

Asymmetries in phylogenetic diversification and character change can be untangled

Emmanuel Paradis

► To cite this version:

Emmanuel Paradis. Asymmetries in phylogenetic diversification and character change can be untangled. Evolution - International Journal of Organic Evolution, 2007, 62 (1), pp.241 - 247. 10.1111/j.1558-5646.2007.00252.x . hal-01822133

HAL Id: hal-01822133 https://hal.science/hal-01822133

Submitted on 23 Jun 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ASYMMETRIES IN PHYLOGENETIC DIVERSIFICATION AND CHARACTER CHANGE CAN BE UNTANGLED

Emmanuel Paradis

Institut de Recherche pour le Développement, UR175 CAVIAR, GAMET – BP 5095, 361 rue Jean François Breton, F-34196 Montpellier Cédex 5, France E-mail: Emmanuel.Paradis@mpl.ird.fr

LRH: BRIEF COMMUNICATIONS

RRH: BRIEF COMMUNICATIONS

Abstract.—The analysis of diversification and character evolution using phylogenetic 1 data attracts increasing interest from biologists. Recent statistical developments have 2 resulted in a variety of tools for the inference of macroevolutionary processes in a 3 phylogenetic context. In a recent paper Maddison (2006) pointed out that uncareful 4 use of some of these tools could lead to misleading conclusions on diversification or 5 character evolution, and thus to difficulties in distinguishing both phenomena. I here 6 present guidelines for the analyses of macroevolutionary data that may help to avoid 7 these problems. The proper use of recently developed statistical methods may help to 8 untangle diversification and character change, and so will allow us to address 9 important evolutionary questions. 10

¹¹ Keywords.—diversification, extinction, maximum likelihood, phylogeny, speciation.

The last twenty years have witnessed a remarkable change in paradigm for 12 evolutionists: the variation in speciation and extinction rates (i.e., the tempo of 13 evolution), and the variation in the rates of character change (the mode of evolution) 14 are now ideally studied with molecular phylogenies and data collected on recent 15 species. These issues, traditionally called macroevolution, were the domain of 16 paleontologists during several decades (Simpson 1953), whereas molecular data were 17 used to address microevolutionary mechanisms (to have an idea of this changing 18 paradigm, compare the two editions of the same book by Futuyma 1986, 1998). 19 In a recent paper, Maddison (2006) used simulated data following two scenarios to 20 bring attention to some limits of this new paradigm. In the first scenario, some clades 21 were simulated starting from the root, and a character evolved along the tree: two 22 states were allowed (0 and 1) with a constant probability of change between them. 23 The speciation rate was related to the state of the character so that species in state 1 24 split at a higher rate than those in state 0. Maddison showed that the estimates of 25 the ratio of character transition rates was correlated with the ratio of the speciation 26 rates. In the second scenario, Maddison used a similar setting except that the 27 speciation rate was constant, but the rates of transition of the character in one 28 direction was four times higher than in the other. He then showed that the analysis of 29 diversification led to infer, more frequently than by chance, that the speciation rate 30 was different between the states of the character. 31

Maddison (2006) concludes that, if simple data analyses are done, biased 32 diversification with respect to a character may lead to wrong conclusions on the 33 evolution of this character. On the other hand, biased character evolution may lead to 34 falsely associate this character to biased diversification. In other words, this raises the 35 question whether we may be able to untangle differential evolution of character from 36 differential diversification. The question may be framed in more specific terms from 37 Maddison's simulation study: if a character state is observed to be relatively rare in a 38 clade, can we distinguish whether this state is associated with a low speciation rate. 39 or evolution towards this state occurs at a low rate? (Or both?) 40

⁴¹ This issue is of great importance because, if we can answer positively, we could

point which effect most contributes to the generation of diversity. Heterogeneous diversification rates and heterogeneous character change rates are likely linked to different evolutionary mechanisms. Character states with a high transition rate may be the result of counter-selection, developmental instability, or low plasticity of the character. On the other hand, a high speciation rate related to a character state may be the result of its adaptative value, or its association with the rate of reproductive isolation.

My aim in the present paper is to show that appropriate data analysis methods can separate the effects of biased character change and biased diversification. This suggests that heterogeneous character evolution and heterogeneous diversification rates can be untangled in future analyses of macroevolutionary data. I point to some recommendations for data analyses in macroevolutionary studies with recent species, as well as further needs for future research in this area.

55

56

Methods

Analysis of Character Change

Maddison (2006) showed that the maximum likelihood estimates of the ratio of 57 character transition rates was correlated with the actual ratio in speciation rates. One 58 may be tempted to interpret a high ratio of character transition rates as evidence for 59 a strong bias in character change, thus concluding that rates are actually different. 60 However, Maddison did not address the issue of testing whether these rates are 61 significantly different. Indeed, looking at parameter estimates is not the ideal way to 62 assess the validity of an hypothesis. The results from Maddison's analyses illustrate 63 that these parameter estimates can be strongly biased when a wrong model is used, 64 not that they are significantly different. 65

With simulated data, we know the model used to generate the data, but with real data we have to test whether a given model is appropriate. In a likelihood framework, the likelihood ratio test (LRT) is the canonical way for hypothesis testing (Lindsey 1996; Ewens and Grant 2005). To fix ideas, let us write explicitly the model used to

$$\begin{array}{ccc}
0 & 1 \\
0 & \left[\begin{array}{c}
-r & r \\
r & -r \end{array} \right],
\end{array}$$
(1)

where the row labels denote the initial states of the character (denoted x in this paper), the column labels the final ones, and r is the instantaneous rate of change, that is the probability that x changes of state during a very short interval of time. It is difficult to interpret the parameter r biologically, but its virtue is that it is independent of time. To obtain real probabilities, we have to fix a time interval, say t, and compute the matrix exponential e^{tQ} (see below).

Maddison estimated the rates ratio using the following two-parameter model(referred to as model 2):

$$\begin{bmatrix} -r_1 & r_1 \\ r_2 & -r_2 \end{bmatrix}.$$
 (2)

We know, in this particular situation, that this model has an extra parameter because the data were simulated with $r_2 = r_1$. Models 2 and 1 can be compared with a LRT because they are nested: the latter is a particular case of the former (Pagel 1994; Nosil and Mooers 2005). The test follows a χ^2 distribution with df = 1 (the difference in number of parameters).

In real situations we cannot be sure that either model 1 or model 2 generated the 85 observed data. If neither of them are the true model, the tests of hypotheses could be 86 biased in the same way than Maddison showed that parameter estimates are biased. 87 It is thus necessary to assess the goodness-of-fit of a model in a general way. In the 88 present context, the shape of the likelihood function is an indication of the poor or 89 good fit of a model. If a model poorly fits the data, the likelihood function is 90 relatively flat. It is possible to examine the shape of the likelihood function by 91 plotting it against a range of parameter values, but an easier procedure is to look at 92

the estimated standard-errors of the parameters which are derived, under the 93 standard likelihood method, from the second derivatives of the likelihood function. 94 Thus, the smaller these standard-errors, the narrower the likelihood function. When 95 no analytical expression of the second derivatives are possible, which is precisely the 96 case for the Markovian models considered here, a numerical computation may be done 97 using nonlinear optimization (Schnabel et al. 1985). The ratio of the parameter 98 estimate, \hat{r} , on its standard-error, $se(\hat{r})$, can be used as an indication of the shape of 99 the likelihood. This ratio is in fact a formulation of the Wald test which, under the 100 null hypothesis that r is equal to zero, follows a standard normal distribution (see 101 Rao 1973): 102

$$\frac{\hat{r}}{\operatorname{se}(\hat{r})} \sim \mathcal{N}(0,1) \quad \text{if } r = 0.$$

The LRT and the Wald test can thus be used to test the same hypothesis, but the latter is rarely used because it has generally poorer statistical performance than the LRT, particularly for small sample sizes (Agresti 1990; McCulloch and Searle 2001). However, both tests are expected to give the same results for large sample sizes (Rao 1973; McCulloch and Searle 2001).

I propose the following rule of thumb: when this ratio is less than two, it is likely 108 that the model under consideration is not appropriate. A justification for this rule is 109 that, under the assumption that the maximum likelihood estimates are normally 110 distributed, (a standard assumption of the likelihood theory of estimation), then a 111 95% confidence interval may be calculated with $\hat{r} \pm 1.96 \times \text{se}(\hat{r})$. Consequently, if 112 $\hat{r}/\text{sec}(\hat{r}) < 2$ then zero is included in this interval suggesting the presence of a flat 113 likelihood function. The method should be used as follows. First, perform the LRT 114 comparing both models. If this test is significant, then examine the ratios of the rate 115 estimates of model 2 on their standard-errors: if one of them is less than two, then it 116 is likely that the LRT is biased and the null hypothesis is true. 117

Analysis of Diversification

Maddison (2006) inferred differences in speciation rates by comparing the number of 119 species between sister-clades that are different with respect to the character, but all 120 species within each clade have the same character state. This method does not use all 121 available information as it considers only a subset of the nodes of the tree, and it 122 ignores branch lengths. Several methods have been proposed to analyze diversification 123 that essentially differ in the type of data under consideration (see reviews in 124 Sanderson and Donoghue 1996; Mooers and Heard 1997; Pagel 1999). For instance, 125 some methods consider tree topology and balance (e.g., Aldous 2001), whereas others 126 consider the distribution of branch lengths (e.g., Pybus and Harvey 2000). When the 127 topology and branch lengths of the analyzed phylogeny are available, it is possible to 128 refine the analyses and use more elaborate methods. Recent developments have been 129 done in the inference of diversification, particularly the Yule model with covariates 130 which takes full account of all phylogenetic information (tree topology and branch 131 lengths) to infer the effects of species traits on speciation rates (Paradis 2005). In this 132 model the speciation rate depends on a linear combination of some variables 133 measured on each species. This approach is similar to a standard linear regression 134 where the mean of the response is given by a linear combination of variables. 135 Consequently, a wide variety of models may be fitted to the same phylogenetic and 136 species traits data. The latter could be continuous and/or discrete. 137

Because the Yule model with covariates is fitted by maximum likelihood, different models can be compared with a LRT if they are nested. In the present context of testing the effect of x on the speciation rate (λ) , the general model is:

$$\log \frac{\lambda}{1-\lambda} = \beta x + \alpha$$

where α and β are parameters, so that the right-hand side of the above equation is equal to α if x = 0, or to $\alpha + \beta$ if x = 1 (see Paradis 2005, for details). The LRT comparing this model to the standard Yule model tests the hypothesis of the effect of x on λ , i.e., whether β is significantly different from zero. The Wald test may also be performed by computing $\hat{\beta}/\text{se}(\hat{\beta})$. The same criterion proposed above for the analysis of character was applied as well.

Data Simulation

147

I simulated some data with known parameters in order to assess the statistical 148 performance of the methods described above. The idea was to generate some data 149 sets where the majority of the species are in a state due to, either different speciation 150 rates, or different rates of character change, or both. The data were then analyzed to 151 assess whether the two processes can be untangled. The simulations were started 152 from the root of the tree, that is the initial bifurcation. At each time-step, each 153 species present in the clade had a given probability (λ) to split into two, and a 154 probability (p) to change the state of its character x. When a species split, the 155 daughter-species inherited its value of x. Both λ and p depend on x, and so are 156 hereafter denoted λ_0 , p_0 , λ_1 , and p_1 , where the subscript indicates the value of x. 157 Two of these parameters were fixed for all simulations: $\lambda_0 = 10^{-4}$ and $p_0 = 5 \times 10^{-5}$. 158 Three different combinations of parameters were used for λ_1 and p_1 : (1) $\lambda_1 = 5\lambda_0$, 159 $p_1 = p_0$, (2) $\lambda_1 = \lambda_0$, $p_1 = p_0/10$, and (3) $\lambda_1 = 2.5\lambda_0$, $p_1 = p_0/4$. 160

These settings correspond to the three plausible biological scenarios leading to the 161 abundance of a trait in a clade: (1) this trait is associated with a high speciation rate, 162 (2) species without this trait tend to evolve towards acquiring it, and (3) a mixture of 163 both processes. The parameter values were chosen so that ca. 90% of species had 164 x = 1 at the end of the simulation. These were found analytically in setting (2), since 165 speciation was homogeneous, using the matrix exponentiation explained above, and 166 the fact that the expected number of species after t time-steps is given by $2e^{\lambda t}$ 167 (Kendall 1948). It was obviously unnecessary to consider here the null setting $\lambda_1 = \lambda_0$ 168 and $p_1 = p_0$ because this would yield 50% of species in each state, and so little 169 difficulty for data analysis. The first setting is similar to Maddison's (2006) first 170 scenario, whereas the second setting is close to his second one: the difference is that 171 he used a ratio of 4 instead of 10. 172

The time-step of the simulations was transformed in time unit equal to 0.001.

Giving a probability of change of 5×10^{-5} for t = 0.001, we can find by 174 back-transformation with the matrix exponentiation and interpolation that the actual 175 parameter of the first setting was $r \approx 0.05$. All simulations were run until 100 species 176 were present. This was replicated 100 times for each possible initial state at the root 177 (0 or 1) and each combination of the parameters. The trees and values for x at the 178 end of the simulation were saved (the files are available from the author). All 179 simulations were programmed in R version 2.4.0 (R Development Core Team 2006). 180 The data were analyzed as if they were real data: the procedures described above 181 were applied to all replicates using R's standard looping functions. Both LRTs and 182 Wald tests were computed, as well as the proportion of cases where the tests agreed in 183 rejecting the null hypothesis. The analyses of character change and of diversification 184 were done with the package ape (Paradis et al. 2004). 185

186

194

Results

The results were slightly affected by the state of the root as it changed slightly the proportion of species in state 1 at the end of the simulation: 83.8% and 91.6%, for a root in state 0 or 1, respectively (data pooled over all simulations). This proportion had in fact a skewed distribution when calculated for each replicate; the corresponding medians were thus slightly larger: 85% (range: 28–99) and 92% (63-99). For simplicity, the results below are presented for the two series of simulations pooled since they were overall consistent.

Analysis of Character Change

Table 1

The proportions of replicates where the LRT comparing models 1 and 2 was significant at the 5% level were 0.18, 0.405, and 0.175 for the three scenarios of equal rate of change, strong, and moderate biases, respectively (Table 1). On the other hand, the Wald test had relatively poor performance as the rejection rate of the null hypothesis was almost the same whatever the scenario (*ca.* 0.25). After examining the standard-errors of the parameter estimates, the proportions of cases where the LRT was significant and the ratios $\hat{r}_1/\text{se}(\hat{r}_1)$ and $\hat{r}_2/\text{se}(\hat{r}_2)$ were greater than 2 were lowered

to 0.055, 0.26, and 0.1. Only the first one was not significantly different from 0.05202 (two-sided exact binomial test: P = 0.744, P < 0.0001, and P = 0.003, respectively). 203 Figure 1 gives a summary of the distributions of the estimates under both models 204 for the three settings. As observed by Maddison (2006), a few cases resulted in 205 extremely dispersed estimates, so I focused on the median and first and third 206 quartiles. In the first setting, the median of the estimator \hat{r} was very close to the true 207 value (0.049, instead of 0.05). Remarkably, the dispersion of all parameter estimates 208 and their standard-errors were smaller when model 2 was true. 209

The reconstruction of ancestral states is also informative. Under model 2 nearly all nodes (except the most terminal ones) had a relative likelihood of 0.5 for each state, meaning that these reconstructions were in fact indeterminate under this model. Under model 1 most nodes were estimated to be in state 1 with a high probability (relative likelihood greater than 0.9): this was particularly the case for the root even if the actual state was 0 (Fig. 2). Thus the models failed to estimate correctly the ancestral states of the simulated character.

Analysis of Diversification

The simulated trees and phenotypic values were analyzed with the Yule model with 218 covariates testing for the effect of x on λ . The method requires us to reconstruct the 219 values of x at the nodes of the tree: this was done with a simple parsimony criterion. 220 The branch lengths were back-transformed to the original time scale of the 221 simulations prior to analysis in order to ease the fitting procedure. The proportion of 222 significant LRT were 0.415, 0.01, and 0.13. All these proportions are significantly 223 different from 0.05 (two-sided exact binomial test: $P \leq 0.005$ in all cases). The Wald 224 test gave similar results, but with slightly larger rejection rates of the null hypothesis. 225 However, in all cases where the LRT was significant the Wald test was as well 226 (Table 1). It is interesting to note that though the estimated ancestral states were 227 inaccurate, the test was able to reject the null hypothesis in more than 5% of the 228 cases, and so is robust, at least partially, to inexact ancestral character estimation (at 229 least in the present situation). 230

Fig. 2

217

DISCUSSION

The combined use of the LRT and Wald test here revealed an interesting contrast 232 between analyses. In the analysis of character change, the proposed criterion derived 233 from the Wald test helped to keep the type I error rate at a reasonable value. In the 234 analysis of diversification, the LRT performed well and there is, in the present 235 scenarios, no need to use the proposed criterion. How to explain this difference? 236 There are three possible, nonexclusive explanations of the relatively poor performance 237 of the models of character change. Firstly, the departures from the Markovian 238 assumptions may be too strong so that the models considered here are too 'ill-defined' 239 for the present data. Secondly, the maximization of the likelihood functions of these 240 models is not straightforward as it requires iterative calculations of likelihoods along 241 the nodes of the tree which may be difficult for current optimization methods. By 242 contrast, the likelihood function of the Yule model with covariates is a linear function 243 of the parameters (Paradis 2005). Thirdly, the current formulation of the model of 244 character change may not be adequate for likelihood maximization, and a 245 reparametrization resulting in a linear combination of parameters might be more 246 appropriate. 247

Mooers and Schluter (1999) were cautious about the approximation of the LRT 248 with a χ^2 distribution, and recommended using conservative values to assess the 249 increase in likelihood when fitting model 2. They also called for more work to verify 250 the validity of this approximation. In the present study, I observed in some cases a 251 difficulty to fit the models which required to repeat the fitting procedure with 252 different starting values. Some further study is needed to assess whether this is due to 253 the inadequacy of these models, or a failure of our current algorithms of likelihood 254 maximization. 255

All simulations were run with parameter values that gave equal number of species, and similar frequencies of the states of x. So the differences in the results of the tests were due to the internal structure of the data, not to the prevalence of species with x = 1. The Yule model with covariates will tend to be accepted when shorter branches are associated with a character state. Such an association may be significant
even if a state is at low frequency. Another interesting property of the Yule model
with covariates revealed here is that it is robust to inexact estimation of ancestral
states. This is likely to be important giving difficulties in reconstructing ancestral
states (see below).

The simulations strongly suggest that the methods considered here are 265 statistically consistent: they had greater power to reject the null hypothesis when the 266 difference between the parameters were larger. Mooers and Schluter (1999) suggested 267 that most data sets are too small to yield enough statistical power to reject model 1. 268 Interestingly, most data sets they considered have many fewer than 100 species. With 269 the increasing size of biological databases, and the advent of supertree techniques 270 (e.g., Agnarsson et al. 2006; Beck et al. 2006; Bininda-Emonds et al. 2007), we are 271 likely to have now the possibility to fit more complex models of character change. 272

A possible limitation of the current Markovian models of character change is the assumption of equilibrium (Nosil and Mooers 2005). When this is violated, and this is certainly true in real situations, then our methods will lose some power. However, the simulations presented here suggest that this is apparently a question of degree.

Though the method loses some power, it remains statistically consistent. In a more 277 general context, the statistical consistency of data analysis methods is important in 278 the face of confounding factors in realistic situations. The critical point is that the 279 methods can still detect significant effects even when their assumptions are not met. 280 An important confounding factor not considered here is extinction and the variation 281 of its rate. If extinction rate is related with a species character, then this will likely 282 result in biases for both methods considered here because this character will be 283 associated with long branches in the tree. With respect to the Yule model with 284 covariates, a previous simulation study showed that random extinction resulted in a 285 loss of statistical power but the method remained statistically consistent (Paradis 286 2005). 287

The present paper considered separate analyses of character change and diversification with currently available methods. From a statistical point of view, this

can be viewed as each analysis was conditioned on the other: the analysis of character 290 change was done assuming constant and homogeneous diversification, whereas the 291 analysis of diversification was done assuming given reconstruction of the ancestral 292 states. An interesting perspective would be to develop a joint analysis of these 293 processes. Since each analysis is done by maximum likelihood, it is possible to 294 combine both likelihood functions in a joint likelihood function that would be 295 maximized over all parameters (r and λ). It remains to be seen whether this would 296 increase the statistical performance of our methods. 297

Estimation of Ancestral States

298

The results of the analyses of character change show that the inferred state at the root 299 is likely to be misleading (see also Mooers and Schluter 1999, Fig. 1b). This failure to 300 estimate correctly the state at the root is due to the assumption of equilibrium 301 common to most Markovian models. A consequence of this assumption is that 302 whatever the initial state, the probability to be in any state is given by its relative 303 transition rate. For model 1, these probabilities are 0.5 since the transition rates are 304 equal. This is concretely visualized by calculating e^{tQ} whose values tend to 0.5 when t 305 increases. Consequently, if one state is rare this implies that the initial state (i.e., the 306 root) was in the other state, and so the system is in transition towards its equilibrium. 307 On the other hand, it was observed that the estimates of the rate of change were 308 nearly unbiased. This suggests that, while we are able to correctly assess how 309 frequently a character changed, our estimates of its ancestral states may not be 310 meaningful. This is an important issue because there is more interest in ancestral 311 states than on rates of change (e.g., Webster and Purvis 2002; Oakley 2003). This 312 clearly needs further study because if the relative poor performance of inferring 313 ancestral states is confirmed, this would require a revision of some previous results 314 and ideas. 315

It must be pointed out that the ancestral states were estimated assuming uniform priors on the root (i.e., it could *a priori* be in any state). This can be relaxed, for instance, by assuming that the prior probabilities of the root are equal to the observed proportions of the states (Maddison 2006). It will be interesting to examine
whether this has an effect on ancestral state estimates.

Conclusion

To conclude, I would echo Maddison's (2006) view that uncareful analyses of 322 evolutionary data can lead to wrong conclusions. Though the modern approach to 323 macroevolution has certainly some limitations, there are reasons to be optimistic. 324 Future data analyses should include several tools based on model fitting, hypothesis 325 testing, model checking, and parameter estimation. There is clearly a need to study 326 the statistical behavior of our models in a wide range of realistic situations. There is 327 also space for many methodological developments. One issue that still remains 328 challenging is how extinction can be considered efficiently in these methods. 329

330

321

Acknowledgments

I am grateful to Wayne Maddison for reading a previous version of this paper, to
Gemain Chen for clarifying some issues on the Wald test, and to three anonymous
referees and Arne Mooers for their constructive comments on the text. This research
was supported by a project SPIRALES from the Institut de Recherche pour le
Développement (IRD).

336

LITERATURE CITED

337 Agnarsson, I., L. Avilés, J. A. Coddington, and W. P. Maddison. 2006. Sociality in

theridiid spiders: repeated origins of an evolutionary dead end. Evolution

³³⁹ 60:2342–2351.

340 Agresti, A. 1990. Categorical Data Analysis. John Wiley & Sons, New York.

³⁴¹ Aldous, D. J. 2001. Stochastic models and descriptive statistics for phylogenetic

trees, from Yule to today. Statist. Sci. 16:23–34.

- 343 Beck, R. M. D., O. R. P. Bininda-Emonds, M. Cardillo, F. G. R. Liu, and A. Purvis.
- ³⁴⁴ 2006. A higher-level MRP supertree of placental mammals. BMC Evol. Biol. 6:93.

- Bininda-Emonds, O. R. P., M. Cardillo, K. E. Jones, R. D. E. MacPhee, R. M. D.
- Beck, R. Grenyer, S. A. Price, R. A. Vos, J. L. Gittleman, and A. Purvis. 2007.
- ³⁴⁷ The delayed rise of present-day mammals. Nature 446:507–512.
- ³⁴⁸ Ewens, W. J. and G. R. Grant. 2005. Statistical Methods in Bioinformatics: An
- ³⁴⁹ Introduction (Second Edition). Springer, New York.
- Futuyma, D. J. 1986. Evolutionary Biology (Second Edition). Sinauer Associates,
 Sunderland, MA.
- ³⁵² Futuyma, D. J. 1998. Evolutionary Biology (Third Edition). Sinauer Associates,
- 353 Sunderland, MA.
- Kendall, D. G. 1948. On the generalized "birth-and-death" process. Ann. Math. Stat.
 19:1–15.
- ³⁵⁶ Lindsey, J. K. 1996. Parametric Statistical Inference. Clarendon Press, Oxford.
- Maddison, W. P. 2006. Confounding asymmetries in evolutionary diversification and
 character change. Evolution 60:1743–1746.
- ³⁵⁹ McCulloch, C. E. and S. R. Searle. 2001. Generalized, Linear, and Mixed Models.
- John Wiley & Sons, New York.
- Mooers, A. Ø. and S. B. Heard. 1997. Evolutionary process from phylogenetic tree
- ³⁶² shape. Quart. Rev. Biol. 72:31–54.
- ³⁶³ Mooers, A. Ø. and D. Schluter. 1999. Reconstructing ancestor states with maximum
- likelihood: support for one- and two-rate models. Syst. Biol. 48:623–633.
- Nosil, P. and A. Ø. Mooers. 2005. Testing hypotheses about ecological specialization
 using phylogenetic trees. Evolution 59:2256–2263.
- Oakley, T. H. 2003. Maximum likelihood models of trait evolution. Comment. Theor.
 Biol. 8:1–17.
- ³⁶⁹ Pagel, M. 1994. Detecting correlated evolution on phylogenies: a general method for
- the comparative analysis of discrete characters. Proc. R. Soc. Lond. B 255:37–45.
- ³⁷¹ Pagel, M. 1999. Inferring the historical patterns of biological evolution. Nature
- 401:877-884.
- Paradis, E. 2005. Statistical analysis of diversification with species traits. Evolution
 59:1–12.

- Paradis, E., J. Claude, and K. Strimmer. 2004. APE: analyses of phylogenetics and 375 evolution in R language. Bioinformatics 20:289–290. 376
- Pybus, O. G. and P. H. Harvey. 2000. Testing macro-evolutionary models using 377
- incomplete molecular phylogenies. Proc. R. Soc. Lond. B 267:2267-2272. 378
- R Development Core Team. 2006. R: A Language and Environment for Statistical 379
- Computing. R Foundation for Statistical Computing, Vienna, Austria. URL: 380
- http://www.R-project.org. 381
- Rao, C. R. 1973. Linear Statistical Inference and its Applications (Second Edition). 382 Wiley, New York.
- Sanderson, M. J. and M. J. Donoghue. 1996. Reconstructing shifts in diversification 384 rates on phylogenetic trees. Trends Ecol. Evol. 11:15–20. 385
- Schnabel, R. B., J. E. Koontz, and B. E. Weiss. 1985. A modular system of algorithms 386 for unconstrained minimization. ACM Trans. Math. Software 11:419–440. 387
- Simpson, G. G. 1953. The Major Features of Evolution. Columbia University Press, 388
- New York. 389

383

393

- Webster, A. J. and A. Purvis. 2002. Testing the accuracy of methods for 390
- reconstructing ancestral states of continuous characters. Proc. R. Soc. Lond. B 391
- 269:143-149.392

Corresponding Editor: A. Mooers

TABLE 1. Proportions of cases where the null hypothesis was rejected (200 replications). A star indicates where the tested null hypothesis was actually true. The columns labeled 'Both' give the proportions of cases where the likelihood-ratio test (LRT) and the Wald test agreed on rejecting the null hypothesis.

		Character change			Diversification		
p_1	λ_1	LRT	Wald	Both	LRT	Wald	Both
p_0	$5\lambda_0$	0.18*	0.265*	0.055^{*}	0.415	0.53	0.415
$0.1p_{0}$	λ_0	0.405	0.31	0.26	0.01*	0.01*	0.01*
$0.25p_{0}$	$2.5\lambda_0$	0.175	0.235	0.1	0.13	0.23	0.13

FIG. 1. Distribution summaries of the estimates of the rates of character change and their standard-errors. The dots indicate the median, and the error-bars the first and third quartiles. The panels correspond to the three simulated scenarios.

- ³⁹⁷ FIG. 2. For each simulated data set, the relative likelihood that the root was in state
- x = 0 was calculated under both models of character change. This series of
- ³⁹⁹ histograms shows the distribution of these values for each combination of parameters
- (as rows) and each actual initial state of the root (as pairs of columns). A value close
- to zero implies that the root state was inferred to be x = 1, a value close to one
- implies it was x = 0, and a value close to 0.5 implies that the inference of the root
- ⁴⁰³ state was indeterminate by this model.



Fig. 1.



Fig. 2.