rates but performed relatively poorly overall. However, λ_{TB}

310 performed better than λ_{ND}, and just as well as λ_{DR} at estimating speciation rates for diversity-

311 dependent trees (Figure 2, S5). All metrics performed best for multi-regime trees, followed by

312 evolving rates and diversity-dependent trees, respectively. For diversity-dependent trees, λ_{ND}

313 rates are effectively uncorrelated with the true rates (Figure 2). Additionally, the performance of

314 the different tip rate metrics for multi-regime phylogenies is not sensitive to the source of the

315 simulated phylogenies (Figure S6). We found that BAMM substantially outperformed all other

316 metrics on datasets from studies that independently assessed BAMM's performance (Figure S6:

317 Moore et al. 2016; Meyer & Wiens 2017). Tip rates were also generally but more weakly

318 correlated with true net diversification rates, with the exception of λ_{ND}, which was uncorrelated

319 with true rates for diversity-dependent trees, presumably because this metric equally weights the

320 full depth of the tree (Figure S7).

321   In terms of mean per-tip error, λ_{BAMM} consistently outperformed the other metrics for

322 multi-regime, diversity-dependent and evolving rates trees (Figures 3, S8). Error in λ_{BAMM}

323 increased as a function of rate heterogeneity for evolving rate phylogenies, but was largely

324 independent of the magnitude of rate heterogeneity for the other scenarios. λ_{DR} generally

325 exhibited greater error than λ_{BAMM}, and this error increased as a function of the level of

326 heterogeneity for both the evolving rates and multi-regime trees. Error in λ_{DR} was generally

327 invariant to the number of rate regimes for the diversity-dependent scenarios. However, λ_{DR}

328 tended to have greater error than tip estimates from a simple model that assumes no variation in

329 rates through time or among lineages (λ_{CRBD}; all tips assigned the tree-wide CRBD rate). λ_{ND}

330 performed somewhat similarly to λ_{DR} for constant-rate and evolving rates trees, but worse for

331 diversity-dependent trees. Error in λ_{TB} increased with increasing rate heterogeneity for constant-

332 rate and evolving rates trees, but was relatively unaffected by rate heterogeneity in diversity-

333 dependent trees (Figure S9). However, error for this metric was far greater than for all other tip

334 metrics.

335

336 <u>Effects of regime size on performance</u>

337   Both metrics of performance assessment – the Pearson correlation and OLS slope –

338 generally increased with increasing regime size (Figure 4). This was found to be true for all tip

339 rate metrics, although λ_{TB} and λ_{ND} never achieved high performance. λ_{DR} tended to perform

340 better than other metrics when small rate regimes were included (e.g., 10 tips or fewer);

341 however, the slope between estimated and true rates was greater than 1 across the majority of

342 minimum regime sizes, indicating that λ_{DR} overestimates speciation rates (see also Figure S3).

343 Similar patterns were observed for net diversification rates with λ_{DR}, but the magnitude of the

344 overestimation was greater than for speciation (Figure S10). λ_{BAMM}, in contrast, approached a

345 slope of 1 when estimating speciation rates and slightly underestimated net diversification rates

346 (regimes with > 30 tips: OLS slope = 0.96 for λ, 0.87 for r).

347   Absolute error in regime mean tip rates was lowest for λ_{DR} and λ_{BAMM}, regardless of the

348 size of the rate regime (Figure 5). BAMM's ability to accurately estimate tip rates improved with

349 regime size, whereas absolute error was relatively consistent across regime sizes for λ_{DR} for

350 regimes greater than 10 species. We also found that λ_{DR} slightly outperformed λ_{BAMM} for small

351 rate regimes.

352   Note that, in Figures 4 and 5, each rate regime is treated as a single data point. Rate

353 regimes of sizes 1000, 100, and 1 tip are equivalent under this method of error assessment.

354 Figure 4 assesses how well these methods estimate rates for individual regimes, regardless of the

355 size of those regimes. In contrast, Figures 1-3 ask how well these methods perform at estimating

356 rates for a given tip.

357

358 **Discussion**

359   We assessed several tip rate metrics and confirmed that these are more accurate

360 estimators of the rate of speciation, rather than net diversification (Figures 1, 4, S7, S10). This

361 distinction was especially pronounced at high extinction fractions, where the rate of lineage

362 turnover is high, and rates of speciation and net diversification have the potential to be more

363 divergent. These results are consistent with supplemental analyses performed in Jetz et al. 2012.

364 It is also important to note that recent extinction will have a much greater influence on these

365 metrics than extinction events deeper in time (Quental & Marshall 2011). Net diversification rate

366 is a critical determinant of species richness, yet this quantity is potentially independent of the

367 underlying rate of speciation. Misinterpretation of tip rate metrics could therefore lead to highly

368 misleading perspectives on large-scale diversity dynamics. As we demonstrate (Figures 1, S2,

369 S3), tip rate metrics (λ_{ND}, λ_{DR}) provide relatively little information about net diversification, and

370 high values of these metrics are fully consistent with equilibrial models of speciation where the

371 true net diversification rate is zero. Thus, λ_{DR} and λ_{ND} should not be used to support claims

372 about the dynamics of species richness or net diversification *per se* without independent

373 evidence bearing on plausible levels of extinction.

374   In terms of accuracy, we found that BAMM performed better than non-model-based

375 metrics across all datasets we considered: estimated tip rates were most highly correlated with

376 true tip rates, and mean per-tip error in rates was lower across a range of rate-variable simulation

377 scenarios. This performance is likely to be at least partially due to the inclusion of extinction in

378 the BAMM inference model. BAMM is expected to perform well for phylogenies with discrete

379 shifts in diversification rates as this type of rate variation is most consistent with BAMM's

380 assumptions (Rabosky 2014; Mitchell & Rabosky 2016; Rabosky, Mitchell & Chang 2017;

381 Mitchell, Etienne & Rabosky 2018). However, BAMM performed surprisingly well for the

382 evolving rates phylogenies, which conform poorly to the assumptions of the inference model. In

383 these trees, the rate of speciation changes continuously under a diffusion process, and as a result,

384 the phylogeny exhibits rate heterogeneity without discrete rate shifts.

385   On evolving rates phylogenies, λ_{BAMM} performed better than λ_{DR} (Figure 2; Spearman's

386 ρ for λ_{BAMM} = 0.83, ρ for λ_{DR} = 0.62), despite the fact that λ_{DR} does not rely on the detection of

387 distinct rate regimes to estimate tip rates (Figure 5). λ_{BAMM} also exhibited the lowest mean per-

388 tip error across varying levels of rate heterogeneity (Figure 3).

389   Why do λ_{BAMM} and λ_{DR} exhibit such striking differences in performance across the

390 simulation scenarios considered here? To illustrate the differences between inference under these

391 metrics, we compared true tip rates to λ_{BAMM} and to λ_{DR} on a simulated birth-death tree with a

392 single rate shift (Figure 6), as well as on one evolving rates tree simulated for this study (Figure

393 7). It is clear that if BAMM has the statistical power to detect true rate shifts, then it will perform

394 well under rate shift scenarios. In contrast, λ_{DR} tracks true rate shifts but exhibits high sample

395 variance. With an evolving rates tree (Figure 7), the simulation model is very different from the

396 inference model in BAMM. However, it conservatively places rate shifts in order to

397 accommodate rate heterogeneity that is spread across the phylogeny under a rather different

398 model of rate variation. λ_{DR} also broadly tracks the overall pattern of the true rates, but the

399 variance in the corresponding estimates is so high that performance is negatively affected. If we

400 calculate mean (absolute) per-tip error in λ_{BAMM} and $\lambda_{DR,}$ the error is relatively similar between

401 λ_{BAMM} and λ_{DR}, but the variance in per-tip error for λ_{DR} is higher. Overall, BAMM exhibited

402 substantially lower error than λ_{DR} under precisely this scenario (Figure 3).

403   Thus, although BAMM is conservative in the estimation of tip rates relative to λ_{DR}, the

404 method exhibits lower overall error. It appears that λ_{DR} can recover more subtle rate

405 heterogeneity relative to BAMM (see Rabosky, Mitchell & Chang 2017 for discussion of power

406 in BAMM), but this apparent power advantage comes at the cost of increased variance (error) in

407 the resulting estimates. Remarkably, on a per-tip basis, we find that a simple constant-rate birth-

408 death process (λ_{CRBD}) frequently yields tip estimates with lower median error and less error

409 variance than those obtained with λ_{DR} (Figure 3), despite the simplifying (and incorrect)

410 assumption that rates are identical across all tips in a given tree. For example, across all multi-

411 regime simulations (Figure 3), λ_{CRBD} point estimates were more accurate than the corresponding

412 λ_{DR} point estimates for 84% of trees in the simulations; for λ_{BAMM}, the λ_{CRBD} estimates were

413 more accurate for a much smaller fraction of the total (36%). Similar results were noted for

414 diversity-dependent (λ_{CRBD} more accurate than 98% of λ_{DR} estimates, versus 15% of λ_{BAMM}

415 estimates) and evolving rates trees (λ_{CRBD} more accurate than 93% of λ_{DR} estimates, versus 36%

416 of λ_{BAMM} estimates). Given that λ_{DR} can and does track true heterogeneity in speciation rate

417 (Figures 6, 7), this pattern suggests that the metric is especially sensitive to the stochastic

418 variation in branch lengths that can emerge even when all tips have the same underlying

419 speciation rate.

420   Regardless of the performance summaries presented in this article, important questions

421 remain with respect to how well tip rate metrics can estimate the true rate of speciation from

422 empirical phylogenies. The phylogenies analyzed in this study were simulated under idealized

423   processes and neglect potential biases and sources of uncertainty that are present in real datasets.

424   For example, if the process of speciation takes time to complete, as is generally believed to be

425   the case (i.e., the protracted speciation process; Rosindell et al. 2010; Etienne and Rosindell

426   2012), then the most recent speciation events may still be on-going at the present and typical

427   species-level molecular phylogenies may fail to recognize these events. This will lead to an

428   overestimation of terminal branch lengths, as some terminal branches potentially include

429   incipient species. A related bias might arise due to incomplete taxon sampling, which

430   disproportionately affects the length of terminal (or otherwise recent) branch lengths (Pybus &

431   Harvey 2000). Likewise, variation in taxonomic practice across a phylogeny might lead to

432   spurious rate variation, particularly if different species concepts are used, or if some clades in the

433   phylogeny – but not others – have been subject to population genetic analysis or screens for

434   cryptic species diversity. Additionally, it has been shown that BAMM and other methods may

435   fail to infer accurate speciation rate dynamics if the phylogeny is in diversity decline – that is,

436   when extinction rates increase towards the present and ultimately exceed speciation rates

437   (Quental & Marshall 2011; Burin et al. 2018). A major, if obvious, caveat in the interpretation of

438   tip rates is that they apply to recent speciation rates and are necessarily limited with respect to

439   inferences about historical variation in speciation rate.

440   The greater the importance of the terminal branches in tip rate metrics, the greater the

441   impact these biases might have on tip rate estimates. On one end of the spectrum, metrics such as

442   λ_{TB} will be very sensitive to such biases as they rely exclusively on terminal branch lengths.

443   Such approaches may retain utility as summary statistics (e.g., Bromham, Hua & Cardillo 2016),

444   but we found that λ_{TB} exhibited the greatest amount of error in estimating speciation rates. On

445   the other end of the spectrum, a metric like λ_{ND} would be minimally impacted as this metric is

446   attempting to capture an average speciation rate over an entire root-to-tip path and does not

447   upweight the contribution of recent branch lengths. λ_{DR} is likely somewhere in the middle of this

448   spectrum, as it gives decreasing weight to branches towards the root. λ_{BAMM} is potentially

449   sensitive to such issues as well, although it may be possible to analytically correct for some

450   biases in the mechanics of the model itself (e.g., Rosindell et al. 2010; Etienne and Rosindell

451   2012).

452   Potential empirical biases aside, tip rates present a number of practical advantages in the

453   study of diversification rate variation. First, tip rates can be summarized and compared across

454   non-monophyletic assemblages of species (Jetz et al. 2012; Kennedy et al. 2016; Belmaker &

455   Jetz 2015; Oliveira et al. 2016; Quintero & Jetz 2018; Rabosky et al. 2018), making it possible to

456   summarize rate characteristics of entire communities or regional assemblages of species. Second,

457   estimation of rates at the present should be more robust to the influence of extinction, as

458   extinction can erase the history of lineage splitting deeper in the phylogeny (Nee et al. 1994;

459   Nee, May & Harvey 1994; Rabosky & Lovette 2008). Third, tip-specific rates can be paired with

460   species-specific trait values or geographic attributes in order to test potential trait- or geography-

461   dependent speciation rates (Freckleton, Phillimore & Pagel 2008; Jetz et al. 2012, Rabosky &

462   Goldberg 2017; Harvey & Rabosky 2017). Tip rates make it possible to relax strong assumptions

463   of rate homogeneity within character states, which are inherent to certain trait-dependent models,

464   including BiSSE and GeoSSE (Maddison, Midford & Otto 2007; Goldberg, Lancaster & Ree

465   2011; Ng & Smith 2014). Recent work has provided a conceptually rich and robust interpretive

466   framework for SSE models that does not assume rate-constancy within character states (Beaulieu

467   & O'Meara 2016; Caetano, O'Meara & Beaulieu 2018), but tip rates nonetheless can provide an

468   important check on results obtained with SSE models by providing a direct means of visualizing

469   the relationship between branch lengths and character states (Bromham, Hua & Cardillo 2016;

470   Hua & Bromham 2016; Harvey & Rabosky 2017). Visual inspection of data in this fashion has

471   the potential to reduce false positives by calling attention to potential outliers and other sources

472   of model inadequacy (Maddison & FitzJohn 2014; Rabosky & Goldberg 2015). A final

473   advantage for non-model-based tip rates, especially λ_{DR}, is that they can profitably be applied to

474   extremely large phylogenies: there are few computational limits to using them on phylogenies

475   with tens of thousands of tips or more, in contrast to formal model-based approaches for which

476   BAMM, HiSSE (Hidden State Speciation and Extinction; Beaulieu & O'Meara 2016), and other

477   methods are poorly suited. This computational efficiency also lends itself to more readily

478   accounting for phylogenetic uncertainty, because tip rate metrics can rapidly be computed across

479   posterior distributions of phylogenies and averaged (for example, see Jetz et al. 2012; Rabosky et

480   al. 2018).

481   In summary, tip rates offer a number of theoretical and practical advantages, particularly

482   in the study of associations between traits and diversification. We found that λ_{BAMM}

483   outperformed other metrics evaluated in this study and proved to be relatively accurate, even

484   under diversification scenarios that depart from the BAMM inference model. λ_{DR}

485   underperformed in comparison to λ_{BAMM}, but in many cases still did reasonably well, particularly

486   for small rate regimes. Despite our performance results, λ_{DR} is likely to remain a useful tool in

487   the study of trait- and geography-dependent diversification (Rabosky & Goldberg 2017; Harvey

488   & Rabosky 2017).

489

490   **Acknowledgements**

491   We thank Jonathan Mitchell for help with compiling the simulation datasets evaluated in this

492   study. We thank Francisco Henao Diaz, Matt Pennell, Brian O'Meara, the O'Meara Lab, and an

493   anonymous reviewer for comments on the manuscript. We also thank Michael Grundler, as well

494   as all other members of the Rabosky and Davis Rabosky labs at the University of Michigan for

495   thoughtful discussion. This work was supported in part by a University of Michigan Rackham

496   Predoctoral Fellowship (P.O.T.) and by a Fellowship from the David and Lucile Packard

497   Foundation (D.L.R.).

498

499   **Data accessibility**

500   All trees, data and code necessary to repeat the analyses have been deposited in the Dryad

501   repository: https://datadryad.org/resource/doi:10.5061/dryad.5hr25nv (Title & Rabosky 2019).

502   **Author contributions**

503   P.O.T. and D.L.R. designed the project. P.O.T. assembled the datasets and performed all

504   analyses. P.O.T. and D.L.R. wrote the manuscript. Both authors contributed critically to

505   subsequent drafts and approved the final publication.

506

507   <u>References</u>

508   Alfaro, M.E., Faircloth, B.C., Harrington, R.C., Sorenson, L., Friedman, M., Thacker, C.E.,

509   Oliveros, C.H., Černý, D. & Near, T.J. (2018). Explosive diversification of marine fishes at

510   the Cretaceous–Palaeogene boundary. *Nature Ecology & Evolution*, **2**, 688–696.

511   Alfaro, M.E., Santini, F., Brock, C., Alamillo, H., Dornburg, A., Rabosky, D.L., Carnevale, G. &

512   Harmon, L.J. (2009). Nine exceptional radiations plus high turnover explain species

513   diversity in jawed vertebrates. *Proceedings of the National Academy of Sciences*, **106**,

514   13410–13414.

515 Barker, F.K., Burns, K.J., Klicka, J., Lanyon, S.M. & Lovette, I.J. (2013). Going to extremes:

516 contrasting rates of diversification in a recent radiation of new world passerine birds.

517 *Systematic Biology*, **62**, 298–320.

518 Beaulieu, J.M. & O'Meara, B.C. (2016). Detecting Hidden Diversification Shifts in Models of

519 Trait-Dependent Speciation and Extinction. *Systematic Biology*, **65**, 583–601.

520 Beaulieu, J.M. & O'Meara, B.C. (2015). Extinction can be estimated from moderately sized

521 molecular phylogenies. *Evolution*, **69**, 1036–1043.

522 Belmaker, J. & Jetz, W. (2015). Relative roles of ecological and energetic constraints,

523 diversification rates and region history on global species richness gradients. *Ecology Letters*,

524 **18**, 563–571.

525 Bromham, L., Hua, X. & Cardillo, M. (2016). Detecting Macroevolutionary Self-Destruction

526 from Phylogenies. *Systematic Biology*, **65**, 109–127.

527 Burin, G., Alencar, L.R.V., Chang, J., Alfaro, M.E. & Quental, T.B. (2018). How Well Can We

528 Estimate Diversity Dynamics for Clades in Diversity Decline? *Systematic Biology*, **105**, 1–

529 17.

530 Caetano, D.S., O'Meara, B.C. & Beaulieu, J.M. (2018). Hidden state models improve state-

531 dependent diversification approaches, including biogeographical models. *Evolution*.

532 Cai, T., Fjeldså, J., Wu, Y., Shao, S., Chen, Y., Quan, Q., Li, X., Song, G., Qu, Y., Qiao, G. &

533 Lei, F. (2017). What makes the Sino-Himalayan mountains the major diversity hotspots for

534 pheasants? *Journal of Biogeography*, **45**, 640–651.

535 Claramunt, S. (2010). Discovering exceptional diversifications at continental scales: the case of

536 the endemic families of Neotropical suboscine passerines. *Evolution*, **64**, 2004–2019.

537 Coyne, J.A. & Orr, H.A. (2004). *Speciation*. Sinauer, Cambridge.

538   Etienne, R.S. & Haegeman, B. (2012). A Conceptual and Statistical Framework for Adaptive
539   Radiations with a Key Role for Diversity Dependence. *The American Naturalist*, **180**, E75–
540   E89.

541   Etienne, R.S. & Rosindell, J. (2012). Prolonging the Past Counteracts the Pull of the Present:
542   Protracted Speciation Can Explain Observed Slowdowns in Diversification. *Systematic*
543   *Biology*, **61**, 204–213.

544   FitzJohn, R.G. (2010). Quantitative traits and diversification. *Systematic Biology*, **59**, 619–633.

545   FitzJohn, R.G., Maddison, W.P. & Otto, S.P. (2009). Estimating Trait-Dependent Speciation and
546   Extinction Rates from Incompletely Resolved Phylogenies. *Systematic Biology*, **58**, 595–
547   611.

548   Freckleton, R.P., Phillimore, A.B. & Pagel, M. (2008). Relating Traits to Diversification: A
549   Simple Test. *The American Naturalist*, **172**, 102–115.

550   Goldberg, E.E., Lancaster, L.T. & Ree, R.H. (2011). Phylogenetic inference of reciprocal effects
551   between geographic range evolution and diversification. *Systematic Biology*, **60**, 451–465.

552   Gomes, A.C.R., Sorenson, M.D. & Cardoso, G.C. (2016). Speciation is associated with changing
553   ornamentation rather than stronger sexual selection. *Evolution*, **70**, 2823–2838.

554   Harvey, M.G. & Rabosky, D.L. (2017). Continuous traits and speciation rates: Alternatives to
555   state-dependent diversification models (N. Cooper, Ed.). *Methods in Ecology and Evolution*,
556   **12**, 751–10.

557   Hua, X. & Bromham, L. (2016). PHYLOMETRICS: an R package for detecting
558   macroevolutionary patterns, using phylogenetic metrics and backward tree simulation.
559   *Methods in Ecology and Evolution*, **7**, 806–810.

560   Jablonski, D. (2008). Species selection: Theory and data. *Annual Review of Ecology, Evolution,*
561   *and Systematics*, **39**, 501–524.

562 Jetz, W., Thomas, G.H., Joy, J.B., Hartmann, K. & Mooers, A.O. (2012). The global diversity of
563   birds in space and time. *Nature*, **491**, 444–448.

564 Kay, K., Voelckel, C., Yang, J.Y., Hufford, K.M., Kaska, D.D. & Hodges, S.A. (2006). Floral
565   characters and species diversification. *Ecology and Evolution of Flowers* (eds L. Harder & S.
566   Barrett), pp. 311–325. Oxford.

567 Kennedy, J.D., Borregaard, M.K., Jønsson, K.A., Holt, B., Fjeldså, J. & Rahbek, C. (2016). Does
568   the colonization of new biogeographic regions influence the diversification and
569   accumulation of clade richness among the Corvides (Aves: Passeriformes)? *Evolution*, 38–
570   50.

571 Maddison, W.P. & FitzJohn, R.G. (2014). The Unsolved Challenge to Phylogenetic Correlation
572   Tests for Categorical Characters. *Systematic Biology*, **64**, 127–136.

573 Maddison, W.P., Midford, P.E. & Otto, S. (2007). Estimating a binary character's effect on
574   speciation and extinction. *Systematic Biology*, **56**, 701.

575 Marin, J. & Hedges, S.B. (2016). Time best explains global variation in species richness of
576   amphibians, birds and mammals. *Journal of Biogeography*, **43**, 1069–1079.

577 Meyer, A.L.S. & Wiens, J.J. (2017). Estimating diversification rates for higher taxa: BAMM can
578   give problematic estimates of rates and rate shifts. *Evolution*, **72**, 39–53.

579 Mitchell, J.S. & Rabosky, D.L. (2016). Bayesian model selection with BAMM: effects of the
580   model prior on the inferred number of diversification shifts. *Methods in Ecology and*
581   *Evolution*, **8**, 37–46.

582 Mitchell, J.S., Etienne, R.S. & Rabosky, D.L. (2018). Inferring Diversification Rate Variation
583   From Phylogenies With Fossils. *Systematic Biology*.

584 Mittelbach, G.G., Schemske, D.W., Cornell, H.V., Allen, A.P., Brown, J.M., Bush, M.B.,
585   Harrison, S.P., Hurlbert, A.H., Knowlton, N., Lessios, H.A., McCain, C.M., McCune, A.R.,
586   McDade, L.A., McPeek, M.A., Near, T.J., Price, T.D., Ricklefs, R.E., Roy, K., Sax, D.F.,

587   Schluter, D., Sobel, J.M. & Turelli, M. (2007). Evolution and the latitudinal diversity

588   gradient: speciation, extinction and biogeography. *Ecology Letters*, **10**, 315–331.

589   Moen, D. & Morlon, H. (2014). Why does diversification slow down? *Trends in Ecology and*

590   *Evolution*, **29**, 190–197.

591   Moore, B.R., Höhna, S., May, M.R., Rannala, B. & Huelsenbeck, J.P. (2016). Critically

592   evaluating the theory and performance of Bayesian analysis of macroevolutionary mixtures.

593   *Proceedings of the National Academy of Sciences*, **113**, 9569–9574.

594   Near, T.J., Dornburg, A., Kuhn, K.L., Eastman, J.T., Pennington, J.N., Patarnello, T., Zane, L.,

595   Fernandez, D.A. & Jones, C.D. (2012). Ancient climate change, antifreeze, and the

596   evolutionary diversification of Antarctic fishes. *Proceedings of the National Academy of*

597   *Sciences*, **109**, 3434–3439.

598   Nee, S., Holmes, E.C., May, R.M. & Harvey, P.H. (1994a). Extinction Rates Can Be Estimated

599   From Molecular Phylogenies. *Philosophical Transactions of the Royal Society B: Biological*

600   *Sciences*, **344**, 77–82.

601   Nee, S., May, R.M. & Harvey, P.H. (1994b). The reconstructed evolutionary process.

602   *Philosophical Transactions of the Royal Society B: Biological Sciences*, **344**, 305–311.

603   Nee, S., Mooers, A. & Harvey, P.H. (1992). Tempo and Mode of Evolution Revealed From

604   Molecular Phylogenies. *Proceedings of the National Academy of Sciences*, **89**, 8322–8326.

605   Ng, J. & Smith, S.D. (2014). How traits shape trees: new approaches for detecting character

606   state-dependent lineage diversification. *Journal of Evolutionary Biology*, **27**, 2035–2045.

607   Oliveira, B.F., Machac, A., Costa, G.C., Brooks, T.M., Davidson, A.D., Rondinini, C. &

608   Graham, C.H. (2016). Species and functional diversity accumulate differently in mammals.

609   *Global Ecology and Biogeography*, **25**, 1119–1130.

610   Pybus, O.G. & Harvey, P.H. (2000). Testing macro-evolutionary models using incomplete

611   molecular phylogenies. *Proceedings of the Royal Society of London. Series B: Biological*

612   *Sciences*, **267**, 2267–2272.

613 Quental, T.B. & Marshall, C.R. (2011). The Molecular Phylogenetic Signature of Clades in
614   Decline. *PLoS ONE*, **6**, e25780–9.

615 Quintero, I. & Jetz, W. (2018). Global elevational diversity and diversification of birds. *Nature*,
616   **555**, 246–250.

617 Rabosky, D.L. (2014). Automatic detection of key innovations, rate shifts, and diversity-
618   dependence on phylogenetic trees. *PLoS ONE*, **9**, e89543.

619 Rabosky, D.L. (2010). Extinction rates should not be estimated from molecular phylogenies.
620   *Evolution*, **64**, 1816–1824.

621 Rabosky, D.L. & Goldberg, E.E. (2017). FiSSE: A simple nonparametric test for the effects of a
622   binary character on lineage diversification rates. *Evolution*, **106**, 13410–11.

623 Rabosky, D.L. & Goldberg, E.E. (2015). Model Inadequacy and Mistaken Inferences of Trait-
624   Dependent Speciation. *Systematic Biology*, **64**, 340–355.

625 Rabosky, D.L. & Lovette, I.J. (2008). Explosive evolutionary radiations: decreasing speciation or
626   increasing extinction through time? *Evolution*, **62**, 1866–1875.

627 Rabosky, D.L., Chang, J., Title, P.O., Cowman, P.F., Sallan, L., Friedman, M., Kaschner, K.,
628   Garilao, C., Near, T.J., Coll, M. & Alfaro, M.E. (2018). An inverse latitudinal gradient in
629   speciation rate for marine fishes. *Nature*, **559**, 392–395.

630 Rabosky, D.L., Donnellan, S.C., Grundler, M. & Lovette, I.J. (2014). Analysis and Visualization
631   of Complex Macroevolutionary Dynamics: An Example from Australian Scincid Lizards.
632   *Systematic Biology*, **63**, 610–627.

633 Rabosky, D.L., Grundler, M., Anderson, C., Title, P.O., Shi, J.J., Brown, J.W., Huang, H. &
634   Larson, J.G. (2014). BAMMtools: an R package for the analysis of evolutionary dynamics
635   on phylogenetic trees. *Methods in Ecology and Evolution*, **5**, 701–707.

636 Rabosky, D.L., Mitchell, J.S. & Chang, J. (2017). Is BAMM Flawed? Theoretical and Practical
637   Concerns in the Analysis of Multi-Rate Diversification Models. *Systematic Biology*, **66**,
638   477–498.

639 Rabosky, D.L., Title, P.O. & Huang, H. (2015). Minimal effects of latitude on present-day
640   speciation rates in New World birds. *Proceedings of the Royal Society of London. Series B:*
641   *Biological Sciences*, **282**, 20142889.

642 Redding, D.W. & Mooers, A.O. (2006). Incorporating Evolutionary Measures into Conservation
643   Prioritization. *Conservation Biology*, **20**, 1670–1678.

644 Ricklefs, R.E. (2006). Global variation in the diversification rate of passerine birds. *Ecology*, **87**,
645   2468–2478.

646 Rohde, K. (1992). Latitudinal Gradients in Species Diversity: The Search for the Primary Cause.
647   *Oikos*, **65**, 514–527.

648 Rosindell, J., Cornell, S.J., Hubbell, S.P. & Etienne, R.S. (2010). Protracted speciation revitalizes
649   the neutral theory of biodiversity. *Ecology Letters*, **13**, 716–727.

650 Sanderson, M.J. & Donoghue, M.J. (1996). Reconstructing shifts in diversification rates on
651   phylogenetic trees. *Trends in Ecology and Evolution*, **11**, 15–20.

652 Schluter, D. (2016). Speciation, Ecological Opportunity, and Latitude. *The American Naturalist*,
653   **187**, 1–18.

654 Silvestro, D., Schnitzler, J. & Zizka, G. (2011). A Bayesian framework to estimate
655   diversification rates and their variation through time and space. *BMC Evolutionary Biology*,
656   **11**, 311.

657 Stadler, T. (2011). Simulating Trees with a Fixed Number of Extant Species. *Systematic Biology*,
658   **60**, 676–684.

659 Steel, M. & Mooers, A. (2010). The expected length of pendant and interior edges of a Yule tree.
660   *Applied Mathematics Letters*, **23**, 1315–1319.

661   Title, P. & Rabosky D. (2019). Data from: Tip rates, phylogenies and diversification: what are

662   we estimating, and how good are the estimates? Methods in Ecology and Evolution

663   doi:10.5061/dryad.5hr25nv

664   Zink, R.M., Klicka, J. & Barber, B. (2004). The tempo of avian diversification during the

665   Quaternary. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359**,

666   215–220.

667   **Figures/Captions**

668

669   **Table 1.** Summary of simulated phylogenies used in this study.

simulation model	number of trees	tree size	regime number	source
single-regime, constant-rate birth-death	100	100	1	Mitchell & Rabosky 2016
single- and multi-regime, constant-rate birth-death	100	51-148	1-6	Moore et al. 2016
single- and multi-regime, constant-rate birth-death	400	10-4296	1-67	Rabosky, Mitchell & Chang 2017
multi-regime, constant-rate birth-death	20	939-3708	11	Meyer & Wiens 2017
single- and multi-regime, constant-rate birth-death	188	4-3955	1-73	Mitchell, Etienne & Rabosky 2018
single-regime, constant-rate birth-death, lambda uniform	1000	100	1	this study
single-regime, constant-rate birth-death, net diversification uniform	1000	100	1	this study
pure birth root regime, 1-4 discrete shifts to diversity-dependent regimes	1200	54-882	1-5	Rabosky 2014; Mitchell & Rabosky 2016
speciation rate evolves via diffusion process	1200	25-1208	1	Rabosky 2010; Beaulieu & O'Meara 2015; Rabosky 2016; this study

670

671

672

673

674

675
676
677
678
679
680

681
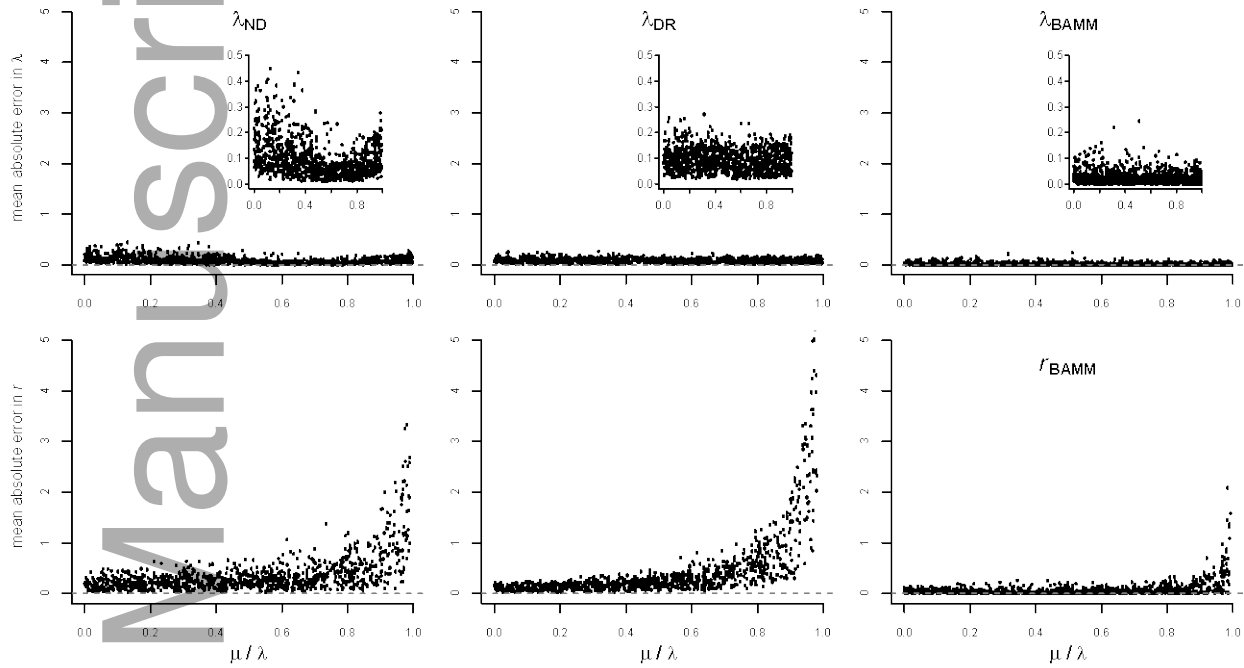
Figure 1. Mean absolute error in λ (top) and *r* (bottom) for three different tip rate metrics, across a range of relative extinction rates. For BAMM, the estimated speciation and net diversification rates are presented in the top and bottom panels, respectively. Absolute error of zero implies perfect accuracy. Inset plots show error in λ with truncated y-axis scale to facilitate comparison among metrics. All tip rate metrics track λ more accurately than they track *r*. See Figure S4 for λ_{TB}, which performed much worse than the other metrics.

688
689
690
691
692

693
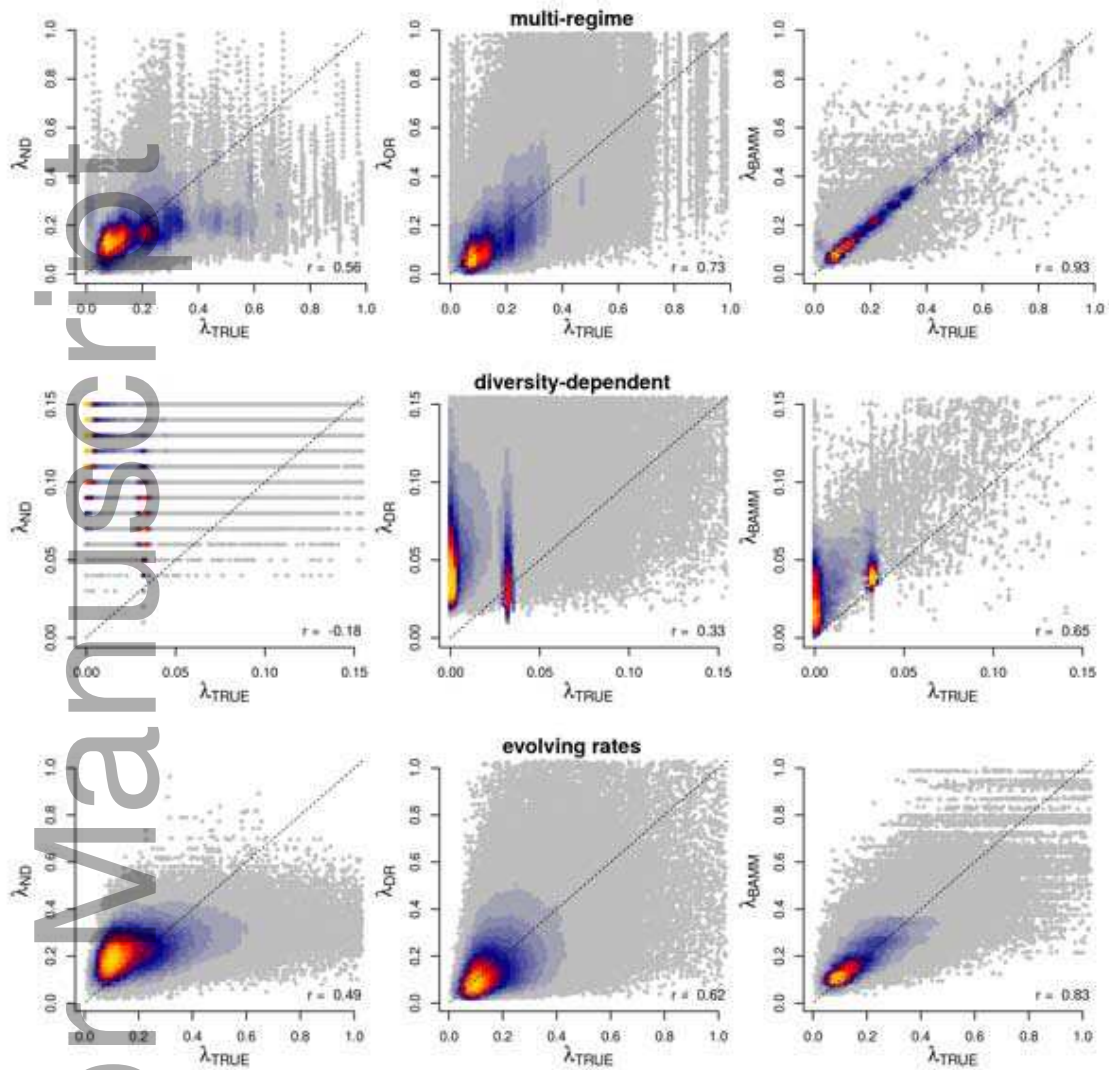
Figure 2. True tip rates (λ_{TRUE}) in relation to estimated tip rates. Tip rates were compared separately for different major categories of phylogeny simulations (rows) and are plotted separately by inference method (columns). Plotting region is restricted to the 99th percentile of true rates, but Spearman correlations between true and estimated rates (lower right of each figure panel) are based on the full range of the data. Colors indicate the density of points in the scatter plots. The horizontal gaps in λ_{ND} for diversity-dependent trees are an artefact of all trees having the same crown age. λ_{BAMM} exhibited the strongest correlation with true rates for all simulation categories.
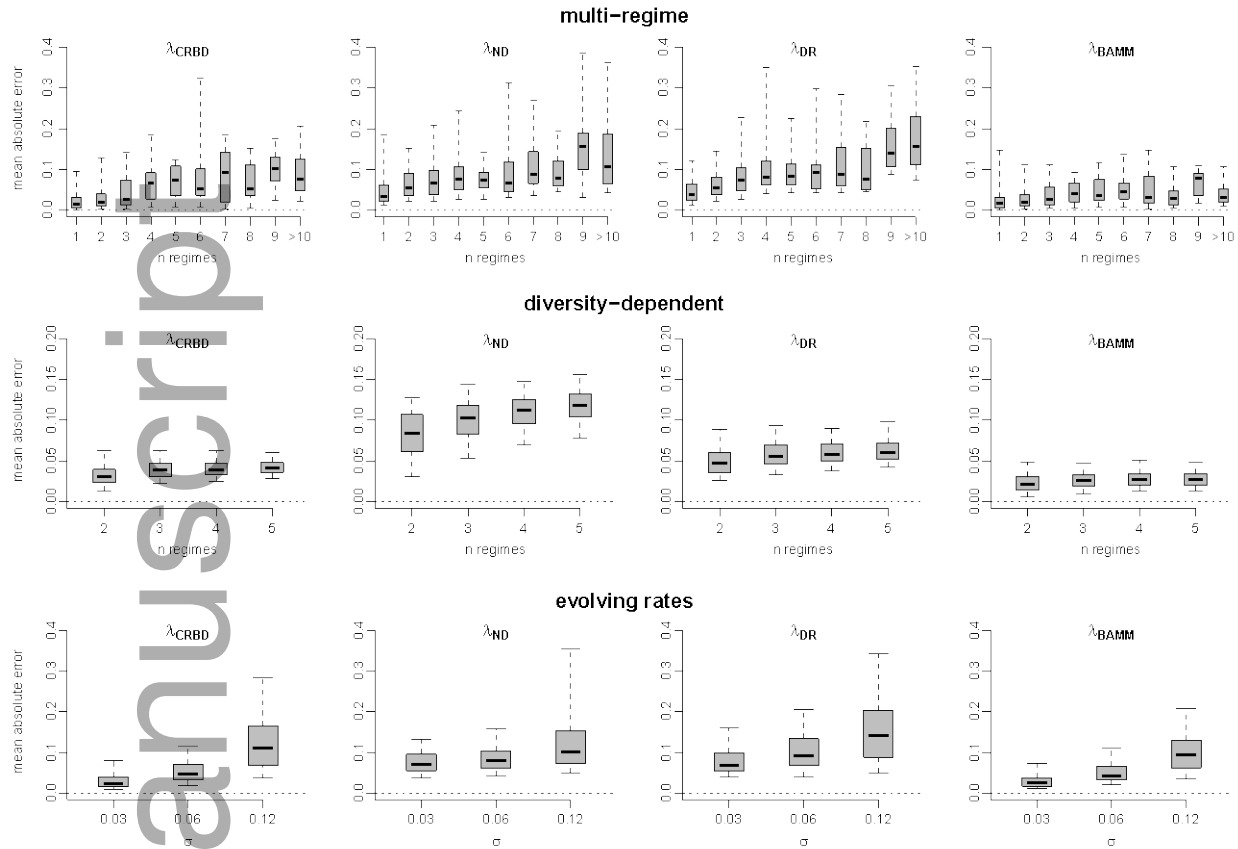
Figure 3. Mean per-tip absolute error in speciation rates as a function of the magnitude of rate heterogeneity in each simulated phylogeny. Results are presented separately for different categories of rate variation (Table 1); left column shows estimates from a constant-rate birth-death model for reference. The boxes and whiskers represent the 0.25 – 0.75, and the 0.05 – 0.95 quantile ranges, respectively. In some cases, λ_{ND} and λ_{DR} had more error than a simple CRBD model with no variation in tip rates. λ_{BAMM} had the least amount of error across all amounts of rate heterogeneity. See Figure S9 for λ_{TB}.
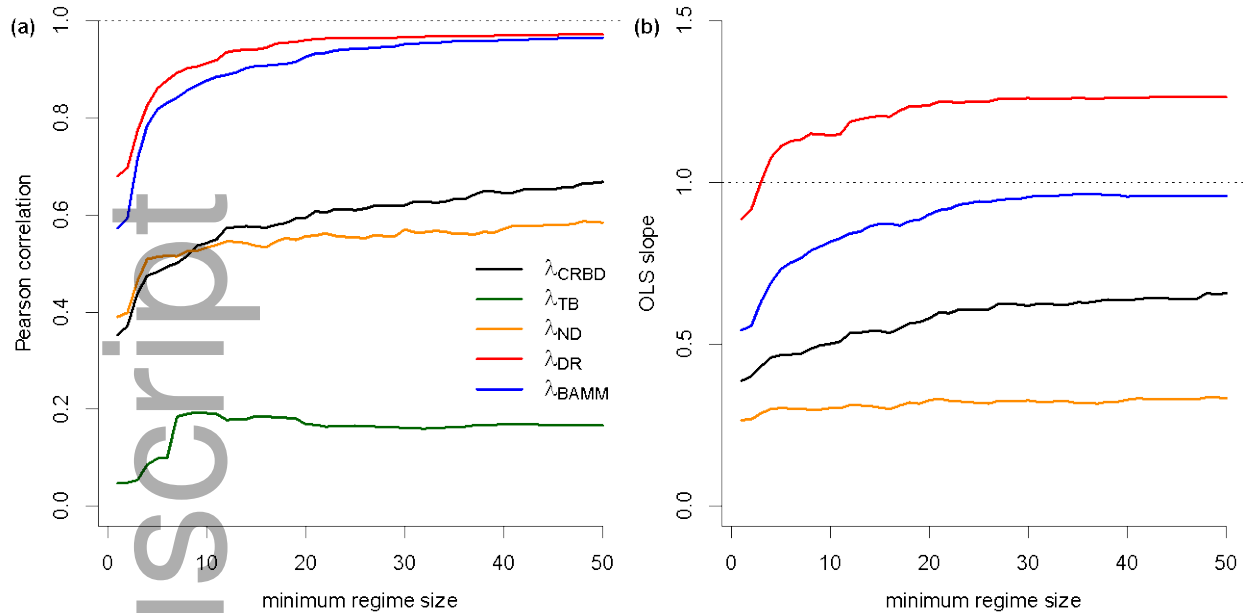
Figure 4. Performance of tip rate metrics as a function of regime size, including Pearson correlation (a) and OLS regression slope (b) for mean rates with respect to λ_{TRUE}. λ_{DR} and λ_{BAMM} outperform the other metrics when summarized in this fashion, although λ_{DR} tends to overestimate the rate of speciation. The x-axis denotes the minimum regime size across which performance was summarized. For example, x = 20 corresponds to the correlations and slopes computed for all regimes with 20 or more tips; a value of x = 1 is the corresponding results for all regimes. The OLS slope for λ_{TB} is not visible as it ranges between 7 and 9.
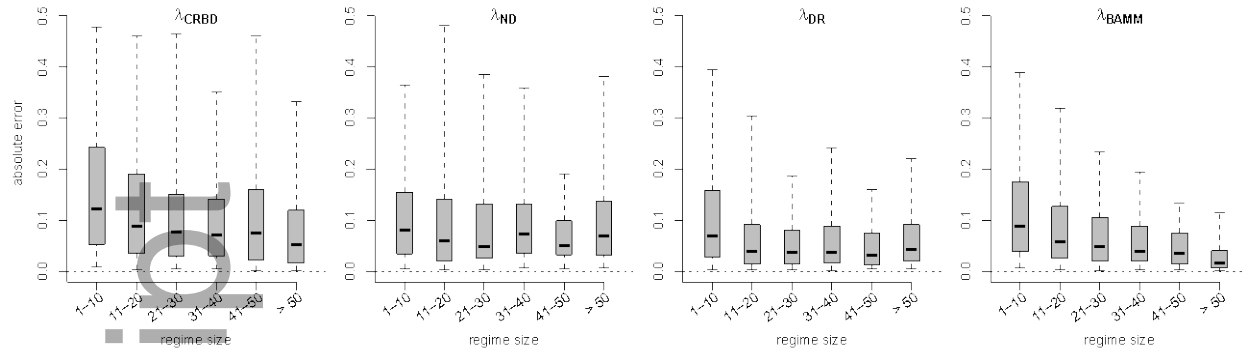
725

726

727 **Figure 5.** Mean per-regime absolute error in relation to true rate regime size, as binned into 10

728 size categories. The boxes and whiskers represent the 0.25 – 0.75, and the 0.05 – 0.95 quantile

729 ranges, respectively. Perfectly estimated rates have an error of zero. λ_{DR} and λ_{BAMM} exhibit the

730 least error when averaged by regimes, and λ_{DR} does slightly better for small clades (10-clade

731 median error 0.07 for λ_{DR}, and 0.08 for λ_{BAMM}).

732

733

734

735

736

737

738

739

740

741

742

743
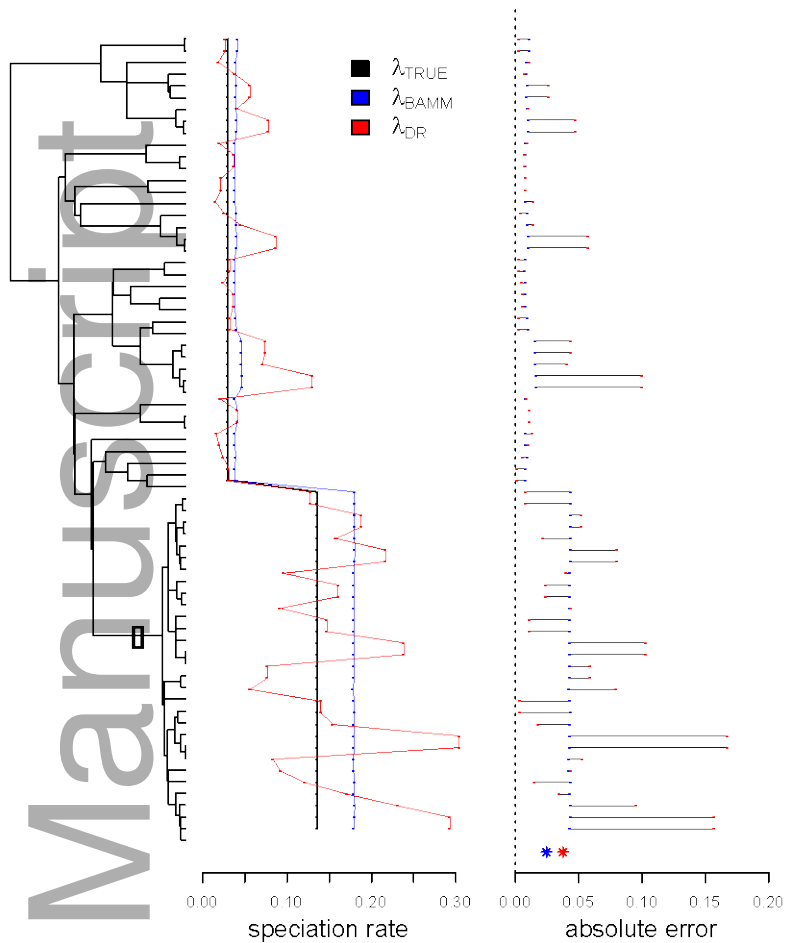
Figure 6. Relationship between λ_{TRUE}, λ_{BAMM}, and λ_{DR} for a simulated phylogeny containing a

single rate shift (orange circle). Subplots to the right of the tree illustrate true and estimated rates

for each tip (left) and corresponding absolute error (right). Asterisks at the bottom denote mean

per-tip error in tip rate metrics. Mean per-tip error is relatively low and similar between λ_{DR} and

λ_{BAMM}, but the sample variance in λ_{DR} tip rates is high. In this example, the variance in absolute

per-tip error in λ_{DR} is 0.002 versus 0.0003 for λ_{BAMM}. On average, λ_{DR} tends to either

overestimate or underestimate rates relative to λ_{BAMM}, even if the mean per-tip error is relatively

low for both metrics.

752

753
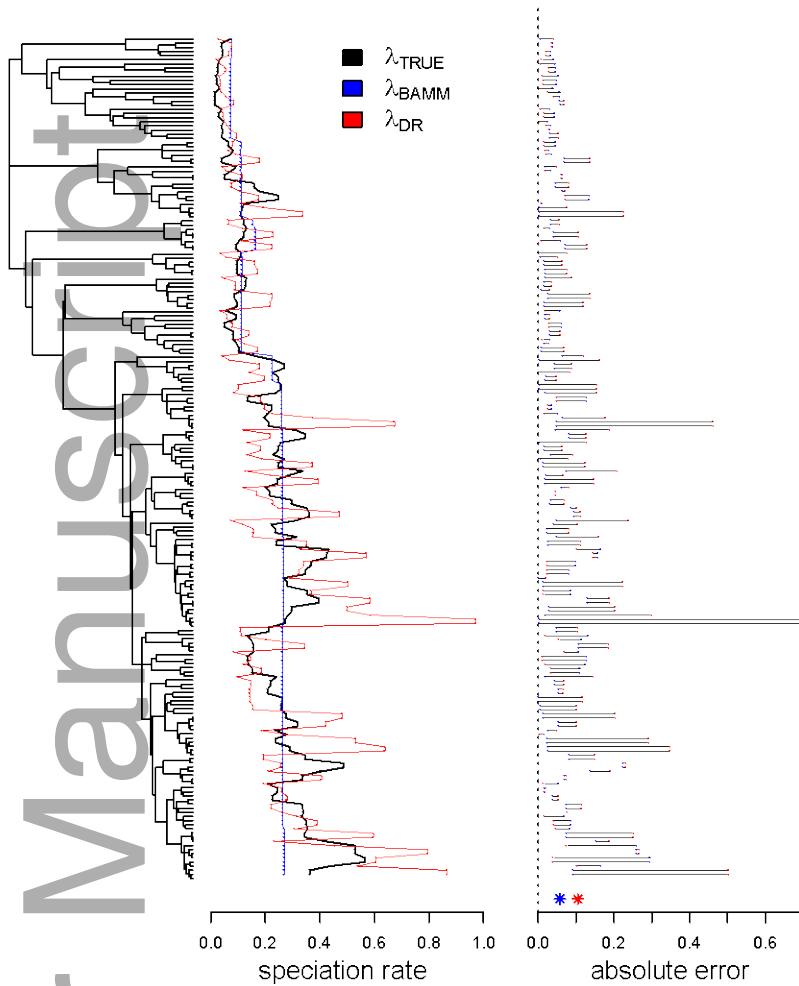
Figure 7. Relationship between λ_{DR}, λ_{BAMM}, and λ_{DR} for a phylogeny simulated under an "evolving rates" model, such that the speciation rate itself varies under a diffusion model. See Figure 6 for additional details. Neither metric is particularly well equipped to infer the true rate variation in this case. However λ_{BAMM}'s conservative estimates are still more accurate relative to λ_{DR}, which is negatively impacted by high variance in tip rates. Here, variance in absolute per-tip error in λ_{DR} is 0.012 versus 0.003 for λ_{BAMM}.
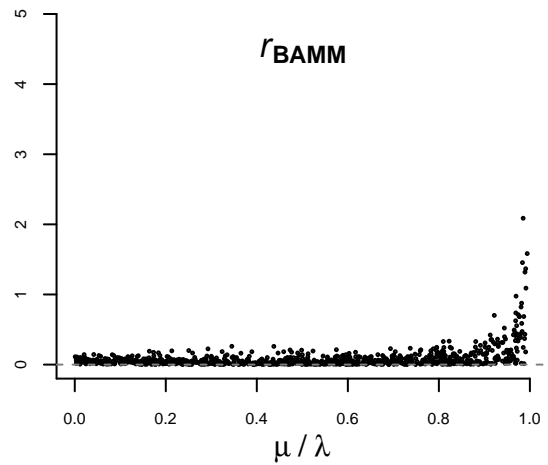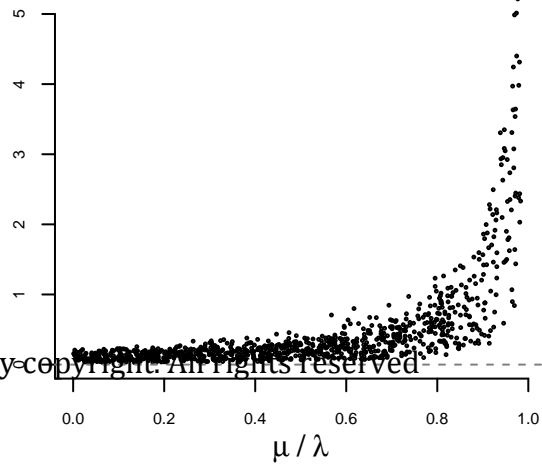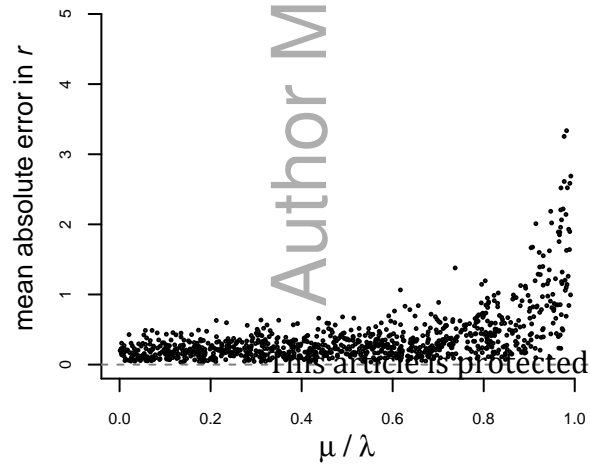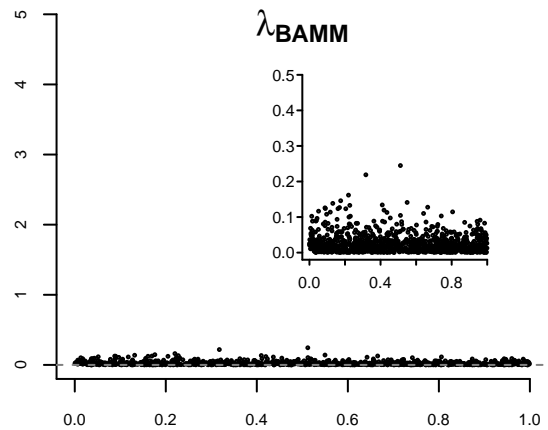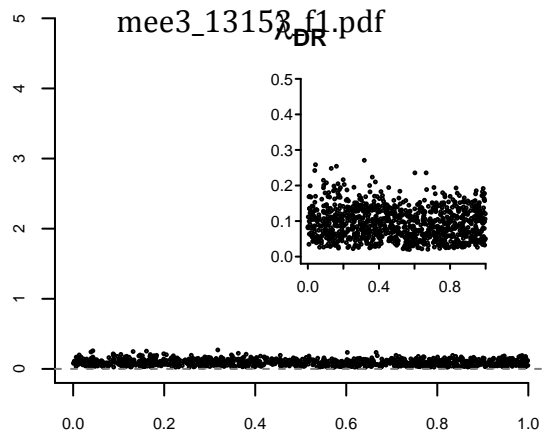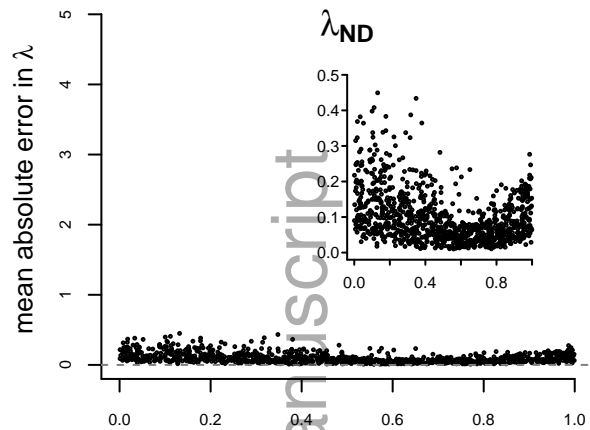
mee3_13152_f1.pdf
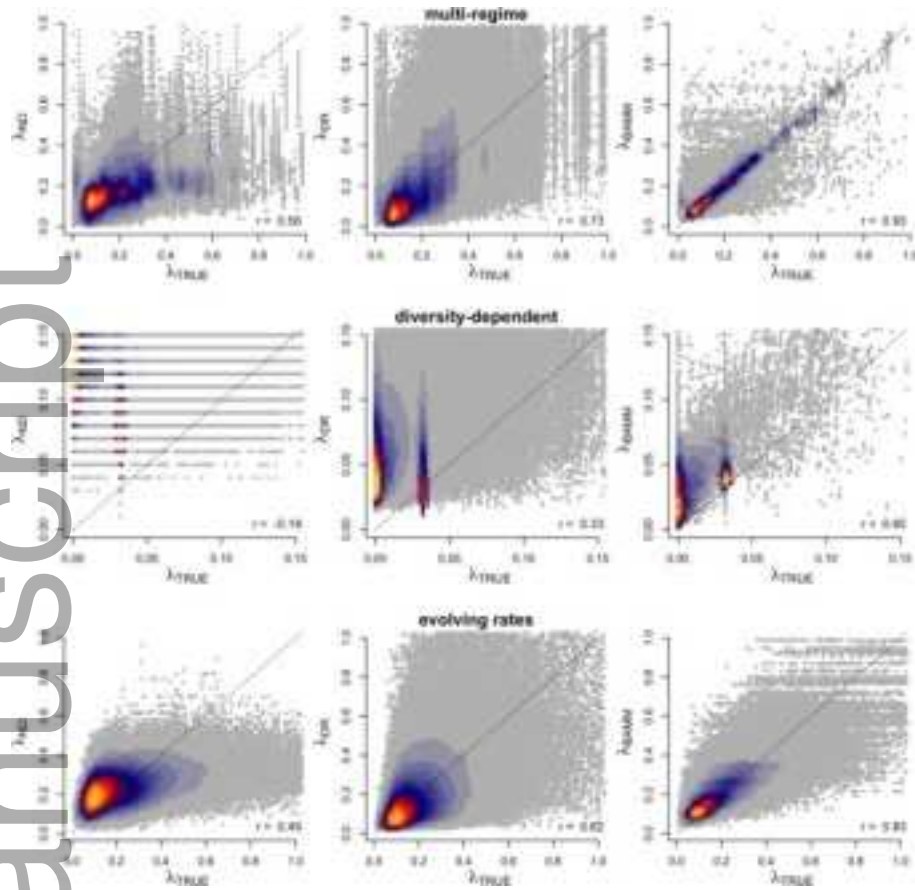
mee3_13153_f2.tiff

multi-regime

diversity–dependent

evolving rates

(a)

mee3_13155_f4.pdf **(b)**

Legend:
- λ_{CRBD} (black)
- λ_{TB} (green)
- λ_{ND} (orange)
- λ_{DR} (red)
- λ_{BAMM} (blue)

Left panel: Pearson correlation vs minimum regime size

Right panel: OLS slope vs minimum regime size