

1 **sensiPhy: an R-package for sensitivity analysis in phylogenetic comparative methods**

2 Gustavo Brant Paterno^{1*}; Caterina Penone^{2*}; Gijbert D.A. Werner^{3,4*}

3

4 ¹Departamento de Ecologia, Universidade Federal do Rio Grande do Norte, Natal, Brazil.

5 ²Institute of Plant Sciences, University of Bern, CH

6 ³Department of Zoology, University of Oxford, UK

7 ⁴Balliol College, University of Oxford, UK

8

9 corresponding author: Gustavo B. Paterno (paternogbc@gmail.com)

10 * These authors contributed equally to this work

11 **Summary**

12 **1.** Biological conclusions drawn from phylogenetic comparative methods can be sensitive to
13 uncertainty in species sampling, phylogeny and data. To be confident about our conclusions, we
14 need to quantify their robustness to such uncertainty.

15 **2.** We present sensiPhy, an R-package to easily and rapidly perform sensitivity analysis for
16 phylogenetic comparative methods. sensiPhy allows researchers to evaluate the sampling effort,
17 detect influential species and clades, assess phylogenetic uncertainty and quantify the effects of
18 intraspecific variation, for phylogenetic regression and for metrics of phylogenetic signal,
19 diversification and trait evolution.

20 **3.** Uniquely, sensiPhy allows users to simultaneously quantify the effects of different types of
21 uncertainty and potential interactions among them.

22 **4.** Using real data, we show how conclusions from comparative methods can be affected by
23 uncertainty and how sensiPhy can help determine if a conclusion is robust.

24 **5.** By providing a single, intuitive and user-friendly resource that can evaluate various sources of
25 uncertainty, sensiPhy aims to encourage researchers, and particularly less experienced users, to
26 incorporate sensitivity analyses in their phylogenetic comparative analyses.

27

28 **Keywords:** PGLS, Phylogenetic Regression; Robustness; Diversification; Trait evolution; Bias

29

30 **Introduction**

31 Over the last few decades, phylogenetic comparative methods have become a central
32 approach in ecology and evolutionary biology, boosted by the expansion of comparative methods
33 available in R (Paradis, 2012; Garamszegi, 2014). Like all statistical models, phylogenetic
34 comparative methods are subject to several types of uncertainty which can affect conclusions we
35 draw from these analyses (Donoghue & Ackerly, 1996; Huelsenbeck *et al.* 2000; Felsenstein,
36 2008). Yet, the sensitivity of (biological) conclusions to uncertainty is seldom considered (Cooper
37 *et al.* 2016). This can cause researchers to overestimate the reliability of their findings, for instance
38 by estimating too narrow confidence intervals or by providing biased parameter estimates (Rangel
39 *et al.* 2015; Silvestro, 2015).

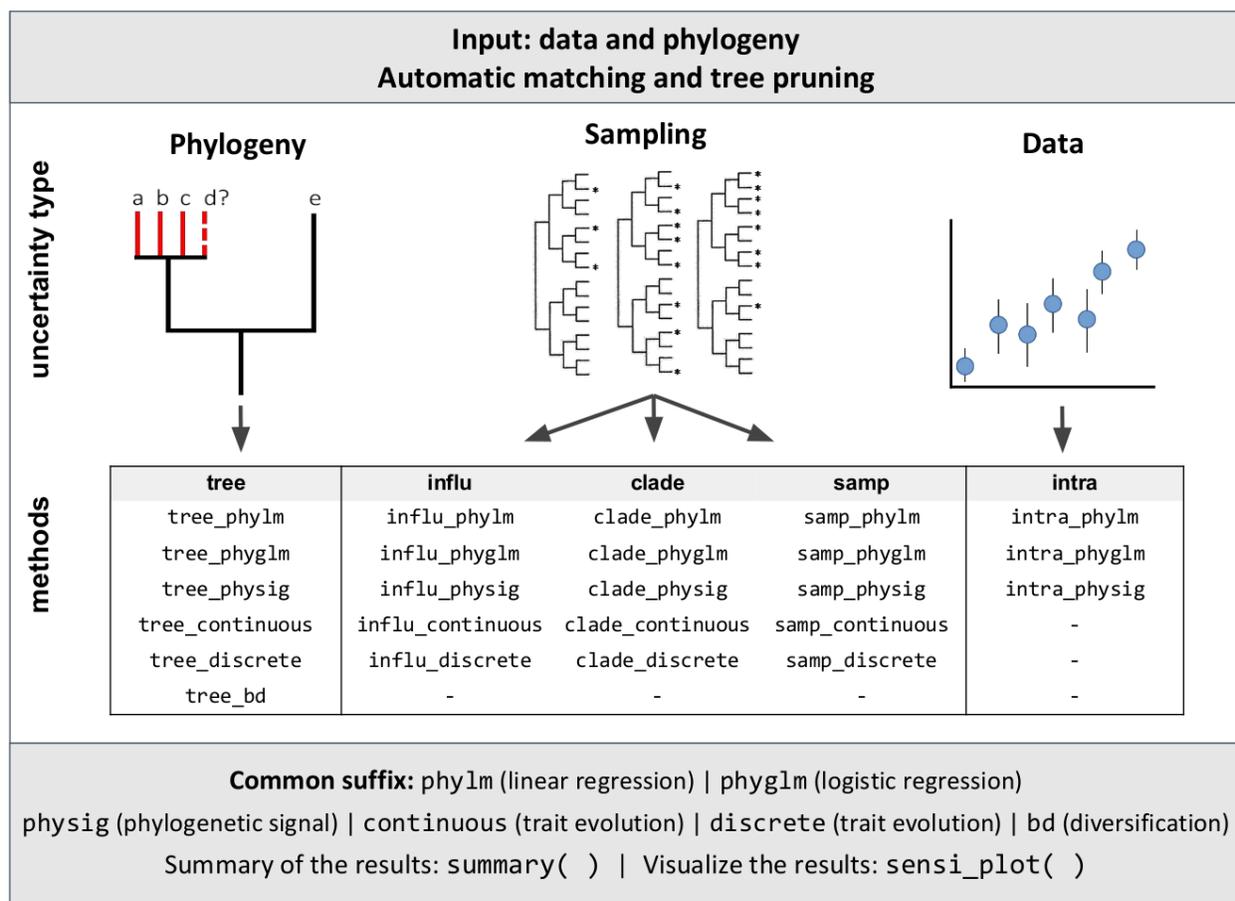
40 Three main sources of uncertainty can affect comparative methods (Fig. 1). (i) Species
41 sampling uncertainty encompasses uncertainty in parameter estimates resulting from (arbitrary)
42 variation in the species set included. (ii) Phylogenetic uncertainty encompasses uncertainty in
43 phylogenies used in comparative analyses. (iii) Data uncertainty includes both within-species
44 variation in trait values as well as measurement error that might occur when determining trait
45 values. Sensitivity analysis is a powerful approach to evaluate if conclusions are influenced by
46 these uncertainties in comparative biology (Donoghue and Ackerly, 1996; Cooper *et al.*, 2016;
47 Cornwell & Nakagawa, 2017). Here, we present sensiPhy, an R-package to perform sensitivity
48 analysis for the most frequently used phylogenetic comparative methods. Our main goal is to make
49 it easier for less-experienced users to implement the best practices when running comparative

50 analyses. To our knowledge, this is the first effort to combine in a single resource functions to
51 account for three types of uncertainty in commonly used comparative methods.

52

53 **The sensiPhy package**

54 SensiPhy is written in the R-language (R Core Team 2017) and is available on the CRAN
55 repository. The package provides an umbrella of statistical and graphical methods to estimate and
56 report sensitivity to uncertainty in phylogenetic comparative analysis (PGLS, phylogenetic signal,
57 diversification and trait evolution). We leverage methods implemented in the R-packages *phylolm*,
58 *phytools* and *geiger* (Ho & Ané 2014; Revell 2012; Harmon et al. 2008) and implement functions
59 to perform sensitivity analysis for phylogenetic generalized least squares models (PGLS; both
60 using linear and logistic regression models), for estimates of phylogenetic signal in trait data
61 (Blomberg et al. 2003, Pagel 1999), for macroevolutionary models (both continuous and discrete,
62 binary, traits) and estimates of diversification rates (Magallón & Sanderson, 2001; Harmon et al.
63 2008). For each type of sensitivity analysis, a specific set of diagnostics graphics and summary
64 statistics are provided (Fig. 1). In all PGLS functions, the evolutionary model to use can be
65 specified (e.g. Brownian Motion and Ornstein-Uhlenbeck; Ho & Ané 2014), allowing the user to
66 analyse the fit of different models and select the most appropriate one (Cornwell & Nakagawa
67 2014, Garamszegi 2014; Pennell et al. 2015). Scientists can use sensiPhy to analyse results
68 originally obtained from other software (e.g. PGLS with *caper* or *gls*) when available analysis use
69 the same macroevolutionary models implemented in *phylolm*, *phytools* and *geiger* (e.g. Brownian
70 Motion, OU, lambda; see package vignette for examples and details).



71

72 Fig1: Overview of the main functions in sensiPhy organized by source of uncertainty. sensiPhy contains functions to
 73 quantify the effects of the three types of uncertainty and of interactions among them: phylogenetic uncertainty (**tree**),
 74 uncertainty arising from species sampling (**influ**, **clade** and **samp**) and uncertainty in the underlying trait data
 75 (**intra**).

76

77 Sources of uncertainty

78 We briefly highlight the three main sources of uncertainty, indicating how they can affect
 79 conclusions, and then provide two examples on how researchers can use sensiPhy. A full tutorial,

80 highlighting examples for all sources of uncertainty and implemented functions, can be found in
81 the package vignette and on Github (<https://github.com/paternogbc/sensiPhy/wiki>).

82

83 *Species sampling uncertainty*

84 Some species, or clades of species, are particularly important drivers of parameter
85 estimates. However, often the set of species sampled in a comparative analysis is determined by
86 considerations that are arbitrary from an evolutionary perspective, like presence in a trait database
87 or easy access in the field. Also, conclusions can be sensitive to the number of species being
88 studied, or the sampling effort. Moreover, particular species or clades can represent influential
89 cases and can drive key results because they show a pattern that is different in strength or direction
90 than the general pattern. Since in all of these cases, the source of uncertainty is driven by the set
91 of species considered, we group all these issues under the name of species sampling uncertainty.

92 The `samp` functions (`samp_phylm`, `samp_phyglm`, `samp_physig`,
93 `samp_continuous` and `samp_discrete`; Fig. 1) uses a jackknifing method to test if
94 models are robust to variation in the set of species and sample size (Efron 1982; Werner et al
95 2014). The function fits PGLS regressions, tests for phylogenetic signal or calculates metrics for
96 trait evolution after iteratively removing user-defined fractions of species at random and compares
97 simulations with the model using the full dataset.

98 The `influ`-functions (Fig. 1) perform leave-one-out-deletion analysis to test if specific
99 species are strongly driving the results. For all species, these functions fit a new model without a

100 given species (reduced data) and compare the estimated parameters using the full dataset. This
 101 analysis can reveal influential cases (species driving relatively large changes in parameter
 102 estimates) and test model stability across samples (Field, 2013). The `clade`-functions (Fig.1)
 103 extend the same leave-one-out approach to detect influential clades (or more generally, groupings
 104 of species). The functions remove all species belonging to a clade and compare the reduced and
 105 the full datasets using a randomization test to correct for the number of species removed.

106 Three simple measures are used to estimate sensitivity in model parameters.

107 (i) the raw difference:

$$108 \quad db_i = b_i - b_0 \quad \text{eqn 1}$$

109 where b_i is the estimated parameter for the reduced dataset and b_0 is the estimated parameter for
 110 the full dataset;

111 (ii) the standardized difference:

$$112 \quad Sdb_i = db_i / SD_{db_i} \quad \text{eqn 2}$$

113 where SD_{db_i} is the standard deviation of db_i , thus Sdb_i is a simple z-score of db_i ; and

114 (iii) the percentage of change:

$$115 \quad Pdb_i = (|db_i| / b_0) * 100 \quad \text{eqn 3}$$

116 where $|db_i|$ is the absolute raw difference (eqn 1). While these functions provide useful estimates
 117 of how subsets of the dataset change key results, they do not account for potential structural biases

118 in the available data (e.g. bias in missing data). For instance, a common problem in comparative
119 analyses occurs when data is missing non-randomly with respect to the phylogeny. To help detect
120 this problem, we provide a supplementary function (`miss.phylo.d`), which detects
121 phylogenetic signal in missing data (D-statistics; Fritz and Purvis, 2010, Orme et al 2013).

122

123 ***Phylogenetic uncertainty***

124 Phylogenetic uncertainty refers to the notion that there are usually a number of alternative
125 phylogenetic hypotheses with different topologies and/or branch lengths. Yet, comparative studies
126 often analyse a single tree which is thought of as the ‘best’ estimate out of a family of candidate
127 phylogenies, without accounting for phylogenetic uncertainty, potentially biasing statistical
128 inference (Donoghue & Ackerly, 1996, Hernandez et al. 2013; Rangel et al. 2015). A simple way
129 to account for phylogenetic uncertainty in comparative methods is to repeat the analysis using a
130 sample of relevant phylogenetic trees (Donoghue & Ackerly, 1996). The influence of phylogenetic
131 uncertainty can be quantified by the amount of variation in model parameters between competing
132 models fitted with alternative trees (Hernandez et al. 2013; Martinez et al. 2015). The `tree-`
133 `functions` (Fig.1) account for multiple phylogenetic hypotheses, by rerunning the models over a
134 *multiPhylo* object containing different candidate phylogenies and comparing parameter estimates
135 across these reruns.

136

137 ***Data uncertainty***

138 Intraspecific variation due to differences between individuals or to measurement errors is
139 an important source of uncertainty and can influence both parameter estimation and hypothesis
140 testing (Felsenstein, 2008; Garamszegi & Møller, 2010; Silvestro et al. 2015). One way to account
141 for intraspecific variation is by simulating trait values for each species derived from the
142 intraspecific standard deviation of the mean, which users can calculate from their own data if they
143 have multiple measurements per tip (Martinez et al 2015). Rather than assuming a single trait value
144 per species, this approach tests the sensitivity of comparative models to variation in the underlying
145 trait data, accounting for the confidence range around the estimate (Garamszegi 2014). The
146 `intra`-functions (Fig.1) account for such uncertainties both in response and explanatory
147 variables. While the statistical distribution of such intraspecific variation may not always be
148 known, the functions implement two potential trait distributions (normal and uniform).

149

150 ***Interactions among uncertainty types***

151 Most users of phylogenetic comparative methods will face multiple sources of uncertainty
152 simultaneously (Cooper et al. 2016; Cornwell & Nakagawa 2017). Different types of uncertainty
153 can interact, potentially further reducing the robustness of a result. Yet, the interaction between
154 types of uncertainty is rarely studied (but see: Martinez et al. 2015), even in cases where sensitivity
155 to single uncertainties is quantified (Werner et al. 2014), potentially because of a lack of available
156 tools. We implemented functions to study interactions of both phylogenetic uncertainty (`tree`-
157 `functions`) and data uncertainty (`intra`-functions) with sampling uncertainty (`clade`-, `infl`-,
158 and `samp`-functions), as well as interactions between data and phylogenetic uncertainty.

159

160 Example 1: Influential clades

161 We included two datasets in sensiPhy: "primates" (Jones et al. 2009) and "alien"
162 (Gonzalez-Suarez et al. 2015). Each dataset contains a *multiPhylo* file with 101 phylogenetic trees
163 originated from pseudo-posterior distribution and pruned to match species in data (Fritz et al 2009;
164 Kuhn et al. 2011). As an example, we use the “primates” dataset to investigate how the deletion
165 of entire clades (families) can influence model parameters for a PGLS linear regression between
166 sexual maturity (days) and adult body mass (g).

167

```
168 > data("primates")  
169 > fit <- clade_phylm(log(sexMaturity) ~ log(adultMass),  
170 phy = primates.phy[[1]], data = primates.data, clade.col = "family",  
171 n.sim = 500, model = "lambda")
```

172

173 The function *clade_phylm* reruns the phylogenetic regression between sexual maturity and
174 body mass, iteratively leaving out individual families. This is defined by the argument ‘clade.col’
175 which indicates the grouping variable defining which species to include. Typically, these will be
176 taxonomically defined, but other groupings can be used, for instance based on geographic
177 locations, sampling methods or data sources. The function *sensi_plot* can be used to visualize

178 the results (Fig. 2) while `summary` shows the effect of each clade on model parameters (Table 1;
179 complete output in supplementary material).

180

```
181 > summary(fit) # table 1
```

```
182 > sensi_plot(fit, clade = "Cercopithecidae") # Fig 2AB
```

```
183 > sensi_plot(fit, clade = "Cebidae") # Fig 2CD
```

184

185 The analysis reveals that without species from the *Cercopithecidae* the regression slope is
186 22.8% higher than the full dataset model (Table 1; Fig 2a), indicating that this family has a major
187 negative influence on the relationship between sexual maturity and mass. Removal of *Cebidae*
188 species had a smaller and inverse effect (Table 1; Fig 2b) while *Lemuridae* species had only a
189 minor effect on model parameters (Table 1).

Table 1: Subset of the summary output from `clade_phylm`. Estimated model parameters after removing clades. `DIFestimate` indicates the shift in slope when excluding a species grouping (eqn 1), ‘change %’ expresses this as a percentage (eqn 3). `Pval.randomization` indicates the P-value for the randomization test (main text).

Clade removed	estimate	DIFestimate	change (%)	Pval	Pval.randomization
Cercopithecidae	0.308	0.057	22.8	5.7E-11	0.168
Cebidae	0.220	-0.031	12.2	7.3E-07	0.006
Callitrichidae	0.226	-0.024	9.8	5.3E-08	0.004
Lemuridae	0.258	0.008	3.1	1.3E-09	0.430

190

191 However, *Cercopithecidae* contains substantially more species (N=32) than *Cebidae* (N=

192 19). We would therefore expect *Cercopithecidae* to have a larger effect on parameter estimates,

193 by virtue of it containing a larger proportion of the species analysed. To correct for clade size, a

194 randomization test analyses if the change in parameter estimate is significantly different from a

195 null distribution when randomly removing the same number of species as the focal clade. The

196 randomisation test shows that in fact the *Cercopithecidae* are an influential clade only because

197 they contain a large number of species, not because the biological pattern is substantially different

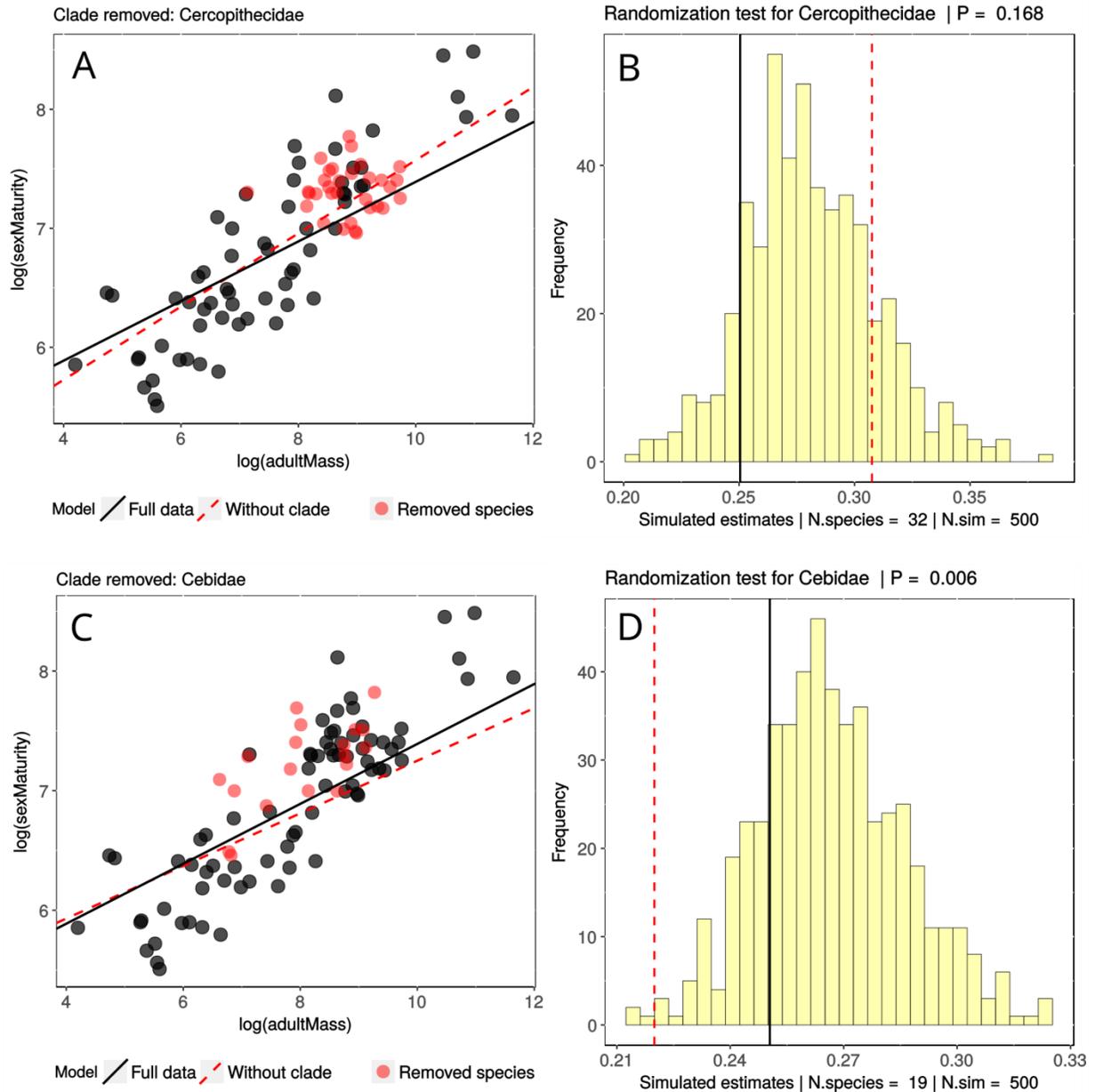
198 (P = 0.168, Table 1, Figure 2AB). This is different for the *Cebidae* (and the *Callitrichidae*), which

199 strongly influence our parameter estimates even when correcting for clade size, indicating a

200 substantially different pattern (P = 0.006, Table 1, Fig. 2CD). The exclusion of the *Lemuridae*

201 continues to have no effect, both in absolute terms and when correcting for clade size (Table 1).

202



203

204 **Fig 2.** Diagnostic graphs from the function **clade_phylm** for the clade Cercopitheciidae (A;B) and Cebidae (C;D).
 205 The effect of clade removal on the phylogenetic regression between sexual maturity and adult body mass of 95
 206 primates species (A;C). Null distribution of estimates after randomly removing the same number of species as the
 207 focal clade (B;D).

208

209

210 **Example 2: *Interaction among influential clades & phylogenetic uncertainty***

211 In the first example, we considered only a single primate phylogeny. However, a range of
212 alternative phylogenetic hypotheses is available for this group (Fritz et al. 2009; Kuhn et al. 2011).
213 We can use the function `tree_clade_phylm` to evaluate potential interactions among these
214 two uncertainty types.

215

```
216 > fit2 <- tree_clade_phylm(log(sexMaturity) ~ log(adultMass),  
217 phy = primates.phy, data = primates.data, clade.col = "family",  
218 n.sim = 100, n.tree = 30)
```

219

220 This function reruns Example 1 across multiple trees to test if the effect of clade removal
221 on model parameters interacts with phylogenetic uncertainty. The number of trees evaluated is set
222 with the argument 'n.trees'.

223

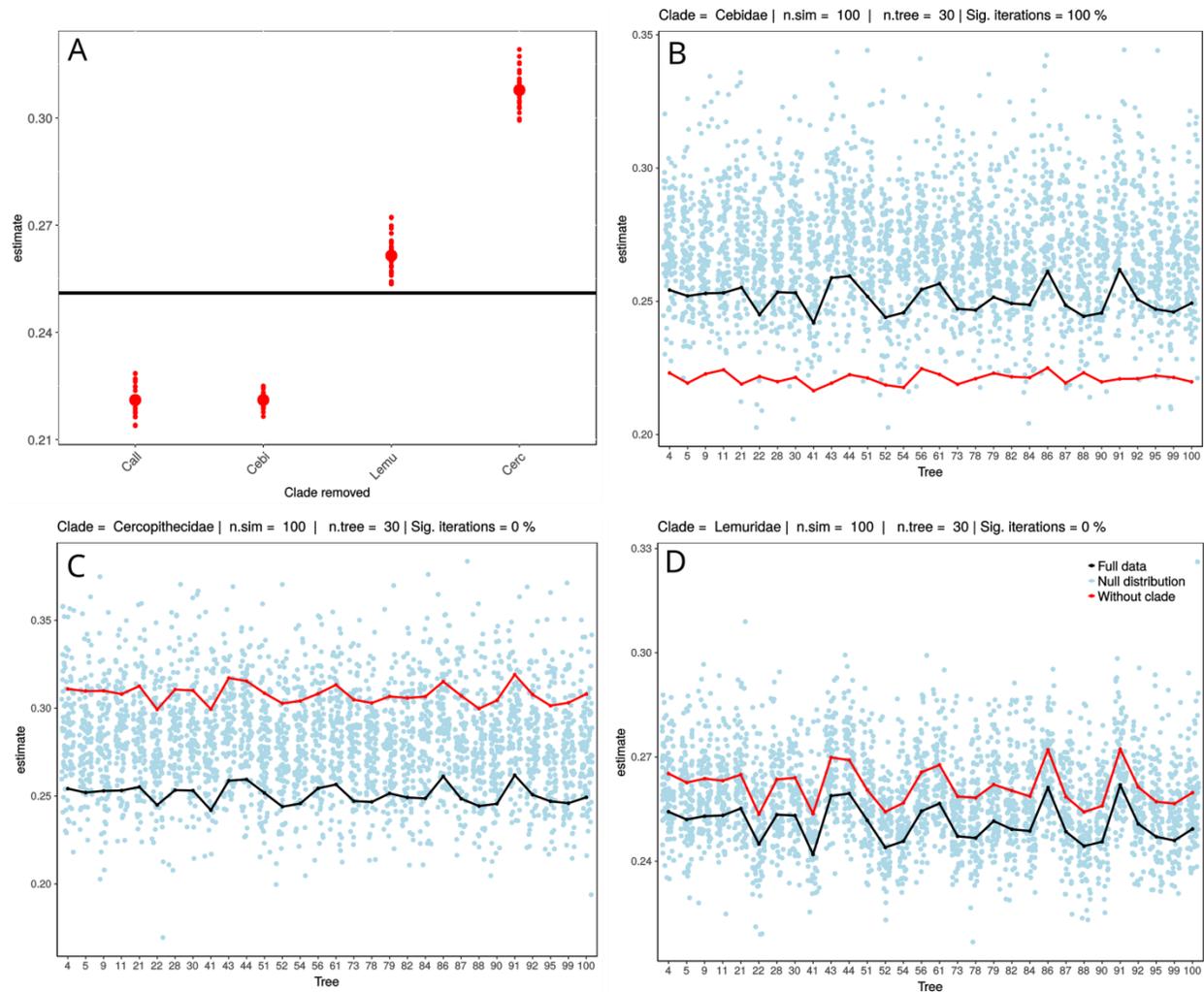
```
224 > summary(fit2) #Supplementary Table S1  
225 > sensi_plot(fit2, graphs = 1) # Fig 3A  
226 > sensi_plot(fit2, graphs = 2, clade = "Cercopithecidae") # Fig 3B  
227 > sensi_plot(fit2, graphs = 2, clade = "Cebidae") # Fig 3C
```

```
228 > sensi_plot(fit2, graphs = 2, clade = "Lemuridae") # Fig 3D
```

229

230 This analysis reveals that clade effects on estimates remained the same after taking into
231 account multiple phylogenetic trees (Fig. 3, Supplementary Table 1). For instance, the removal of
232 the *Cercopithecidae* family continues to cause a strong increase in slope (Fig. 3A). Furthermore,
233 the effect of *Cebidae* (and *Callitrichidae*) on parameter estimates is significantly different from
234 the null expectation across all alternative phylogenies tested (few blue dots below the red line in
235 Fig. 3B), while the effect of *Cercopithecidae* and *Lemuridae* falls within the null distribution (Fig.
236 3CD). Therefore, this analysis confirms the robustness of previous results, suggesting there is no
237 interaction among sampling and phylogenetic uncertainty.

238



239

240 **Fig 3.** Diagnostic graphs from the function `tree_clade_phylm`. (A) Estimated slopes after clade removal across
 241 multiple trees. Solid black line: average slope estimate among trees using the full dataset. Red dots: reruns between
 242 phylogenetic trees (small dots) and average estimate (larger dot). (B-D) The effect of clade removal on slope estimate
 243 across individual trees for Cebidae (B), Cercopithecidae (C) and Lemuridae (D). Blue dots: null expectation estimates
 244 after removing the same number of species as in the focal clade.

245

246 Implications & Solutions of a sensitive result

247 Sensitivity analyses from `sensiPhy` can be a starting point for further analyses (Table 2).

248 Considering our examples, a first step could be to verify if the *Cebidae* data are somehow biased,

249 resulting in a substantially different pattern. For instance, perhaps a different method to estimate
250 sexual maturity was used than in the other primates, which may have overestimated age of sexual
251 maturity in this clade. Alternatively, there could be biological reasons why the *Cebidae* show a
252 stronger correlation among traits, which could provide interesting biological insight. New
253 biological hypotheses could in turn be tested using comparative analyses. For instance, if an
254 interaction with climate might drive the differential effects of body mass on sexual maturity in the
255 *Cebidae* and the *Callitrichidae*, an expanded comparative analysis could test that hypothesis.

256 We highlight that a sensiPhy-analysis cannot directly reveal the underlying reason why a
257 biological effect is not robust to a given type of uncertainty. This can be for various methodological
258 reasons or reflect an actual biological effect. While the implications of finding that a biological
259 conclusion is sensitive to some, or multiple, forms of uncertainty will be highly context and model-
260 system specific, we provide general pointers and solutions that users can explore (Table 2).

261

262 **Table 2:** Potential implications and solutions when finding sensitive results

Biological question	sensiPhy method	Implications / potential solutions
Do influential species or clades drive result?	clade or influ	<ol style="list-style-type: none"> 1. Verify if data is biased in influential species/clades? 2. Identify biological drivers of influential species/clades. 3. Ideally, verify (2) by including as term in comparative model.
Does sampling effort influence results?	samp	<ol style="list-style-type: none"> 1. Increase sample size (overall). 2. If interaction with specific clades, increase sample size in those clades. 3. Consider test for phylogenetic signal in missing data.
Does intraspecific variation influence results?	intra	<ol style="list-style-type: none"> 1. Verify if driven by imprecise measurements. Can we measure variables to greater precision? 2. Explicitly quantify intraspecific vs interspecific variation in phylogenetic context (Garamszegi, 2014). 3. Consider if species level is the most appropriate level of analysis for this variable.
Does phylogenetic uncertainty influence results?	tree	<ol style="list-style-type: none"> 1. Verify if specific (influential) trees have methodological issues. 2. Can we increase resolution/precision of our phylogenetic tree (e.g. include more/better genetic markers)?

263

264

265 **Conclusions and future directions**

266 The sensiPhy-package offers a quick and easy approach to check the robustness of
267 frequently used comparative methods to multiple types of uncertainties. Performing sensitivity
268 analysis can greatly benefit authors by providing ways to estimate and account for uncertainties
269 and to detect and report possible bias in inference. The package helps researchers to be extra
270 careful with their results in an easy and straightforward way, increasing transparency in reporting
271 results from comparative analyses. We hope sensiPhy will encourage the inclusion of sensitivity
272 analysis as a common practice in comparative biology. The statistical reasoning implemented in

273 sensiPhy can be applied more generally to many other types of analyses. The package is open-
274 platform and welcomes users to contribute with new functionalities, facilitating new developments
275 for sensitivity analysis in phylogenetic comparative methods through the Github platform.

276

277 **Acknowledgements**

278 We are grateful to Carlos Fonseca, László Garamszegi, Marcelo Salas and Pablo Martinez for their
279 suggestions and insightful comments on sensiPhy and on earlier versions of this manuscript. We
280 also thank the editor and three reviewers for their many helpful comments. GBP is supported by a
281 CAPES Doctoral scholarship (Brazil). GDAW is supported by a Newton International Fellowship
282 (Royal Society).

283

284 **Author Contributions statement**

285 GBP, CP and GDAW conceived the ideas, developed the statistical reasoning, wrote the code and
286 the manuscript. All authors contributed equally to this work and gave final approval for
287 publication.

288

289 **Data accessibility**

290 All data and code used in this manuscript are available on Github
291 (<https://github.com/paternogbc/sensiPhy>) and deposited at Zenodo
292 (<http://doi.org/10.5281/zenodo.1179248>).

293

294 **Supporting information**

295 **Appendix S1.** A reproducible report containing the source code used to generate all statistical
296 results, figures and tables in this manuscript.

297

298 **References**

299 Blomberg, S.P., Garland, T. & Ives, A.R. (2003). Testing for phylogenetic signal in comparative
300 data: behavioral traits are more labile. *Evolution*, **57**, 717–745.

301 Cooper N, Thomas GH, FitzJohn RG (2016) Shedding light on the “dark side” of phylogenetic
302 comparative methods. *Methods in Ecology and Evolution* **7**(6):693–699.

303 Cornwell & Nakagawa (2017) Phylogenetic comparative methods. *Current Biology* **27**(9):33-336

304 Donoghue, M.J. & Ackerly, D.D. (1996). Phylogenetic Uncertainties and Sensitivity Analyses in
305 Comparative Biology. *Philosophical Transactions: Biological Sciences*, **351**(1345), 1241-
306 1249.

307 Efron, B. (1982). The Jackknife, the Bootstrap and Other Resampling Plans. CBMS-NSF Regional
308 Conference Series in Applied Mathematics, Monograph 38, SIAM, Philadelphia.

- 309 Felsenstein, J. (2008). Comparative methods with sampling error and within-species variation:
310 contrasts revisited and revised. *The American naturalist*, **171**, 713–725.
- 311 Fritz, S.A., Bininda-Emonds, O.R.P. & Purvis, A. (2009). Geographical variation in predictors of
312 mammalian extinction risk: big is bad, but only in the tropics. *Ecology Letters*, **12**, 538–549.
- 313 Fritz, S. A. and Purvis, A. (2010). Selectivity in mammalian extinction risk and threat types: a new
314 measure of phylogenetic signal strength in binary traits. *Conservation Biology*, **24**(4):1042-
315 1051.
- 316 Garamszegi, L.Z. & Møller, A.P. (2010). Effects of sample size and intraspecific variation in
317 phylogenetic comparative studies: A meta-analytic review. *Biological Reviews*, **85**(4), 797-
318 805.
- 319 Garamszegi, L.Z. (2014). *Modern Phylogenetic Comparative Methods and Their Application in*
320 *Evolutionary Biology* (L.Z. Garamszegi, Ed.). Springer
- 321 González-Suárez, M., Bacher, S. & Jeschke, J.M. (2015). Intraspecific trait variation is correlated
322 with establishment success of alien mammals. *The American naturalist*, **185**, 737–46.
- 323 Harmon, L. J., Weir, J. T., Brock, C. D., Glor, R. E., Challenger, W. (2008). GEIGER:
324 investigating evolutionary radiations. *Bioinformatics*. **24** (1), 129-131.
- 325 Hernandez, C. E., Rodríguez-Serrano, E., Avaria-Llautureo, J., Inostroza-Michael, O., Morales-
326 Pallero, B., Boric-Bargetto, D., Canales-Aguirre, C.B., Marquet P.A. & Meade, A. (2013).
327 Using phylogenetic information and the comparative method to evaluate hypotheses in
328 macroecology. *Methods in Ecology and Evolution*, **4**(5), 401-415.

- 329 Ho, L.S.T. & Ané, C. (2014). A linear-time algorithm for Gaussian and non-Gaussian trait
330 evolution models. *Systematic Biology*, **63**(3), 397-408.
- 331 Huelsenbeck, J.P., Rannala, B. & Masly, J.P. (2000). Accommodating phylogenetic uncertainty in
332 evolutionary studies. *Science (New York, N.Y.)*, **288**, 2349–2350.
- 333 Jones, K.E., et al. (2009). PanTHERIA: a species-level database of life history, ecology, and
334 geography of extant and recently extinct mammals. *Ecology*, **90**, 2648–2648.
- 335 Orme, D., et al. (2013). caper: Comparative Analyses of Phylogenetics and Evolution in R. R
336 package version 0.5.2. <https://CRAN.R-project.org/package=caper>
- 337 Martinez, P. a., Zurano, J.P., Amado, T.F., Penone, C., Betancur-R, R., Bidau, C.J. & Jacobina,
338 U.P. (2015). Chromosomal diversity in tropical reef fishes is related to body size and depth
339 range. *Molecular Phylogenetics and Evolution*, **93**, 1–4.
- 340 Magallon, S. & Sanderson, M. J. (2001), Absolute diversification rates in angiosperm clades.
341 *Evolution*, **55**, 1762–1780.
- 342 Pagel, M. (1999). The Maximum Likelihood Approach to Reconstructing Ancestral Character
343 States of Discrete Characters on Phylogenies. *Systematic Biology*, **48**, 612–622.
- 344 Pennell, M.W., R.G. FitzJohn, W.K. Cornwell, and L.J. Harmon. 2015. Model adequacy and the
345 macroevolution of angiosperm functional traits. *The American Naturalist* **186**(2): E33-E50.
- 346 Paradis, E. (2012). *Analysis of Phylogenetics and Evolution with R*. Springer New York, NY.

- 347 Kuhn, T.S., Mooers, A.Ø. & Thomas, G.H. (2011). A simple polytomy resolver for dated
348 phylogenies. *Methods in Ecology and Evolution*, **2**, 427–436.
- 349 R Core Team. (2017). R: A language and environment for statistical computing. R Foundation for
350 Statistical Computing. Vienna, Austria.
- 351 Rangel, T.F., Colwell, R.K., Graves, G.R., Fučíková, K., Rahbek, C. & Diniz-Filho, J.A.F. (2015).
352 Phylogenetic uncertainty revisited: Implications for ecological analyses. *Evolution*, **69**,
353 1301–1312.
- 354 Revell, L. J. (2012) phytools: An R package for phylogenetic comparative biology (and other
355 things). *Methods in Ecology and Evolution*, **3**, 217-223.
- 356 Silvestro, D., Kostikova, A., Litsios, G., Pearman, P.B. & Salamin, N. (2015). Measurement errors
357 should always be incorporated in phylogenetic comparative analysis. *Methods in Ecology
358 and Evolution*, **6**, 340–346.
- 359 Werner, G.D.A., Cornwell, W.K., Sprent, J.I., Kattge, J. & Kiers, E.T. (2014). A single
360 evolutionary innovation drives the deep evolution of symbiotic N₂-fixation in angiosperms.
361 *Nature Communications*, **5**, 4087.