

# RevGadgets: an R Package for visualizing Bayesian phylogenetic analyses from RevBayes

Carrie M. Tribble<sup>1,2,3,\*</sup>, William A. Freyman<sup>4</sup>, Michael J. Landis<sup>5</sup>, Jun Ying Lim<sup>6</sup>, Joëlle Barido-Sottani<sup>7</sup>, Bjørn Tore Kopperud<sup>8,9</sup>, Sebastian Höhna<sup>8,9</sup>, and Michael R. May<sup>1,2</sup>

<sup>1</sup>*Department of Integrative Biology University of California, Berkeley, CA 94709, USA*

<sup>2</sup>*University Herbarium, University of California, Berkeley, CA 94709, USA*

<sup>3</sup>*Current address: School of Life Sciences, University of Hawai'i at Mānoa, Honolulu, HI, 96822, USA*

<sup>4</sup>*23andMe, Inc., Sunnyvale, CA, 94086, USA*

<sup>5</sup>*Department of Biology, Washington University in St. Louis, MO 63130, USA*

<sup>6</sup>*School of Biological Sciences, Nanyang Technological University, Singapore 639798*

<sup>7</sup>*Department of Ecology, Evolution and Organismal Biology, Iowa State University, Ames, IA 50011, USA*

<sup>8</sup>*GeoBio-Center, Ludwig-Maximilians-Universität München, 80333 Munich, Germany*

<sup>9</sup>*Department of Earth and Environmental Sciences, Paleontology & Geobiology, Ludwig-Maximilians-Universität München, 80333 Munich, Germany*

\*E-mail: [ctribble09@gmail.com](mailto:ctribble09@gmail.com)

## Summary

1. Statistical phylogenetic methods are the foundation for a wide range of evolutionary and epidemiological studies. However, as these methods grow increasingly complex, users often encounter significant challenges with summarizing, visualizing, and communicating their key results.
2. We present *RevGadgets*, an R package for creating publication-quality figures from the results of a large variety of phylogenetic analyses performed in *RevBayes* (and other phylogenetic software packages).
3. We demonstrate how to use *RevGadgets* through a set of vignettes that cover the most common use cases that researchers will encounter.
4. *RevGadgets* is an open-source, extensible package that will continue to evolve in parallel with *RevBayes*, helping researchers to make sense of and communicate the results of a diverse array of analyses.

[Bayesian phylogenetics, data visualization, R, RevBayes]

## 1 Introduction

Beyond being a graphical representation of the Tree of Life, phylogenetic trees provide a rigorous basis for a wide range of evolutionary and epidemiological inferences. Phylogenetic methods allow researchers to understand how molecular and morphological traits evolve (Nei, 1987; Yang, 2014; Felsenstein, 1985; Harvey and Pagel, 1991), how lineages disperse over geographic space (Ronquist and Sanmartín, 2011), and how lineages diversify over time (Morlon, 2014), among other evolutionary phenomena. Additionally, phylogenetic methods can be used to inform conservation decisions (Faith,

1992) and are powerful epidemiological tools (Volz et al., 2013; Baele et al., 2017).

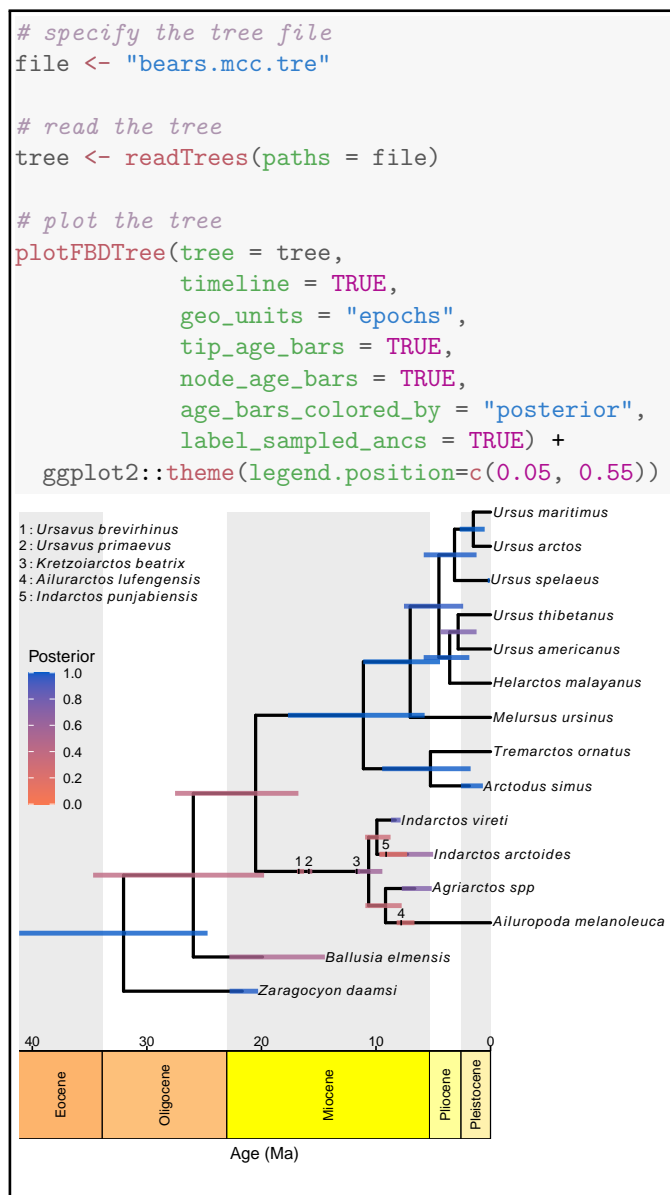
Phylogenetic methods are increasingly based on explicit probabilistic models with parameters that describe underlying evolutionary processes. As datasets grow and evolutionary hypotheses become more nuanced, these models necessarily become more complex. *RevBayes* (Höhna et al., 2016) is a Bayesian phylogenetic inference program that was developed to accommodate this increasing complexity and allows users to explore a vast space of phylogenetic models. Models in *RevBayes* are specified as probabilistic graphical models (Höhna et al., 2014), which are graphical representations of the

26 underlying dependencies among parameters (and their  
 27 corresponding prior distributions), similar to individual  
 28 Legos being used to build a complex city. Using this  
 29 graphical modeling framework, users can design cus-  
 30 tomized models and tailor analyses to their particular  
 31 datasets and research questions. However, this flexibil-  
 32 ity comes at a cost: because of the nearly infinite vari-  
 33 ety of possible models (and model combinations) that  
 34 users can explore in RevBayes, the results of these anal-  
 35 yses are often challenging to summarize and visualize  
 36 using standard software. This is a significant limitation  
 37 for RevBayes users because, in addition to being the pri-  
 38 mary method for reporting results of phylogenetic anal-  
 39 yses, graphical summaries are a valuable tool for mak-  
 40 ing sense of scientific results (Tufte, 2001), and for diag-  
 41 nosing modeling and analytical problems (Kerman et al.,  
 42 2008).

43 Historically, RevBayes users have had to process and  
 44 plot their results using *ad hoc* scripts written for each  
 45 analysis, which imposed a significant barrier to entry for  
 46 users not familiar with the structure of RevBayes out-  
 47 put or comfortable with developing their own graphical  
 48 summaries. To address these challenges, we developed  
 49 RevGadgets. RevGadgets is an R package (R Core Team,  
 50 2020) that adds to the diverse ecosystem of phyloge-  
 51 netic visualization tools—*e.g.*, ape (Paradis and Schliep,  
 52 2019), Tracer (Rambaut et al., 2018), phytools (Rev-  
 53 ell, 2012), ggtree (Yu et al., 2017), FigTree (Rambaut,  
 54 2014), IcyTree (Vaughan, 2017), among many others—  
 55 but is specialized for output produced by RevBayes.  
 56 RevGadgets serves as a bridge between RevBayes anal-  
 57 yses and existing tools for phylogenetic data processing  
 58 and plotting in R, especially the ggtree package suite,  
 59 which includes the ggtree, tidytree, and treeio pack-  
 60 ages (Wang et al., 2020; Yu et al., 2017). RevGadgets pro-  
 61 vides tools for plotting summary trees (including sum-  
 62 maries of parameters for each branch), ancestral-state  
 63 estimates, and posterior distributions of parameters for  
 64 a variety of models. Using the general framework of  
 65 ggplot2, the tidyverse, and associated packages (Wick-  
 66 ham, 2011; Wickham et al., 2019), plotting functions re-  
 67 turn plot objects with default, but customizable, aesthet-  
 68 ics. Here, we present five vignettes demonstrating how  
 69 to use RevGadgets to summarize results for a variety of  
 70 phylogenetic analyses.

## 71 Phylogenies

72 Phylogenies are central to all analyses in RevBayes, so  
 73 accurate and information-rich visualizations of evolu-  
 74 tionary trees are critical. In this case study, we demon-  
 75 strate the tree-plotting functionality of RevGadgets, with  
 76 methods to visualize phylogenies and their associated  
 77 posterior probabilities, divergence-time estimates, and



**Figure 1: Plotting a time-calibrated phylogeny of extinct and extant taxa.** Top) RevGadgets code for reading in and plotting a time-calibrated phylogeny of extant and extinct bears. We use the theme function from ggplot2 to add the posterior-probability legend. Bottom) The maximum sampled-ancestor clade-credibility (MSACC) tree for the bears. Sampled ancestors are indicated by numbers along the branches (legend, top left). Bars represent the 95% credible interval of the age of the node, tip or sampled ancestor in millions of years (geological timescale, x-axis); the color of the bar corresponds to the posterior probability (legend, middle left) of that a clade exists, the posterior probability that a fossil is a sampled ancestor, or the posterior probability that a tip is not a sampled ancestor. (Data from Abella et al., 2012; Heath et al., 2014.)

branch-specific parameter estimates.

RevGadgets provides paired functions for (1) reading in and processing data, and (2) summarizing and visualizing results. For phylogenies, the function readTrees()

loads trees (either individual trees, or sets of trees) in either Newick or NEXUS (Maddison et al., 1997) formats, then processes associated branch or node annotations, and finally stores the tree(s) as treedata object(s) (as defined by treeio; Wang et al., 2020). Users can then visualize the treedata object using either plotTree() or plotFBDTree(), as we demonstrate below.

RevGadgets can plot both unrooted and rooted trees, and creates plots that are compatible with plotting options from ggtree. Additionally, RevGadgets provides extensive functionality for plotting trees with non-contemporaneous tips, such as those produced by total-evidence analyses under the fossilized birth-death [FBD] process (Heath et al., 2014; Zhang et al., 2016). The fossilized birth-death process (and the related serially-sampled birth-death process; Stadler, 2010) produces sampled ancestors (samples that are directly ancestral to another sampled taxon and thus are not represented as tips in the tree), and the ages of the samples are often subject to uncertainty (e.g., because of imperfect knowledge about the age of the strata from which the samples were collected). As a consequence, conventional tree plotting tools are unsuitable for plotting FBD trees. We demonstrate how to use RevGadgets to plot the results of an FBD analyses of living and extinct bears (Figure 1; data from Abella et al., 2012 and Heath et al., 2014). We include age bars colored by the posterior probability of the corresponding node, a geological time scale and labeled epochs from the package deptime (Gearty, 2021), and fossils estimated to be direct ancestors of other samples (i.e., sampled ancestors).

In addition to visualizing trees themselves, RevGadgets allows researchers to visualize branch-specific parameters, for example rates of evolution or diversification for each branch in the phylogeny. In Figure 2, we demonstrate how to use plotTree() to visualize the estimated optimal body size as it varies across the cetacean phylogeny under a relaxed Ornstein-Uhlenbeck process (Butler and King, 2004; Uyeda and Harmon, 2014; data from Steeman et al., 2009; Slater et al., 2010). Under this model, a quantitative character evolves towards an adaptive optimum that changes along the branches of the tree, and thus the optimum associated with each branch is a focal inference.

The plotTree() function can also visualize unrooted or circular phylogenies, and users may add text annotations to denote posterior probabilities or other quantities. Users can apply ggtree functions to modify the RevGadgets plot, e.g., to highlight certain clades with geom\_highlight() or to add phylopic (http://phylopic.org/) using geom\_phylopic(). Together, these functions provide user-friendly and customizable tree-plotting functionality for a variety of core research questions in evolutionary biology.

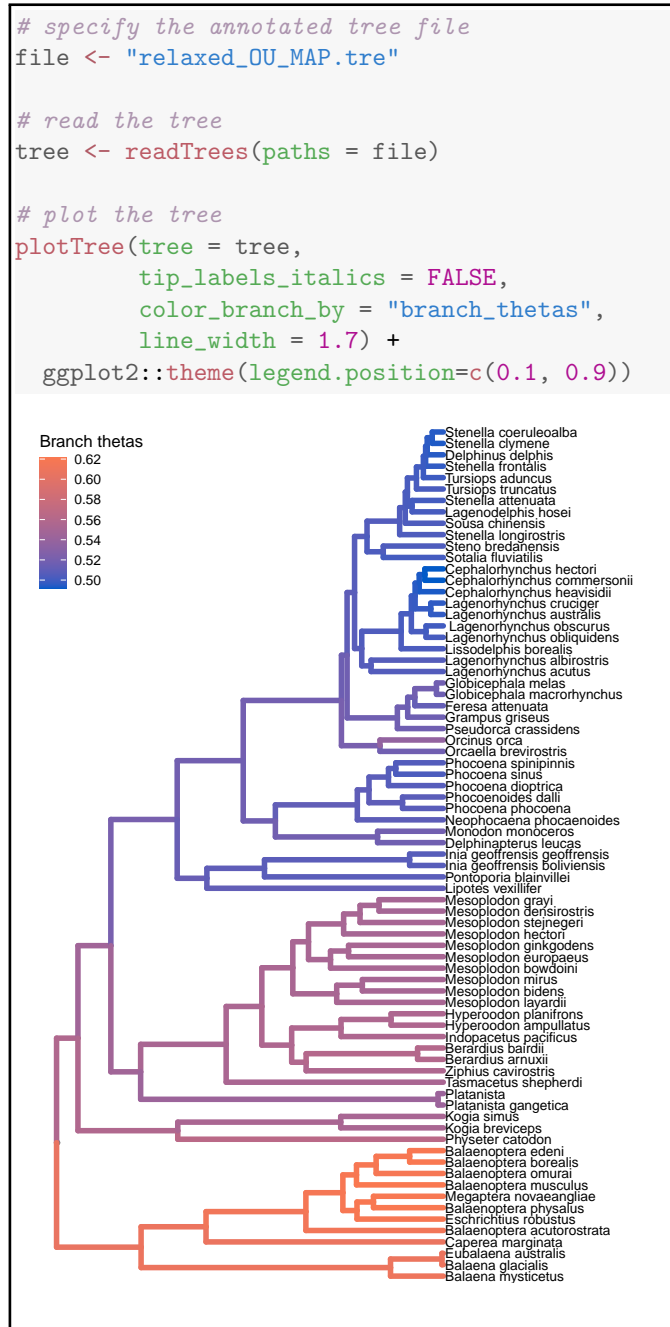


Figure 2: Plotting branch-specific parameter values across a phylogeny. Top) RevGadgets code for reading in and plotting the cetacean phylogeny that has been annotated with branch-specific adaptive optima ( $\theta$ ) inferred under a relaxed Ornstein-Uhlenbeck model. Bottom) The cetacean phylogeny with branches colored according to the posterior-mean estimate of the inferred branch-specific optimum body size,  $\theta$  (legend, tip left). (Phylogeny from Steeman et al., 2009; body size data in units of natural log-transformed meters from Slater et al., 2010.)

## Posterior Estimates of Numerical Parameters

RevGadgets provides several tools to visualize posterior distributions of numerical parameters. The output pro-

140 duced by most RevBayes analyses is a (typically tab-  
141 delimited) text file where rows correspond to samples  
142 from sequential iterations of an MCMC analysis, and  
143 columns correspond to parameters in the model. Most  
144 information of interest to researchers—*e.g.*, most prob-  
145 able parameter values (maximum *a posteriori*, or MAP,  
146 estimates), 95% credible intervals (CIs), or full poster-  
147 ior distributions—requires processing this raw MCMC  
148 output. Here, we demonstrate methods for processing  
149 and visualizing MCMC output for both quantitative and  
150 qualitative parameters.

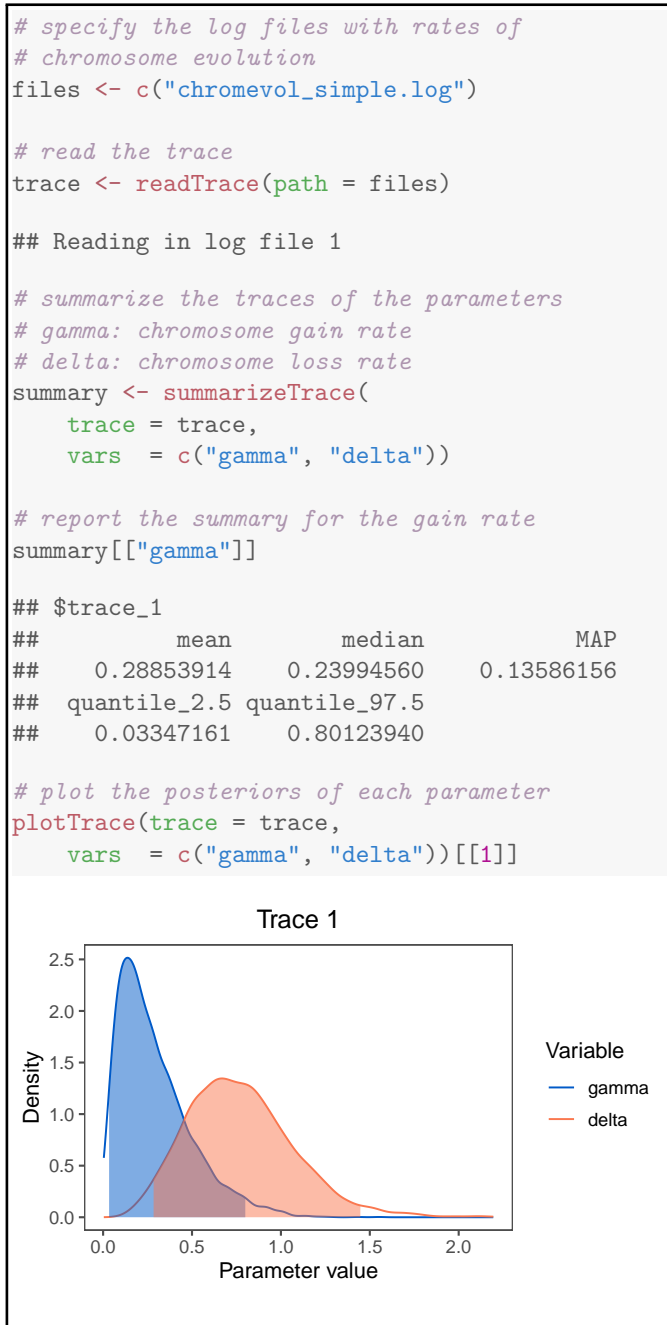
151 We illustrate the core functions for reading, summariz-  
152 ing and visualizing posterior distributions of specific pa-  
153 rameters with an example analysis of chromosome num-  
154 ber evolution (Figure 3; data from Freyman and Höhna,  
155 2018). We use `readTrace()` to read in parameters sam-  
156 pled during one or more MCMC analyses. We then use  
157 `summarizeTrace()` to calculate the posterior mean and  
158 95% credible interval for the focal parameters. Finally,  
159 we plot the marginal posterior distributions of the focal  
160 parameters using `plotTrace()`.

161 Plots of the posterior distributions of parameter val-  
162 ues are key to a thorough understanding of the results  
163 of any Bayesian analysis. These tools encourage users to  
164 explore their results thoroughly rather than relying on  
165 single summary statistics. These summaries and plots  
166 may also be useful as tools for science communication  
167 and education on statistical phylogenetics, as they can  
168 easily be used to demonstrate differences in parameter  
169 estimates that result from changes to basic phylogenetic  
170 models. Additionally, the output of `readTrace()` may  
171 be passed to R packages specializing in MCMC diagno-  
172 sis, *e.g.*, `convenience` (Fabreti and Höhna, 2021) or `coda`  
173 (Plummer et al., 2006). These functions are compatible  
174 with any delimited text file of MCMC samples, and can  
175 be used with the output of most Bayesian phylogenetic  
176 programs.

## 177 Ancestral-State Estimates

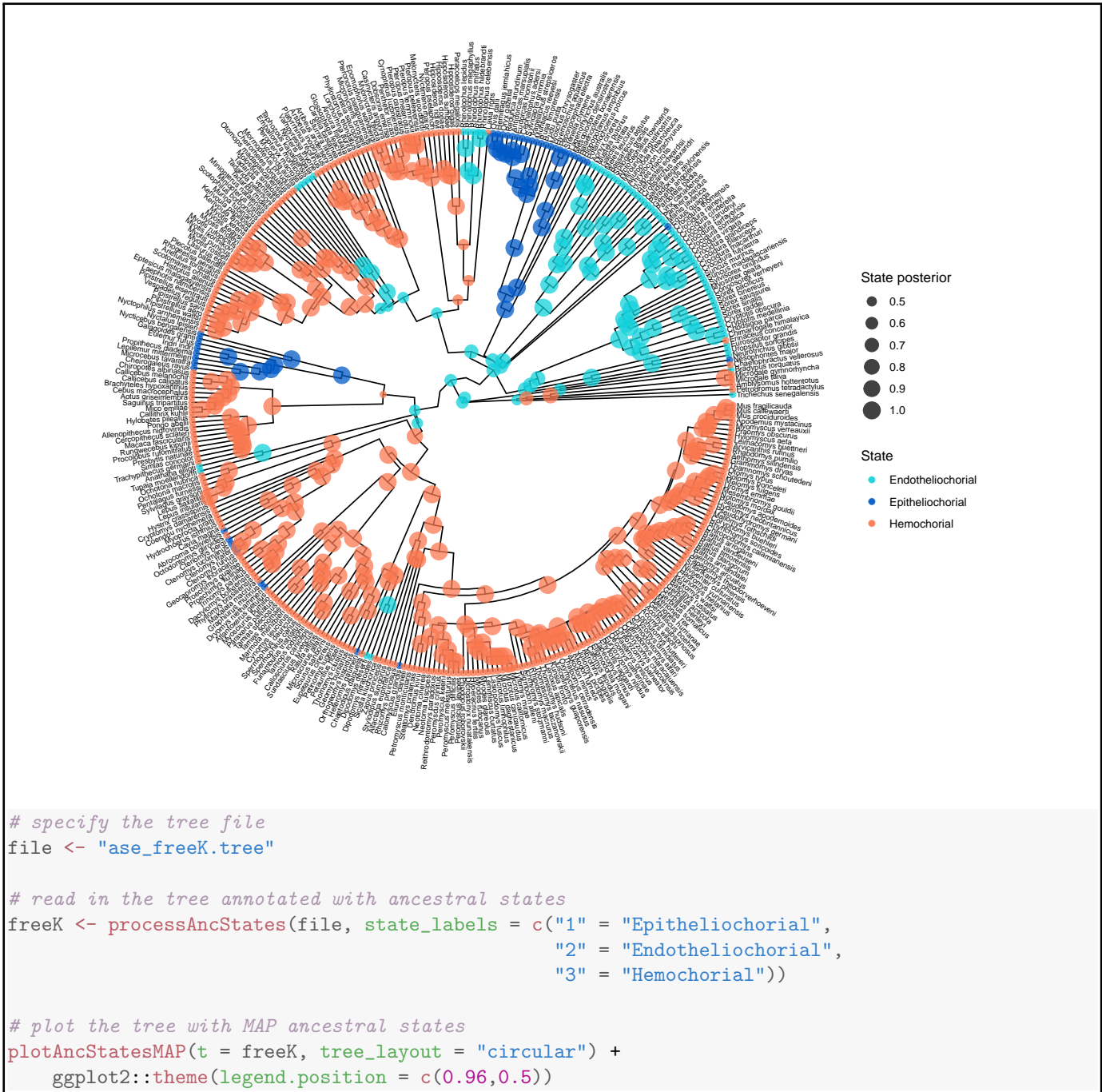
178 In addition to making inferences about the underlying  
179 process of evolution, researchers may be interested in  
180 studying how particular characters evolved across the  
181 branches of the phylogeny. Ancestral-state estimation is  
182 a method for inferring that history.

183 RevGadgets offers two different types of summaries  
184 for ancestral-state estimates: 1) maximum *a posteriori*  
185 (MAP) estimates, *i.e.*, the states with the highest poster-  
186 ior probability at each node, and; 2) pie charts that rep-  
187 resent each state in proportion to its probability at each  
188 node. Ancestral-state estimates may be represented as  
189 text annotations rather than colored symbols. Addition-  
190 ally, RevGadgets can summarize and visualize ancestral-  
191 state estimates at internal nodes and at the “shoulders”,



192 **Figure 3: Plotting posterior distributions of numerical parameter val-**  
193 **ues.** Top) RevGadgets code for reading in and plotting the posterior  
194 distributions of rates of chromosome evolution in *Aristolochia*. Bottom)  
195 Marginal posterior distributions of the two rate parameters. Shaded re-  
196 gions represent the 95% credible interval of each posterior distribution.  
197 (Data from Freyman and Höhna, 2018.)

198 *i.e.*, at the beginning of each branch. Plotting the states at  
199 internal nodes is appropriate for standard evolutionary  
200 models of anagenetic (within-lineage) change. However,  
201 models of evolution that include a cladogenetic compo-  
202 nent (*e.g.*, models of biogeographic or chromosome-  
203 number evolution; Ree and Smith, 2008; Goldberg and

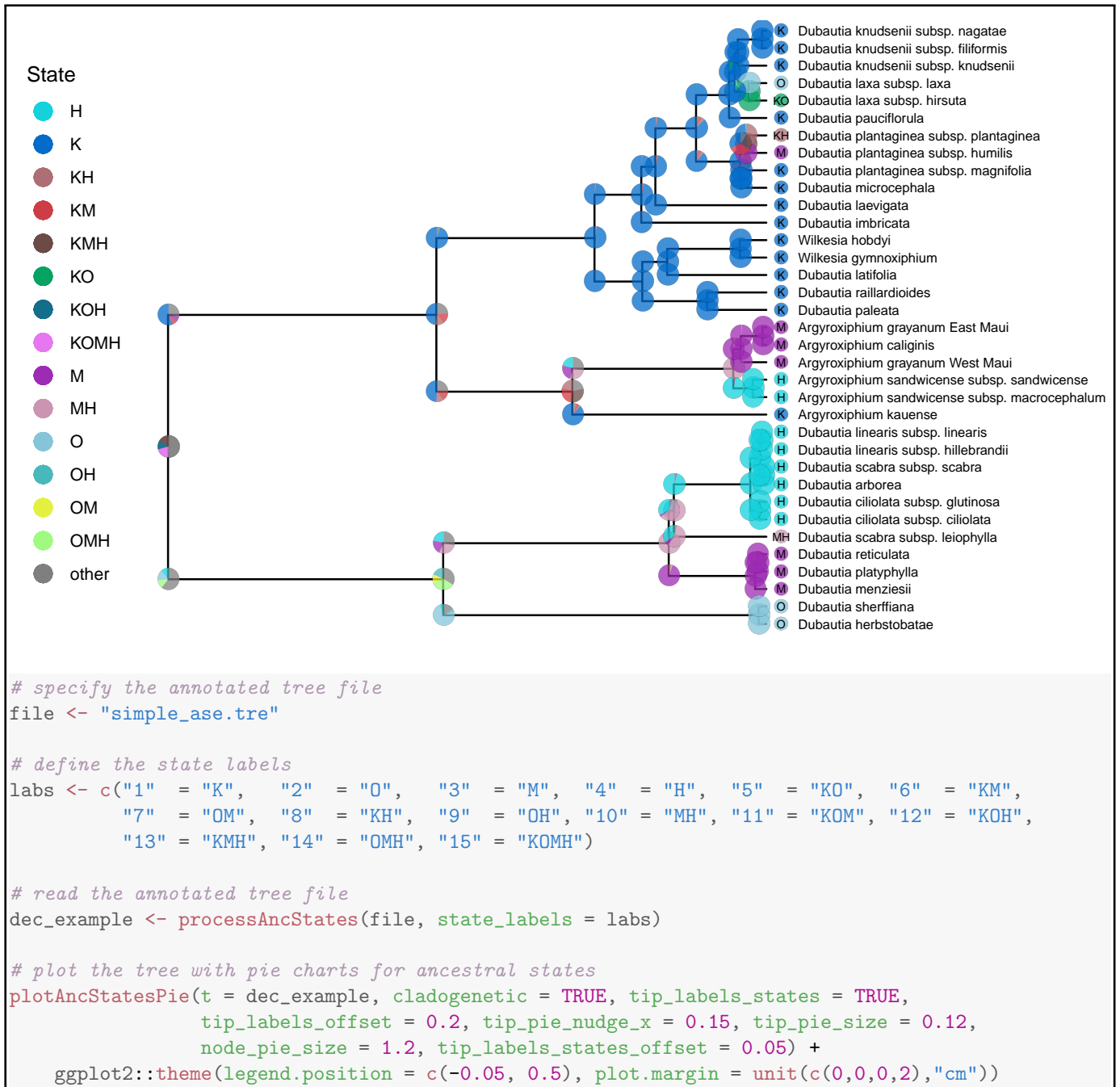


**Figure 4: Plotting maximum a posteriori (MAP) estimates of ancestral states on a circular phylogeny.** Top) MAP estimates of ancestral placental states across the phylogeny of mammals. Each node is colored by the MAP state (legend, bottom right); the size of each symbol is proportional to the posterior probability of the map state (legend, top right). Bottom) RevGadgets code for reading in and plotting the MAP estimates for ancestral placental states across the mammals phylogeny. (Data from Elliot and Crespi, 2006.)

198 Igić, 2012; Freyman and Höhna, 2018) also allow states  
 199 to change at speciation events. In this case, researchers  
 200 may also want to plot the shoulder states, which repre-  
 201 sent the ancestral-state estimates for each daughter lin-  
 202 eage immediately following the speciation event.

203 We demonstrate how to plot ancestral-state estimates

of placenta type across the mammal phylogeny un- 204  
 der an asymmetric model of character evolution (Fig- 205  
 ure 4; data from Elliot and Crespi, 2006). First, we use 206  
 processAncStates() to read in and parse the phylogeny 207  
 and ancestral-state estimates inferred using RevBayes. 208  
 Second, we use plotAncStatesMAP() to color each node 209



**Figure 5: Plotting posterior distributions of ancestral states under a cladogenetic model.** Top) The posterior estimates ancestral biogeographic states of the Hawaiian silverswords estimated under a DEC model. The size of each pie slice is proportional to the posterior probability of a given state (legend, top left) for a particular lineage. Pies at nodes represent the state of the ancestral lineage immediately before speciation; pies at “shoulders” represent the states of each daughter lineage immediately following the speciation event. Bottom) RevGadgets code for reading in and plotting the posterior estimates for ancestral geographic range across the phylogeny of Hawaiian silverswords. (Data from Landis et al., 2018.)

210 symbol according to the state with the highest posterior  
 211 probability, and make the radius of the symbol propor-  
 212 tional to that state’s posterior probability. Because of the  
 213 size of the phylogeny, we choose to plot the estimates on  
 214 a circular tree by changing the tree layout parameter.

215 Next, we demonstrate plotting estimates of ancestral

216 ranges of the Hawaiian silversword alliance that were  
 217 generated by a Dispersal-Extinction-Cladogenesis (DEC)  
 218 model (Figure 5; data from Landis et al., 2018). Since  
 219 the DEC model features a cladogenetic component, we  
 220 include shoulder-state estimates. Because of the large  
 221 number of states in this analysis (15 possible ranges

222 and one “other” category), more pre-processing is necessary. As before, we pass the appropriate state names  
223 to `processAncStates()`; however, in this case we plot  
224 pie charts representing the probability of each state using  
225 `plotAncStatesPie()`, and plot states at shoulders using  
226 `cladogenetic = TRUE`.

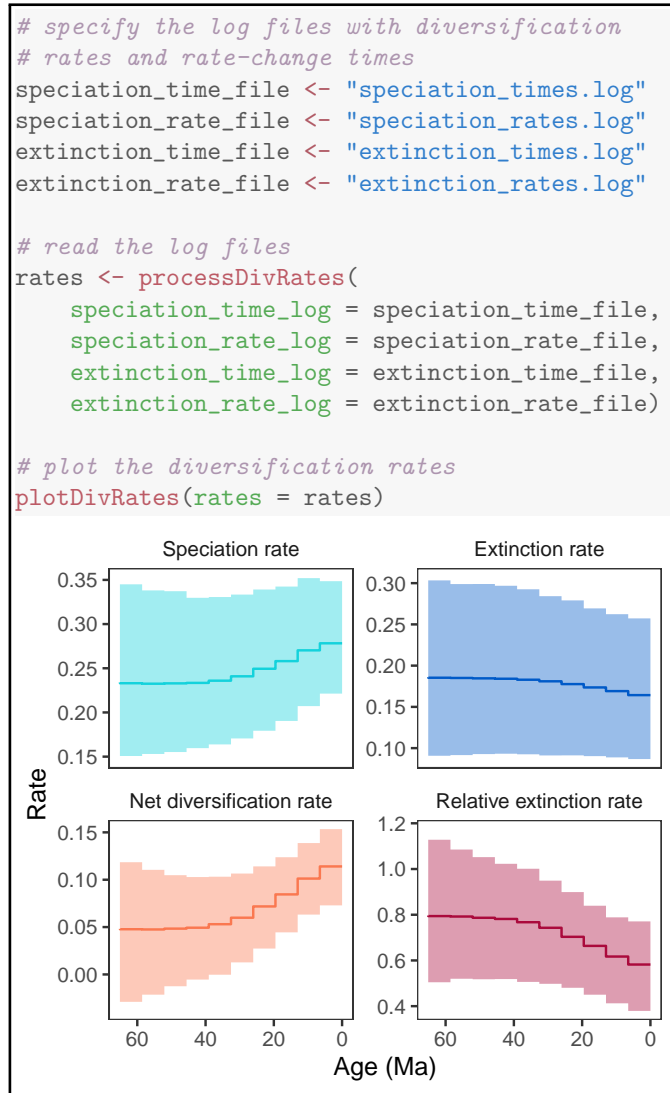
227  
228 Beyond the above examples, these versatile plotting  
229 tools can visualize any discrete ancestral-state estimates  
230 reconstructed by RevBayes, including the results of  
231 chromosome count estimations (Freyman and Höhna,  
232 2018) and discrete state-dependent speciation and ex-  
233 tinction (SSE) models (Freyman and Höhna, 2019; Zenil-  
234 Ferguson et al., 2019).

## 235 Diversification Rates

236 The processes of speciation and extinction (*i.e.*, lineage  
237 diversification) is of great interest to evolutionary bi-  
238 ologists (Morlon, 2014). Rates of speciation and ex-  
239 tinction may be modeled as constant over time and  
240 among branches (as in a constant-rate birth-death pro-  
241 cess; Kendall et al., 1948; Nee et al., 1994), or allowed  
242 to vary over time (Stadler, 2011; May et al., 2016), across  
243 branches of a phylogeny (Rabosky, 2014; Höhna et al.,  
244 2019), or based on the character states of the evolving  
245 lineages (Maddison et al., 2007; Freyman and Höhna,  
246 2019). For example, rates that vary across branches of the  
247 phylogeny can be visualized using `plotTree()` to color  
248 the branches by their inferred rate. State-dependent  
249 diversification models provide estimates of the specia-  
250 tion and extinction rates associated with each charac-  
251 ter state, and may also be used to estimate ancestral  
252 states. `plotTrace()` or specific processing and plot-  
253 ting functions for diversification rates—`processSSE()`,  
254 `plotMuSSE`, and `plotHiSSE`—may be used to visu-  
255 alize the estimated rates. `plotAncStatesMAP()` or  
256 `plotAncStatesPie()` may be used to visualize the  
257 ancestral-state estimates.

258 We demonstrate how to plot the results of a  
259 time-varying model—the episodic birth-death process  
260 (Stadler, 2011; Höhna, 2015)—applied to primate phy-  
261 logeny (Figure 6; Springer et al., 2012). The episodic  
262 birth-death analysis in RevBayes produces separate trace  
263 files each type of rate. We read these output files using  
264 `processDivRates()` and plot the resulting parameter es-  
265 timates over time using `plotDivRates()`.

266 Together with the aforementioned functions  
267 for plotting diversification parameter estimates,  
268 `plotDivRates()` allows users to visualize the out-  
269 puts of nearly all diversification analyses available in  
270 RevBayes. Stochastic character mapping of diversifi-  
271 cation estimates, in which the timing and location of  
272 diversification rate shifts are painted along the branches  
273 of the tree, will be added in the future (Freyman and



**Figure 6: Plotting posterior distributions of diversification rates over time.** Top) RevGadgets code for reading in and plotting the posterior estimates of diversification rates over time inferred from the primate phylogeny. Bottom) Posterior distributions of speciation and extinction rates over time, as well as the net diversification rate (speciation minus extinction) and the relative extinction rate (extinction divided by speciation). Dark lines correspond to the posterior-mean estimate of each parameter for each time interval, and shaded regions correspond to the 95% credible interval. (Data from Springer et al., 2012.)

Höhna, 2019; Höhna et al., 2019).

## 275 Model Adequacy

276 In addition to visualizing the results of phylogenetic  
277 inferences with a specific model, RevGadgets provides  
278 tools for exploring the adequacy of the model (*i.e.*,  
279 whether the model provides an adequate description  
280 of the data-generating process; Bollback, 2002; Gelman  
281 et al., 2013; Brown, 2014; Höhna et al., 2018). Posterior-

282 predictive analysis tests whether a fitted model simu-  
283 lates (predicts) data that are similar to the observed data.  
284 This process is distinct from model testing, in which one  
285 model is chosen from a set of possible models, as the best  
286 model of the set may still provide an inadequate descrip-  
287 tion of the underlying process.

288 First, users analyze their data with the model of inter-  
289 est and then use the inferred posterior distribution to  
290 simulate a number of new data sets. The user then se-  
291 lects test statistics that describe important features of the  
292 data (e.g., the number of invariant sites in a nucleotide  
293 alignment) and calculates these statistics for both the ob-  
294 served data and the simulated data. If the statistic from  
295 the empirical data is reasonably included within the dis-  
296 tribution of statistics from simulated datasets (posterior-  
297 predictive  $p$ -value  $> 0.05$ ), the model is considered an  
298 adequate description of the process that produced the  
299 tested data feature.

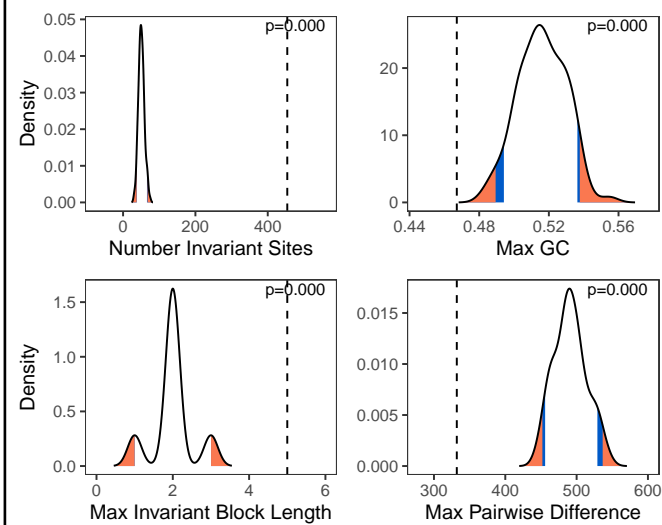
300 Here, we demonstrate the workflow for a posterior-  
301 predictive analysis to test model adequacy of the Jukes-  
302 Cantor model for nucleotide sequence evolution (Jukes  
303 et al., 1969) in a single gene across a sample of 23 pri-  
304 mates (Figure 7; data from Springer et al., 2012). First,  
305 we perform an analysis in RevBayes under a Jukes-  
306 Cantor model of nucleotide sequence data. Second,  
307 we use RevBayes to simulate datasets under the pos-  
308 terior distributions estimated in the first step. Third,  
309 we use RevBayes to calculate statistics from the simu-  
310 lated and empirical datasets. These statistics should  
311 describe aspects of the data that we hope capture a  
312 meaningful component of model performance. Finally,  
313 we use RevGadgets to plot those statistics and compute  
314 posterior-predictive  $p$ -values.

315 Despite being computationally inexpensive compared  
316 to Bayesian model comparison methods (i.e., Bayes fac-  
317 tor calculation), posterior-predictive approaches remain  
318 relatively uncommon in empirical phylogenetic studies.  
319 As genome-scale datasets and increasingly complex sta-  
320 tistical methods become more accessible to researchers,  
321 posterior-predictive simulation will be critical to testing  
322 how well our models describe the underlying generative  
323 processes. This component of RevGadgets functionality  
324 and the associated clear workflows for performing and  
325 interpreting posterior-predictive tests will hopefully in-  
326 crease the adoption of this important tool.

## 327 Conclusions

328 RevBayes is a flexible platform for performing Bayesian  
329 phylogenetic evolutionary inferences. Because of the  
330 almost endless possibilities for building unique combi-  
331 nations of models in RevBayes, these analyses are of-  
332 ten challenging to visualize using standard plotting soft-  
333 ware. We have developed an R package, RevGadgets,

```
# specify the simulated statistics file  
sim <- "simulated_data_pps.csv"  
  
# specify the empirical statistics file  
emp <- "empirical_data_pps.csv"  
  
# read the statistics files  
stats <- processPostPredStats(path_sim = sim,  
                              path_emp = emp)  
  
# create the posterior-predictive plots  
plots <- plotPostPredStats(data = stats)  
  
# plot some of the statistics  
plots[c(1,3,5,7)]
```



**Figure 7: Plotting simulated posterior-predictive distributions to assess model adequacy.** Top) RevGadgets code for reading in and plotting the distributions of summary statistics generated using posterior-predictive simulation posterior. Bottom) Posterior-predictive distributions (black curves) of four statistics simulated under the Jukes-Cantor model fit to primate *cytb*, compared to the same statistics computed on the observed data (dashed vertical lines). The posterior-predictive  $p$ -value (upper right of each panel) is the fraction of simulated statistics that are as or more extreme than the observed statistic. If the observed statistic falls in or beyond the orange region, we deem the model as inadequate at the 5% significance level; if the observed statistic falls in the blue region, the model is marginally adequate at the 10% significance level. In this case, the Jukes-Cantor model provides an inadequate description of the true generating process according to every summary statistic. (Data from Springer et al., 2012.)

to produce publication-quality visualizations of phylogenetic analyses performed in RevBayes. The case studies described above illustrate some of the core functionality available in RevGadgets and demonstrate how to produce plots of the most commonly-performed RevBayes analyses. RevBayes is open source software that is actively maintained and developed. Likewise,



341 RevGadgets is also open source and will continue to pro-  
342 vide new plotting tools to meet new visualization chal-  
343 lenges as they arise. RevGadgets and any future up-  
344 dates will be available on GitHub at <https://github.com/cmt2/RevGadgets>. Additionally, we provide thor-  
345 ough documentation for all functionality in the package  
346 and maintain numerous tutorials demonstrating how to  
347 use RevGadgets on the RevBayes website at [https://  
348 revbayes.github.io/tutorials/](https://revbayes.github.io/tutorials/). Together, the modu-  
349 lar modeling tools from RevBayes and the visualization  
350 gadgets in RevGadgets will help researchers make sense  
351 of and communicate the results of a diverse array of so-  
352 phisticated phylogenetic analyses.  
353

## 354 Authors Contributions

355 CMT and MRM designed the R package. All authors  
356 contributed code and examples. CMT and MRM drafted  
357 the manuscript. All authors revised and approved the  
358 final version of the manuscript.

## 359 Acknowledgements

360 We would like to acknowledge Carl J. Rothfels, Benjamin  
361 K. Blackman, David D. Ackerly, and Chelsea D. Specht  
362 for feedback on initial stages of the manuscript. Ixchel  
363 González Ramírez, Jenna T. B. Ekwealor, Isaac Lichter  
364 Marck, and members of the Rothfels Lab at UC Berkeley  
365 provided valuable feedback on usability and legibility of  
366 figures and code.

367 This research was supported by the Deutsche  
368 Forschungsgemeinschaft (DFG) Emmy Noether-  
369 Program HO 6201/1-1 awarded to SH.

## 370 Data Availability

371 RevGadgets and all example datasets are freely available  
372 on GitHub at <https://github.com/cmt2/RevGadgets>.

## 373 References

374 Abella, J., Alba, D. M., Robles, J. M., Valenciano, A., Rotgers, C., Car-  
375 mona, R., Montoya, P., and Morales, J. (2012). *Kretzoiarctos gen. nov.*,  
376 the oldest member of the giant panda clade. *PLoS One*, 7(11).

377 Baele, G., Suchard, M. A., Rambaut, A., and Lemey, P. (2017). Emerging  
378 concepts of data integration in pathogen phylodynamics. *Systematic  
379 Biology*, 66(1):e47–e65.

380 Bollback, J. P. (2002). Bayesian model adequacy and choice in phyloge-  
381 netics. *Molecular Biology and Evolution*, 19(7):1171–1180.

382 Brown, J. M. (2014). Predictive approaches to assessing the fit of evo-  
383 lutionary models. *Systematic biology*, 63(3):289–292.

384 Butler, M. A. and King, A. A. (2004). Phylogenetic comparative anal-  
385 ysis: a modeling approach for adaptive evolution. *The American  
386 Naturalist*, 164(6):683–695.

Elliot, M. G. and Crespi, B. J. (2006). Placental invasiveness mediates 387  
the evolution of hybrid inviability in mammals. *The American Natu- 388  
ralist*, 168(1):114–120. 389

Fabreti, L. G. and Höhna, S. (2021). Convergence assessment for 390  
bayesian phylogenetic analysis using mcmc simulation. 391

Faith, D. P. (1992). Conservation evaluation and phylogenetic diversity. 392  
*Biological Conservation*, 61(1):1–10. 393

Felsenstein, J. (1985). Phylogenies and the comparative method. *The 394  
American Naturalist*, 125(1):1–15. 395

Freyman, W. A. and Höhna, S. (2018). Cladogenetic and anagenetic 396  
models of chromosome number evolution: a Bayesian model aver- 397  
aging approach. *Systematic biology*, 67(2):195–215. 398

Freyman, W. A. and Höhna, S. (2019). Stochastic character mapping of 399  
state-dependent diversification reveals the tempo of evolutionary 400  
decline in self-compatible Onagraceae lineages. *Systematic Biology*, 401  
68(3):505–519. 402

Gearty, W. (2021). deeptime: Plotting Tools for Anyone Working in 403  
Deep Time. R package version 0.0.5.3. 404

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and 405  
Rubin, D. B. (2013). *Bayesian data analysis*. CRC press. 406

Goldberg, E. E. and Igić, B. (2012). Tempo and mode in plant breeding 407  
system evolution. *Evolution: International Journal of Organic Evolu- 408  
tion*, 66(12):3701–3709. 409

Harvey, P. H. and Pagel, M. D. (1991). *The Comparative Method in Evolu- 410  
tionary Biology*, volume 239. Oxford University Press. 411

Heath, T. A., Huelsenbeck, J. P., and Stadler, T. (2014). The fossilized 412  
birth–death process for coherent calibration of divergence-time esti- 413  
mates. *Proceedings of the National Academy of Sciences*, 111(29):E2957– 414  
E2966. 415

Höhna, S. (2015). The time-dependent reconstructed evolutionary pro- 416  
cess with a key-role for mass-extinction events. *Journal of theoretical 417  
biology*, 380:321–331. 418

Höhna, S., Coghill, L. M., Mount, G. G., Thomson, R. C., and Brown, 419  
J. M. (2018). P3: Phylogenetic posterior prediction in RevBayes. 420  
*Molecular biology and evolution*, 35(4):1028–1034. 421

Höhna, S., Freyman, W. A., Nolen, Z., Huelsenbeck, J., May, M. R., and 422  
Moore, B. R. (2019). A Bayesian approach for estimating branch- 423  
specific speciation and extinction rates. *bioRxiv*, 555805. 424

Höhna, S., Heath, T. A., Boussau, B., Landis, M. J., Ronquist, F., and 425  
Huelsenbeck, J. P. (2014). Probabilistic graphical model representa- 426  
tion in phylogenetics. *Systematic biology*, 63(5):753–771. 427

Höhna, S., Landis, M. J., Heath, T. A., Boussau, B., Lartillot, N., Moore, 428  
B. R., Huelsenbeck, J. P., and Ronquist, F. (2016). RevBayes: Bayesian 429  
phylogenetic inference using graphical models and an interactive 430  
model-specification language. *Systematic Biology*, 65(4):726–736. 431

Jukes, T. H., Cantor, C. R., et al. (1969). Evolution of protein molecules. 432  
*Mammalian protein metabolism*, 3(21):132. 433

Kendall, D. G. et al. (1948). On the generalized “birth-and-death” pro- 434  
cess. *The annals of mathematical statistics*, 19(1):1–15. 435

Kerman, J., Gelman, A., Zheng, T., and Ding, Y. (2008). Visualization 436  
in bayesian data analysis. In *Handbook of Data Visualization*, pages 437  
709–724. Springer. 438

- 439 Landis, M. J., Freyman, W. A., and Baldwin, B. G. (2018). Retracing  
440 the Hawaiian silversword radiation despite phylogenetic, biogeographic,  
441 and paleogeographic uncertainty. *Evolution*, 72(11):2343–  
442 2359.
- 443 Maddison, D. R., Swofford, D. L., and Maddison, W. P. (1997). NEXUS:  
444 an extensible file format for systematic information. *Systematic biology*,  
445 46(4):590–621.
- 446 Maddison, W. P., Midford, P. E., and Otto, S. P. (2007). Estimating a  
447 binary character’s effect on speciation and extinction. *Systematic biology*,  
448 56(5):701–710.
- 449 May, M. R., Höhna, S., and Moore, B. R. (2016). A Bayesian approach  
450 for detecting the impact of mass-extinction events on molecular  
451 phylogenies when rates of lineage diversification may vary. *Methods  
452 in Ecology and Evolution*, 7(8):947–959.
- 453 Morlon, H. (2014). Phylogenetic approaches for studying diversification.  
454 *Ecology letters*, 17(4):508–525.
- 455 Nee, S., May, R. M., and Harvey, P. H. (1994). The reconstructed evolutionary  
456 process. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*,  
457 344(1309):305–311.
- 458 Nei, M. (1987). *Molecular Evolutionary Genetics*. Columbia University  
459 Press.
- 460 Paradis, E. and Schliep, K. (2019). ape 5.0: an environment for modern  
461 phylogenetics and evolutionary analyses in R. *Bioinformatics*,  
462 35:526–528.
- 463 Plummer, M., Best, N., Cowles, K., and Vines, K. (2006). CODA: convergence  
464 diagnosis and output analysis for MCMC. *R news*, 6(1):7–  
465 11.
- 466 R Core Team (2020). *R: A Language and Environment for Statistical Computing*.  
467 R Foundation for Statistical Computing, Vienna, Austria.
- 468 Rabosky, D. L. (2014). Automatic detection of key innovations, rate  
469 shifts, and diversity-dependence on phylogenetic trees. *PLoS one*,  
470 9(2).
- 471 Rambaut, A. (2014). FigTree 1.4. 2 software. *Institute of Evolutionary  
472 Biology, Univ. Edinburgh*.
- 473 Rambaut, A., Drummond, A. J., Xie, D., Baele, G., and Suchard, M. A.  
474 (2018). Posterior summarization in Bayesian phylogenetics using  
475 Tracer 1.7. *Systematic biology*, 67(5):901–904.
- 476 Ree, R. H. and Smith, S. A. (2008). Maximum likelihood inference of  
477 geographic range evolution by dispersal, local extinction, and cladogenesis.  
478 *Systematic biology*, 57(1):4–14.
- 479 Revell, L. J. (2012). phytools: an R package for phylogenetic comparative  
480 biology (and other things). *Methods in ecology and evolution*,  
481 3(2):217–223.
- 482 Ronquist, F. and Sanmartín, I. (2011). Phylogenetic methods in biogeography.  
483 *Annual Review of Ecology, Evolution, and Systematics*, 42.
- 484 Slater, G. J., Price, S. A., Santini, F., and Alfaro, M. E. (2010). Diversity  
485 versus disparity and the radiation of modern cetaceans. *Proceedings of the Royal Society B: Biological Sciences*,  
486 277(1697):3097–3104.
- 487 Springer, M. S., Meredith, R. W., Gatesy, J., Emerling, C. A., Park, J., Rabosky,  
488 D. L., Stadler, T., Steiner, C., Ryder, O. A., Janečka, J. E., et al.  
489 (2012). Macroevolutionary dynamics and historical biogeography of primate  
490 diversification inferred from a species supermatrix. *PLoS one*,  
491 7(11).
- 492 Stadler, T. (2010). Sampling-through-time in birth–death trees. *Journal  
493 of Theoretical Biology*, 267(3):396–404.
- Stadler, T. (2011). Mammalian phylogeny reveals recent diversification  
494 rate shifts. *Proceedings of the National Academy of Sciences*,  
495 108(15):6187–6192.  
496
- Steeman, M. E., Hebsgaard, M. B., Fordyce, R. E., Ho, S. Y., Rabosky,  
497 D. L., Nielsen, R., Rahbek, C., Glenner, H., Sørensen, M. V., and  
498 Willerslev, E. (2009). Radiation of extant cetaceans driven by restructuring  
499 of the oceans. *Systematic Biology*, 58(6):573–585.  
500
- Tufte, E. (2001). The visual display of quantitative information. 501
- Uyeda, J. C. and Harmon, L. J. (2014). A novel Bayesian method for inferring  
502 and interpreting the dynamics of adaptive landscapes from phylogenetic  
503 comparative data. *Systematic biology*, 63(6):902–918.  
504
- Vaughan, T. G. (2017). IcyTree: rapid browser-based visualization for  
505 phylogenetic trees and networks. *Bioinformatics*, 33(15):2392–2394.  
506
- Volz, E. M., Koelle, K., and Bedford, T. (2013). Viral phylodynamics.  
507 *PLoS Computational Biology*, 9(3):e1002947.  
508
- Wang, L.-G., Lam, T. T.-Y., Xu, S., Dai, Z., Zhou, L., Feng, T., Guo, P.,  
509 Dunn, C. W., Jones, B. R., Bradley, T., et al. (2020). treeio: an R package  
510 for phylogenetic tree input and output with richly annotated  
511 and associated data. *Molecular biology and evolution*, 37(2):599–603.  
512
- Wickham, H. (2011). ggplot2. *Wiley Interdisciplinary Reviews: Computational  
513 Statistics*, 3(2):180–185.  
514
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.,  
515 François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., et al.  
516 (2019). Welcome to the Tidyverse. *Journal of Open Source Software*,  
517 4(43):1686.  
518
- Yang, Z. (2014). *Molecular Evolution: A Statistical Approach*. Oxford  
519 University Press.  
520
- Yu, G., Smith, D. K., Zhu, H., Guan, Y., and Lam, T. T.-Y. (2017). ggtree:  
521 an R package for visualization and annotation of phylogenetic trees  
522 with their covariates and other associated data. *Methods in Ecology  
523 and Evolution*, 8(1):28–36.  
524
- Zenil-Ferguson, R., Burleigh, J. G., Freyman, W. A., Igić, B., Mayrose, I.,  
525 and Goldberg, E. E. (2019). Interaction among ploidy, breeding system  
526 and lineage diversification. *New Phytologist*, 224(3):1252–1265.  
527
- Zhang, C., Stadler, T., Klopstein, S., Heath, T. A., and Ronquist, F.  
528 (2016). Total-evidence dating under the fossilized birth-death process.  
529 *Systematic Biology*, 65(2):228–249.  
530