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

Gustavo Brant Paterno[1 *]; Caterina Penone[2 *]; Gijsbert D.A. Werner[3,4 *]

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

[2]Institute of Plant Sciences, University of Bern, CH

[3]Department of Zoology, University of Oxford, UK

[4]Balliol College, University of Oxford, UK

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

* These authors contributed equally to this work

Summary

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

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

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

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

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

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

Introduction

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

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

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

The sensiPhy package

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

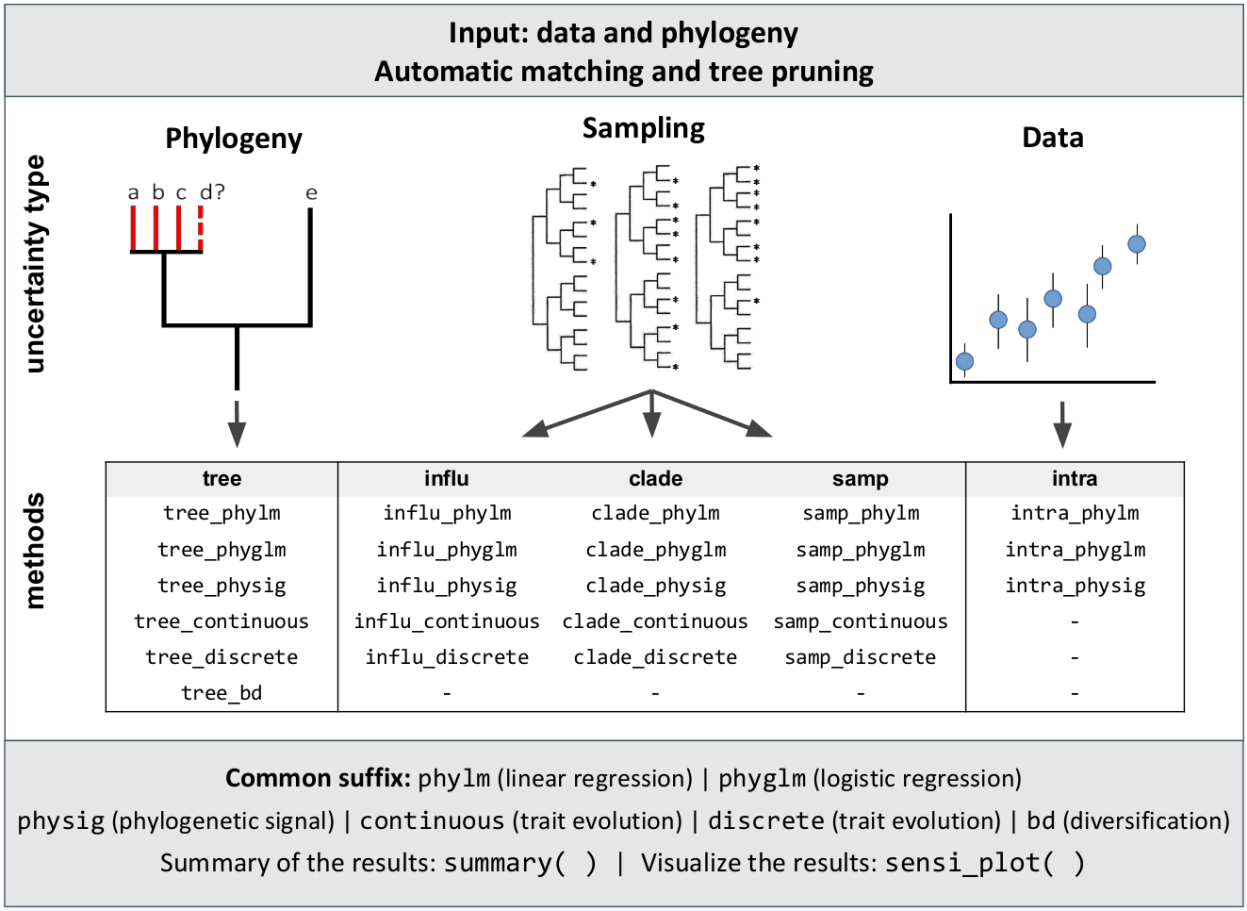


Fig1: Overview of the main functions in sensiPhy organized by source of uncertainty. sensiPhy contains functions to quantify the effects of the three types of uncertainty and of interactions among them: phylogenetic uncertainty (`tree`), uncertainty arising from species sampling (`influ`, `clade` and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

Sources of uncertainty

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

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

Species sampling uncertainty

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

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

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

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

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

(i) the raw difference:

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

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

(ii) the standardized difference:

$$Sdb_i = db_i / SD_{dbi} \qquad \text{eqn 2}$$

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

(iii) the percentage of change:

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

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

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

Phylogenetic uncertainty

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

Data uncertainty

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

Interactions among uncertainty types

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

Example 1: ***Influential clades***

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

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

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

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

```
> summary(fit)                                          # table 1
> sensi_plot(fit, clade = "Cercopithecidae") # Fig 2AB
> sensi_plot(fit, clade = "Cebidae")          # Fig 2CD
```

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

However, *Cercopithecidae* contains substantially more species (N=32) than *Cebidae* (N= 19). We would therefore expect *Cercopithecidae* to have a larger effect on parameter estimates, by virtue of it containing a larger proportion of the species analysed. To correct for clade size, a randomization test analyses if the change in parameter estimate is significantly different from a null distribution when randomly removing the same number of species as the focal clade. The randomisation test shows that in fact the *Cercopithecidae* are an influential clade only because they contain a large number of species, not because the biological pattern is substantially different (P = 0.168, Table 1, Figure 2AB). This is different for the *Cebidae* (and the *Callitrichidae*), which strongly influence our parameter estimates even when correcting for clade size, indicating a substantially different pattern (P = 0.006, Table 1, Fig. 2CD). The exclusion of the *Lemuridae* continues to have no effect, both in absolute terms and when correcting for clade size (Table 1).

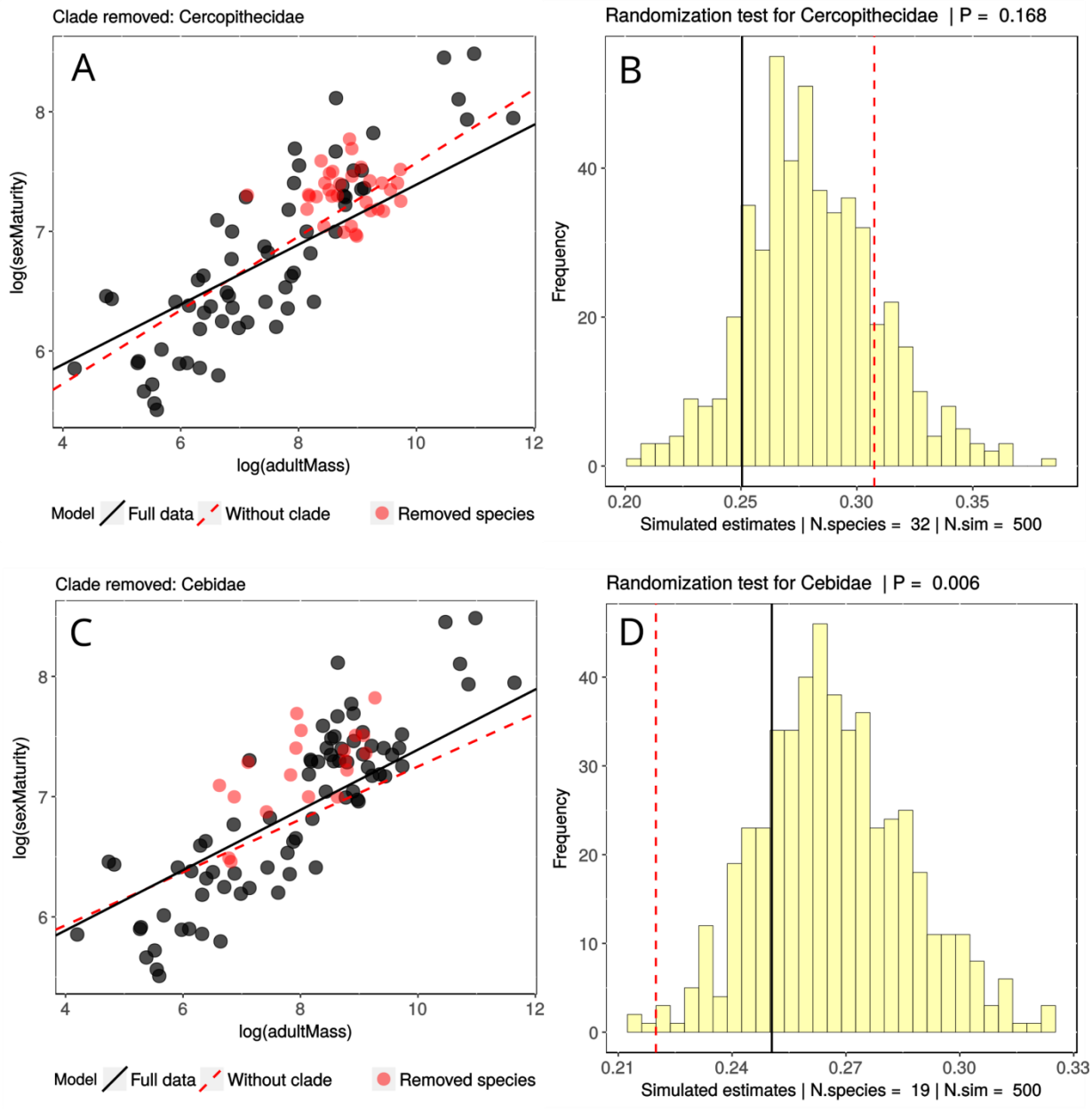


Fig 2. Diagnostic graphs from the function **clade_phylm** for the clade Cercopithecidae (A;B) and Cebidae (C;D). The effect of clade removal on the phylogenetic regression between sexual maturity and adult body mass of 95 primates species (A;C). Null distribution of estimates after randomly removing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

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

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

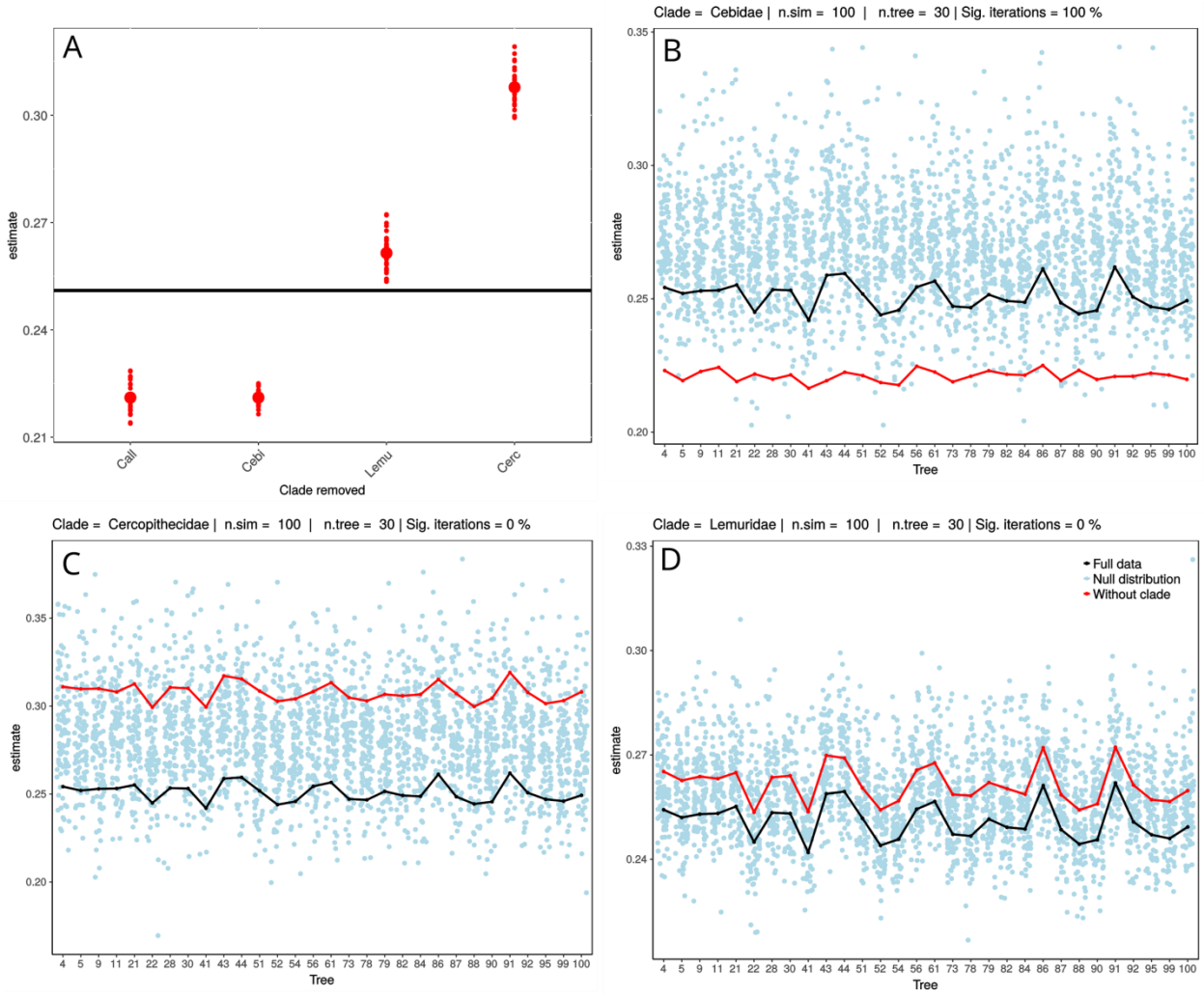
```
> fit2 <- tree_clade_phylm(log(sexMaturity) ~ log(adultMass),
phy = primates.phy, data = primates.data, clade.col = "family",
n.sim = 100, n.tree = 30)
```

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

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

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

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



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,

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

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

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`	**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`	**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`	**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`	**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)?

Conclusions and future directions

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

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

Acknowledgements

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

Author Contributions statement

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

Data accessibility

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

Supporting information

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

References

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

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

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

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

Efron, B. (1982). The Jackknife, the Bootstrap and Other Resampling Plans. CBMS-NSF Regional 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.

Kuhn, T.S., Mooers, A.Ø. & Thomas, G.H. (2011). A simple polytomy resolver for dated phylogenies. *Methods in Ecology and Evolution*, **2**, 427–436.

R Core Team. (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria.

Rangel, T.F., Colwell, R.K., Graves, G.R., Fučíková, K., Rahbek, C. & Diniz-Filho, J.A.F. (2015). Phylogenetic uncertainty revisited: Implications for ecological analyses. *Evolution*, **69**, 1301–1312.

Revell, L. J. (2012) phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, 3, 217-223.

Silvestro, D., Kostikova, A., Litsios, G., Pearman, P.B. & Salamin, N. (2015). Measurement errors should always be incorporated in phylogenetic comparative analysis. *Methods in Ecology and Evolution*, **6**, 340–346.

Werner, G.D.A., Cornwell, W.K., Sprent, J.I., Kattge, J. & Kiers, E.T. (2014). A single evolutionary innovation drives the deep evolution of symbiotic N2-fixation in angiosperms. *Nature Communications*, **5**, 4087.