



HAL
open science

A model with many small shifts for estimating species-specific diversification rates

Odile Maliet, Florian Hartig, H el ene Morlon

► **To cite this version:**

Odile Maliet, Florian Hartig, H el ene Morlon. A model with many small shifts for estimating species-specific diversification rates. *Nature Ecology & Evolution*, 2019, 3 (7), pp.1086-1092. 10.1038/s41559-019-0908-0 . hal-02408107

HAL Id: hal-02408107

<https://hal.science/hal-02408107>

Submitted on 12 Dec 2019

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 destin ee au d ep ot et  a la diffusion de documents scientifiques de niveau recherche, publi es ou non,  emanant des  tablissements d'enseignement et de recherche fran ais ou  trangers, des laboratoires publics ou priv es.

A model with many small shifts for estimating species-specific diversification rates

Odile Maliet^{a,*}, Florian Hartig^b, H el ene Morlon^a

^a Institut de biologie de l'Ecole normale sup erieure (IBENS), Ecole normale sup erieure, CNRS, INSERM, PSL Research University, 75005 Paris, France

^b Theoretical Ecology, Faculty of Biology and Preclinical Medicine, University of Regensburg, Universit atsstra e 31, Regensburg, Germany

* Corresponding author. email: odile.maliet@orange.fr

Abstract

1 Understanding how and why diversification rates vary through time, space, and across species groups is
2 key to understanding the emergence of today's biodiversity. Phylogenetic approaches aimed at identifying
3 variations in diversification rates during the evolutionary history of clades have focused on exceptional shifts
4 subtending evolutionary radiations. While such shifts have undoubtedly affected the history of life (1),
5 identifying smaller but more frequent changes is important as well. We develop ClaDS, a new Bayesian
6 approach for estimating branch-specific diversification rates on a phylogeny, that relies on a model with
7 changes in diversification rates at each speciation event. We show using Monte-Carlo simulations that the
8 approach performs well at inferring both small and large changes in diversification. Applying our approach
9 to bird phylogenies covering the entire avian radiation, we find that diversification rates are remarkably
10 heterogeneous within evolutionary restricted species groups. Some groups such as *Accipitridae* (hawks and
11 allies) cover almost the full range of speciation rates found across the entire bird radiation. As much as 76%
12 of the variation in branch-specific rates across this radiation is due to intra-clade variation, suggesting that
13 a large part of the variation in diversification rates is due to many small rather than few large shifts.

Manuscript

14
15 Several phylogenetic approaches have been developed for understanding when and on which
16 lineages diversification rates have changed during the evolutionary history of clades (2; 1; 3; 4; 5; 6).
17 Most have focused on 'major' rate shifts, which is convenient methodologically; likelihoods of trees
18 under such models have been used for some time (7). These models correspond to the idea that
19 few rare events, such as key innovations, facilitate the invasion of new adaptive zones, with a
20 drastic impact on diversification rates (8; 9). In these models, outside of few remarkable events,
21 diversification rates are assumed to be homogeneous. However, while major rate shifts linked to key
22 innovations have undoubtedly affected the history of life (1), they are not the only – nor necessarily
23 the most important – source of variation in diversification rates.

24 Shifts in diversification rates are likely quite widespread. Speciation and extinction rates may
25 vary across lineages as a response to the particular biotic and abiotic environment experienced by
26 each lineage (10); they may also vary as a response to traits that affect reproductive isolation such
27 as reproduction mode (11) or pollination and dispersal syndromes (12). Such changes in diversi-
28 fication rates probably occur far more frequently than key innovations, resulting in heterogeneous

29 diversification rates at much finer taxonomic scales (4). Accounting for such finer scale hetero-
30 geneity is crucial if we want to obtain refined estimates of lineage-specific diversification rates and
31 to better understand the processes subtending heterogeneity in the diversification of life. Meth-
32 ods of the State-Speciation-Extinction family (13) can in principle better account for these types of
33 heterogeneities, but they require assuming trait-dependency of rates (SI Appendix section 3.5). Non-
34 model-based approaches such as the DR statistic (4) can also account for fine-scale heterogeneities,
35 but they are rather *ad hoc* and generally do not perform as well as model-based approaches (14).

36 Here, we develop a new Bayesian approach (ClaDS) for estimating lineage-specific diversification
37 rates on a phylogeny that better accounts for the diverse sources of variation in diversification rates
38 that occur during the evolutionary history of clades. Using Monte Carlo simulations, we quantify the
39 ability of ClaDS to faithfully recover both small and large changes in diversification rates. Finally,
40 we apply the method to time-calibrated phylogenies for 42 bird clades to evaluate the extent to
41 which differences in the pace of diversification across the entire avian radiation result from few large
42 versus many small events.

43 A new model of diversification rate variation

44 We consider a birth-death diversification process, the cladogenetic diversification rate shift
45 (ClaDS) model, where diversification rates are inherited at speciation, but with a shift (Fig. 1).
46 At the beginning of the process, the clade is composed of one lineage with speciation rate λ_0 and
47 extinction rate μ_0 . At each speciation event, the two daughter lineages inherit new diversification
48 rates $(\lambda_{i1}, \lambda_{i2})$ and (μ_{i1}, μ_{i2}) sampled from a joint probability distribution ν parameterized by
49 the parental rates λ_i and μ_i . If the change in speciation and extinction rates are assumed to be
50 independent, the λ_i are sampled from a distribution ν_λ , the μ_i are sampled from a distribution ν_μ ,
51 and $\nu = \nu_\lambda \times \nu_\mu$. Moreover, we allow for the possibility that some extant species are missing by
52 assuming that each extant species is observed with probability $f \leq 1$. We derive the probability
53 density of a reconstructed phylogeny under this general model and implement its computation in R
54 (Materials and Methods & SI Appendix sections 2 to 5).

55 We then consider several scenarios in ClaDS where: i) ν_λ is a lognormal distribution with
56 parameters $\log(\alpha * \lambda)$ and σ ; the latter ensures that the relative change in rate at speciation λ_i/λ is
57 independent from the parental rate with a mean m given by $\alpha \exp(\sigma^2/2)$; σ controls how constrained
58 daughter rates are (highly constrained for small σ values) and α controls the trend at speciation (i.e.
59 whether daughter rates tend to be higher or lower than parental rates) ii) extinction rates are either
60 negligible ($\mu_i = 0$ for all lineages, ClaDS0), homogeneous across all lineages in the clade ($\mu_i = \mu_0$
61 for all lineages, ClaDS1) or vary across lineages, but with a constant turnover ε (i.e. $\mu_i/\lambda_i = \varepsilon$ for
62 all lineages, ClaDS2). We use Monte Carlo simulations under ClaDS1 and ClaDS2 (Materials and
63 Methods & SI Appendix section 6) to verify that our likelihood expression is correct (SI Appendix
64 section 6, Fig. S6 to S8). Finally, we implement a Monte Carlo Markov Chain (MCMC) sampler
65 that, given a reconstructed phylogeny, simultaneously estimates both the parameters of ClaDS (λ_0 ,
66 α , σ , and either μ_0 or ε) and the speciation rates λ_i at the origin of each branch i of the phylogeny
67 (Materials and Methods & SI Appendix section 7, see also section 8 and Fig. S9 & S10 for a test
68 of the sampler). Branch-specific extinction rates μ_i at the origin of each branch i of the phylogeny
69 are given by μ_0 for ClaDS1 and by $\varepsilon * \lambda_i$ for ClaDS2. In what follows for simplicity we refer to λ_i
70 and μ_i as “branch-specific rates” instead of the more accurate “rates at the origin of each branch”.

71 Under these scenarios of the ClaDS process, heterogeneity in speciation rates across lineages
72 is determined on the one hand by a stochastic component (controlled by σ), and on the other

73 hand by a trend component (controlled by m). When the expected daughter rate is equal to the
74 parental rate ($m = 1$), the resulting trees are relatively imbalanced and tippy (SI Appendix section
75 1, Fig. S1 & S2): lineages that by chance have high speciation rates early in clade’s history spread,
76 leading to rates that are heterogeneous across lineages and average rates that increase through time.
77 This sorting effect is exacerbated when the expected daughter rate is higher than the parental rate
78 ($m > 1$, Fig. S1 & S2), corresponding to a ‘niche-piling’ scenario where diversity begets diversity
79 (15). To the contrary, when the expected daughter rate is lower than the parental rate ($m < 1$),
80 corresponding to a ‘niche-filling’ scenario where diversification gets harder as new species arise
81 (16; 17; 18), the heterogeneity in speciation rates across lineages is reduced, and with a low enough
82 m , the average rate is constant or even decreasing through time (Fig. S1 & S2). Importantly, ClaDS
83 is able to produce the combination of stemmy and imbalanced tree shapes observed in nature, and
84 under a wider set of parameter values for the scenario with constant turnover (ClaDS2) than the
85 scenario with constant extinction rate (ClaDS1, Fig. S1 & S2).

86 Performance of ClaDS

87 We begin by testing the performance of ClaDS under frequent rate changes and in the absence
88 of extinction (ClaDS0) (Materials and Methods). We find that the approach provides unbiased
89 estimates of all model’s parameters for large enough trees (size 200, Fig. 2); the relative change
90 in rate at speciation m is also well estimated (Fig. 2 d). As expected, bias and variability around
91 parameter estimates increase for smaller trees (Fig. S11 to S14).

92 ClaDS provides reliable estimates of branch specific speciation rates on average: while low rates
93 tend to be slightly overestimated and large rates slightly underestimated, ClaDS can detect regions
94 of the tree with relatively high or low rates (Fig. 3 & Fig. S15 & S16).

95 When considering also extinctions, focusing on the scenario with constant turnover (ClaDS2)
96 as it generally produced tree shapes closer to those observed in nature, we found that estimates
97 remain accurate at low levels of extinction ($\varepsilon = 0.1$) for both model parameters (Fig. S20) and
98 branch-specific speciation rates (Fig. S21). At high levels of extinction ($\varepsilon = 0.9$), σ and, when
99 the mean change in rate at speciation m approaches 1, branch-specific speciation rates, remain well
100 estimated. It is not the case, however, of the turnover rate ε , α , and branch-specific speciation
101 rates when $m < 1$, although accounting for extinction does improve inferences over ignoring it (Fig.
102 S20 & S21). When extinction is not accounted for, estimated branch-specific speciation rates are
103 generally lower than realized ones, but higher than realized net diversification rates (Fig. S21C &
104 D).

105 If there are a small number of major rate shifts during the evolution of clades, rather than many
106 small changes (tested here with a single rate shift, Materials and Methods), ClaDS is still able to
107 provide reliable estimates of branch-specific rates (Fig. S17 & S19). The model is also able to
108 detect when two branches in the tree belong to distinct speciation regimes as soon as the difference
109 in rates between the two regimes is large enough (a two-fold increase or decrease in our simulations)
110 and both regimes are represented by a large enough number of branches in the phylogeny (Fig. S19
111 left). The false detection rate associated to this test is low (Fig. S19 right).

112 Finally, when comparing the performance of ClaDS to that of two other popular methods for
113 estimating branch-specific rates (the DR statistic and BAMM (4; 5)) under various simulation
114 schemes (SI Appendix section 9), we find, overall, that ClaDS outperforms the other methods for
115 trees simulated with both many small shifts at speciation (Fig S22 & S26 to S28) and gradual
116 changes along branches (Fig S24), and that it performs as well as other methods for trees simulated

117 with few large shifts (Fig. S23) and variations in extinction rates (Fig. S25). Importantly, ClaDS
118 provides reliable estimates of the variance in rates under both the many small and the few large
119 shifts scenarios (Fig. S30 & S32), while BAMM underestimates the variance in rates under the
120 many small shifts scenario (Fig. S31). ClaDS and BAMM provide low (and similar) estimates of
121 the variance in rates for trees simulated under constant rates (Fig. S29); in the presence of rate
122 heterogeneity, they tend to underestimate rather than overestimate rate variance (Fig. S30 to S33),
123 and BAMM more so than ClaDS.

124 Diversification across the avian radiation

125 When applying ClaDS to major bird clades (Materials and Methods), we found that lineage-
126 specific speciation rates can vary by as much as 2 orders of magnitude within clades (Fig. 4e). In
127 *Accipitridae* (hawks and allies) for example, speciation rates range from 0.013 to 1.2 Mya⁻¹, which
128 almost covers the range found across the entire avian radiation (0.013 – 5 Mya⁻¹). Comparable
129 within-clade heterogeneities occur in other clades, such as *Muscicapidae* & *Turdidae*, *Tyrannidae*
130 and *Parulidae* (Fig. 4e, in orange). Such within-clade heterogeneities are way above heterogeneities
131 arising from estimation error (Fig. S29). A variance partitioning of speciation rates across the bird
132 radiation (Material and Methods) reveals that intra-clade variance accounts for 76% of the total
133 variance. In comparison, BAMM would have estimated much less within-clade heterogeneities,
134 with an intra-clade variance accounting for only 46% of the total variance (Fig. S34). Given our
135 simulation results, this suggests that BAMM underestimates the intra-clade variance, and thus that
136 many small shifts occurred during bird diversification that BAMM cannot detect.

137 While some clades have very heterogeneous rates, others are quite homogeneous, such as *Ram-*
138 *phastides*, *Alcedinidae*, *Charadrii* and *Phasianidae* (Fig. 4e, in blue). We did not find any signifi-
139 cant relationship between the variance in rate values within a clade and the size ($p = 0.49$) or age
140 ($p = 0.93$) of the clade, indicating that rate heterogeneity is not a mere result of time or species
141 richness; rather, rates are pretty constrained in some old and rich clades (e.g. *Phasianidae*) as well
142 as in some younger or less species-rich clades (e.g. *Alcedinidae*), while they can take very different
143 values for distinct species of both old or young clades (e.g. *Parulidae*, *Tyrannidae*). The wide range
144 of σ estimates found across bird clades (Fig. 4a), in comparison with rather tight α and m estimates
145 (Fig. 4b & c), suggests that differences in rate heterogeneity across clades are due to the stochastic
146 component of the model, rather than its trend component. Indeed, α ranges between 0.38 and 1.02
147 (with a mean of 0.71, Fig. 4b), which indicates a universal tendency for daughter rates to be smaller
148 than ancestral ones, with a decline that is comparable in magnitude across clades. There is only
149 one case when m is clearly above 1 (1.12 in *Campephagidae*); this corresponds to a case when most
150 shifts correspond to rate declines, but the few shifts that correspond to rate increases are much
151 bigger in magnitude.

152 Discussion

153 Models of diversification applied to phylogenies of extant taxa are increasingly used to under-
154 stand the long-term evolution of biodiversity. These approaches have highlighted how much variable
155 diversification rates can be across the tree of life, and the importance of these variations for explain-
156 ing current patterns of diversity (the so-called ‘diversification rate hypothesis’ (19)). Yet, despite
157 recent advances in phylogenetic approaches for understanding diversification, detecting diversifica-
158 tion rate variations and the processes underlying these variations remain a challenge spurring a

159 heated debate (20; 21; 22; 23; 24). In this paper, we have developed ClaDS, a new model with
160 frequent small variations in diversification rates together with a method to infer branch-specific
161 diversification rates on a phylogeny. We have shown using simulations that ClaDS accurately es-
162 timates branch-specific rates. Finally, applying ClaDS to the bird phylogeny, we have shown that
163 small but frequent changes have been instrumental in shaping global rate variation during the avian
164 radiation.

165 One of the major advances of our model is to rely on an explicit and exact computation of
166 the likelihood in the presence of extinction. Previous likelihood expressions under diversification
167 models with variable rates were computed with the underlying assumption that shifts do not occur in
168 extinct lineages (1; 3; 5), except in the case of trait-dependent models (see SI Appendix section 3.5 for
169 further discussion); this is biologically implausible and can introduce an important bias depending on
170 the intensity of extinction (22; 23). In ClaDS we relax this inconvenient assumption by integrating
171 appropriate Ordinary Differential Equations (ODEs, SI Appendix section 3). This allows computing
172 likelihoods accounting for rate shifts on extinct lineages, which has so far only been done through
173 intense and impractical Monte Carlo simulations (22). The ODE integration is computationally
174 intensive, but not as much as to prevent running ClaDS on reasonably sized trees, as we illustrated
175 on the bird phylogenies. Despite this significant improvement, our simulations show that estimating
176 extinction remains difficult, in line with the well-known difficulty of estimating extinction from
177 phylogenies of only extant taxa (25). This is true even when simulations and inferences are performed
178 under simple models with constant extinction or turnover rate. Despite difficulties in estimating
179 extinction rates, properly accounting for extinctions in the likelihood computation is satisfying on
180 a biological and theoretical standpoint, and, as we have shown, improves the estimation of both
181 model parameters and branch specific speciation rates.

182 Another advantage of ClaDS is to avoid using model selection to select the number and location
183 of rate shifts, by assuming that shifts happen at each speciation event. In the frequently-used
184 MEDUSA method (1), stepwise AIC is used to perform this selection, with associated statistical
185 limitations (21). In the approach of Morlon et al. (3), likelihood ratio tests are performed to
186 select the number of shifts, but the location of these shifts needs to be fixed *a priori*. Finally, in
187 the popular Bayesian analysis of macroevolutionary mixtures (BAMM, 5), reversible jump mcmc
188 is used, with a prior on the number and location of shifts that may influence the results (22; 26).
189 ClaDS avoids these limitations, while still performing well in the presence of rare rate shifts with
190 large effects.

191 Maybe more importantly than these technical aspects, ClaDS represents a view of evolution
192 distinct from that of previous models: existing models focus on a small number of discrete diversifi-
193 cation shift events spread across the tree, an idea that fits well with the concept of key innovations
194 driving major diversification shifts (3; 1; 5); to the contrary, ClaDS allows for frequent variations
195 linked, for example, to changes in environmental conditions or associations with continuously evol-
196 ving heritable traits. Accordingly, ClaDS does not aim at identifying specific nodes in a phylogeny
197 subtending major diversification rate shifts. Rather, it assumes that rate shifts happen at each spe-
198 ciation event and focuses on estimating branch-specific diversification rates. In nature, both many
199 shifts with small effects and few shifts with large effects are likely to occur, and so it is reassuring to
200 see that ClaDS can properly estimate branch specific rates under these two evolutionary processes.

201 Accurately estimating branch specific diversification rates is a critical step for understanding
202 the processes that lead some species groups to diversify faster than others. For example, species'
203 traits can modulate their propensity to diversify, and tests based on assessing the correlation between

204 trait values at a phylogenies' tips and metrics capturing the diversification rate of the corresponding
205 lineages ('tip-rate correlations' tests) have been developed to detect such effects (27). These types
206 of tests have regained interest lately (see e.g. STRAPP (20), FiSSE (28), ES-sim (29), pNoTO
207 (30; 31)), as an alternative or complement to state-dependent speciation-extinction (SSE) methods
208 that jointly model diversification dynamics and trait evolution (32; 13). However, current metrics
209 of species-level diversification rates have limitations. Some of them are derived from BAMM (5)
210 and thus reflect a limited set of diversification rate regimes rather than lineage-specific rates *per se*.
211 Others are summary statistics describing phylogenetic branching patterns, such as the "node density"
212 (27), the "equal split" (33), or the "diversification rate" (4) statistics; they are not rigorously derived
213 from speciation-extinction models, and they generally perform worse than model-based approaches
214 (14) (SI Fig. S22 to S27). ClaDS provides tip level estimates of diversification rates that should
215 help identifying the specific features of a species that make it more or less prone to diversify. In
216 the future, we could imagine a hybrid between SSE and ClaDS that would account for both trait-
217 dependent diversification and residual rate variation not accounted for by the trait, in the spirit
218 of hidden states models (HiSSE (34), MSBD (6)). This could for example be done by imputing in
219 ClaDS specific trend parameters α corresponding to trait shifts.

220 Changes in biotic and abiotic conditions can also modulate the tempo of diversification, leading
221 diversification to be faster during some time periods than others. ClaDS accommodates temporal
222 trends in rate variation, without the need to specify a specific form for this variation *a priori* as
223 in time-dependent diversification models (35; 16; 3), and with more flexibility than models where
224 a discrete rate shift at a given time point affects the whole clade (36). In the future, the trend
225 parameter α could depend on measured environmental variables; this would allow directly test-
226 ing for an effect of these environmental variables on diversification, as in environment-dependent
227 diversification models (37; 38), while accounting for residual rate variation.

228 Our ClaDS analysis of the avian radiation reveals a series of compelling results. First, and even
229 though these estimates need to be taken with caution, we find significant (non-zero) turnover rates.
230 Second, we find a pervasive pattern of declines in speciation rates over time congruent with previous
231 studies (16; 17; 18). Third, we find a remarkable heterogeneity in speciation rates, with per-lineage
232 rates that vary by two orders of magnitude ($0.01 - 5 \text{ Mya}^{-1}$), peaking around 0.15 Mya^{-1} . Fourth,
233 we find that variability in speciation rates can be as high within than between clades, suggesting that
234 rate variation may be much more widespread than currently thought and implemented in existing
235 models. Finally, we highlight a remarkable difference across clades in terms of how constrained
236 their diversification rates are, with plovers and allies on one extreme, and hawks and allies on the
237 other extreme of a continuum between rates that vary less than 2 fold to more than 80 folds (Fig.
238 4e, f). These differences in how constrained diversification rates are striking and remain to be
239 explained: these could be linked to differences in genetic architecture, developmental constraints,
240 or biogeographies, for example.

241 Together, our results refute the idea that speciation may be clock-like (39) and emphasize the
242 need to consider diversification models that embrace the pervasive heterogeneity of the evolutionary
243 process. Further, they promise a bright future for approaches, such as ours, that relax the speciation
244 clock similarly to the way the molecular clock has been relaxed (40; 41; 42): similar to molecular
245 rates, diversification rates vary according to many small shifts.

246

Material and Methods

247 Likelihood, simulation and Bayesian implementation of ClaDS

248 **Likelihood** We derived the probability density of observing a reconstructed phylogeny with
249 branches delimited by the times $(t_i, s_i)_{i \in \llbracket 1, N \rrbracket}$ and speciation and extinction rates λ_i and μ_i at
250 time t_i (i.e. at the origin of each branch) under the cladogenetic diversification rate shift model
251 (SI Appendix section 2 to 4). We note Θ the parameters of the new rate distribution ν . The
252 probability density can be derived from three main probability functions: $\Phi_{\Theta, \lambda, \mu}(t)$, the probability
253 that a lineage alive at time t has speciation and extinction rates λ and μ and no descendant in
254 the reconstructed phylogeny; $\chi_{\Theta, \lambda, \mu}(t)$, the probability that a lineage alive at time t has speciation
255 and extinction rates λ and μ and exactly one descendant species sampled in the reconstructed
256 phylogeny; and $\xi_{\Theta, \lambda, \mu}(t, s, \lambda_1, \lambda_2, \mu_1, \mu_2)$, the probability that a lineage alive at time t has speciation
257 and extinction rates λ and μ and gives birth at time s to two daughter lineages that respectively
258 have speciation rates λ_1 and λ_2 and extinction rates μ_1 and μ_2 . We obtained ordinary differential
259 equations (ODEs) to solve for Φ , χ and ξ by considering the different events that can happen during a
260 short time interval Δ_t and making Δ_t tend to 0 (SI Appendix section 3.1 to 3.3). Under a pure-birth
261 model and for a completely sampled phylogeny, the ODEs can be solved analytically (SI Appendix
262 section 4). In the presence of extinction and/or if there are missing taxa in the phylogeny, Φ , χ and
263 ξ are computed by integrating the ODEs numerically, which is more computationally intensive (SI
264 Appendix section 5).

265 **Simulation** We implemented a simulation algorithm of ClaDS in the R-package RPANDA
266 (43, function `sim_ClaDS`) (SI Appendix section 1). In this implementation, the speciation rates of
267 daughter lineages are drawn independently from a distribution ν_λ . Their extinction rates are either
268 drawn from a distribution ν_μ , given by μ_0 (constant extinction rate scenario, ClaDS1), or given by
269 $\varepsilon * \lambda_{s_{i,1}}$ and $\varepsilon * \lambda_{s_{i,2}}$ (constant turnover scenario, ClaDS2). ν_λ and ν_μ can be normal, log-normal,
270 or uniform distributions. The simulations are continued until a stopping criterion is met, either a
271 fixed time or a fixed number of species. In addition, `sim_ClaDS` takes as one of its arguments a
272 parameter p controlling the probability that a shift happens at each speciation event (the default
273 value $p = 1$ corresponds to the model investigated here), and a parameter n , controlling a maximum
274 number of shifts (the default value $n = \text{INF}$ corresponds to the model investigated here; if n takes
275 a finite value, then p switches to 0 as soon as n switches have occurred).

276 **Bayesian implementation** We implemented a Bayesian inference approach for fitting ClaDS
277 to reconstructed phylogenies in the R-package RPANDA (43, function `fit_ClaDS`) (SI Appendix 1
278 section 7). In order to fit ClaDS0 (no extinction), we use a Metropolis within Gibbs MCMC (Monte
279 Carlo Markov Chain) sampler with a Bactrian proposal (44), and convergence is monitored by
280 running three MCMC chains in parallel and computing Gelman statistics (45). In order to fit
281 ClaDS1 and ClaDS2 (i.e. in the presence of extinction), and/or if there are missing taxa in the
282 phylogeny, we use the faster blocked Differential Evolution (DE) MCMC sampler, with sampling
283 from the past of the chains (46). We also ran three chains. For both with and without extinction,
284 we use an inverse gamma prior with shape parameter 1 and rate parameter 0.1 for σ and a flat
285 prior for all other parameters. Each estimate was computed as the mean over the iterations and
286 the three chains.

287 **Testing the performance of ClaDS**

288 We performed intensive simulations to test the performance of ClaDS. We tested both the
289 performance of ClaDS under data generated by this model, and its performance for data generated
290 with a discrete speciation rate shift. In order to assess the performance of ClaDS under a large
291 parameter set and for a variety of tree sizes, we considered primarily the pure birth model with
292 completely sampled phylogenies. We also considered the model with extinction and/or missing taxa,
293 but only in a limited, computationally tractable, set of simulations.

294 **Many small rate shifts (ClaDS model)** For each combination of the following parameter
295 values, we simulated 20 pure birth trees, stopping the simulation when a target tip number of 50,
296 100 and 200 was reached. λ_0 was fixed at 0.1, σ was taken in $\{0, 0.1, 0.18, 0.26, 0.34, 0.41\}$, and α in
297 $\{1.2, 1.1, 1, 0.95, 0.9, 0.7\}$. We recorded the realized speciation rate on each branch in each of these
298 simulations. We then ran ClaDS on each simulated tree using our `run_ClaDS0` function. Lastly,
299 we compared the retrieved estimates of λ_0 , σ and α to their simulated values; we also compared
300 the retrieved estimates of branch-specific speciation rates for each tree to their realized values by
301 performing linear regressions and computing relative errors (ratio of estimated versus realized rates).

302 In order to explore the model accounting for extinction, we simulated 5 trees of size 100 under
303 4 scenarios with constant turnover rate (ClaDS2), and for each condition either low ($\varepsilon = 0.1$) or
304 high ($\varepsilon = 0.9$) turnover (8 scenarios in total). We focused on the scenario with constant turnover,
305 because this scenario produced tree shapes similar to those of empirical trees under a wider set
306 of parameter values than the alternative scenario with constant extinction rate (Fig. S2.1 versus
307 S2.2). Maintaining a balance where extinction is neither negligible nor driving clades to extinction is
308 also easier under ClaDS2. The four scenarios were as follows: i) high heterogeneity and decreasing
309 rates : $\lambda_0 = 0.1$, $\sigma = 0.7$, $\alpha = 0.7$ (mean relative change $m = 0.9$), ii) no heterogeneity and
310 constant rates (equivalent to constant rate birth-death trees) : $\lambda_0 = 0.1$, $\sigma = 0$, $\alpha = 1$ ($m = 1$)
311 iii) Low heterogeneity and no average change in rate at speciation : $\lambda_0 = 0.1$, $\sigma = 0.2$, $\alpha = 0.98$
312 ($m = 1$) iv) Low heterogeneity and decreasing rates : $\lambda_0 = 0.1$, $\sigma = 0.2$, $\alpha = 0.88$ ($m = 0.9$). We
313 recorded the realized speciation rate at the beginning of each branch in each of these simulations.
314 We then ran ClaDS on each simulated tree using our R function, both accounting (`run_ClaDS`)
315 and not accounting (`run_ClaDS0`) for extinction, the latter to evaluate the bias resulting from not
316 accounting for extinction when it occurs. Lastly, we compared the retrieved estimates of σ , α , m
317 and ε for each tree to their simulated values. We did not compare the retrieved estimates of λ_0 to
318 the simulated values, because the estimates correspond to the speciation rate at the crown while
319 the simulated values correspond to the speciation rate at the stem. These two rates can be very
320 different in the presence of extinction. We also compared the retrieved estimates of branch-specific
321 speciation rates and net diversification rates (speciation minus extinction) for each tree to their
322 realized values by performing linear regressions and computing relative errors.

323 **Few large rate shifts** We also tested the behavior of ClaDS under a ‘key innovation’ scenario
324 with only a single large rate shift during the history of the clade. In order to simulate this scenario,
325 we used our `sim_ClaDS` function with λ_0 (the background rate in this case) fixed at 0.1, p (the
326 probability that a rate shift happens at each speciation event) fixed at 0.02, and n (the maximum
327 number of shifts) fixed at 1. The new speciation rate took a series of values from lower (uniformly
328 drawn in $[0.025, 0.03]$, $[0.03, 0.05]$, $[0.05, 0.1]$) to higher (uniformly drawn in $[0.1, 0.15]$, $[0.15, 0.2]$,
329 $[0.2, 0.3]$, $[0.3, 0.4]$, $[0.4, 1]$) than the background rate. For each of these rate values, we simulated

330 phylogenies of size 200 until we had a good coverage of subclade new rate/size combination (from
331 300 to 500 phylogenies per parameter set). In such simulations, there are only two distinct rates
332 across the tree: the background rate and the new rate. We then ran ClaDS on each simulated tree
333 using our run_ClaDS0 function and compared the retrieved estimates of branch-specific speciation
334 rates for each tree to their simulated values by performing linear regressions and computing relative
335 errors. Finally, we tested whether the model is able to detect if two branches in the tree belong to the
336 same or distinct speciation regime(s): two branches were considered to have significantly different
337 rates (distinct regime) if the difference in the estimated speciation rates between the two branches
338 was of constant sign on at least 95% of the MCMC chains. We assessed the significance of speciation
339 rate differences (and the corresponding sign) for all pairs of branches in the simulated trees. Finally,
340 we quantified the ‘proper detection’ rate as the proportion of pairs for which a significant difference
341 was inferred when the two branches indeed belonged to distinct speciation regimes (i.e. one had
342 the background speciation rate and the other one had the new rate), and the ‘false detection’ rate
343 as the proportion of pairs for which a significant difference was inferred, while the two branches
344 actually belonged to the same speciation regime (i.e. both had either the background speciation
345 rate or the new rate).

346 **Diversification of the avian radiation**

347 We applied ClaDS, accounting for extinction (ClaDS2, model with constant turnover) and in-
348 complete sampling, to bird phylogenies. We used the MCC trees from Jetz et al. (4) with only the
349 species for which there was molecular data, along with the associated sampling fractions provided
350 by the authors. Most of these are family level phylogenies, with some spawning two or a few more
351 families. We ran the model on the 42 bird phylogenies with more than 50 species. We report the
352 distribution of branch-specific speciation rates across the 42 clades, as well as individual distri-
353 butions for each clade. We partitioned the total variance of the logarithm of the branch specific
354 speciation rates $(\sum_i (\ln(\lambda_i) - \overline{\ln(\lambda)})^2$, where $\overline{\ln(\lambda)}$ is the mean of the log of the speciation rates for
355 all branches in all clades) between the intra-clade $(\sum_i (\ln(\lambda_i) - \overline{\ln(\lambda_{c_i})})^2$, where c_i is the clade to
356 which branch i belongs and $\overline{\ln(\lambda_{c_i})}$ is the mean of the log of the speciation rates for all branches in
357 clade c) and inter-clade variance $(\sum_i (\overline{\ln(\lambda_{c_i})} - \overline{\ln(\lambda)})^2$). We also tested for a potential correlation
358 between the variance in rates and the size (number of tips) and age (crown age) of clades using
359 PGLS (47) (two-sided test) on the Hacket backbone phylogeny provided in Jetz et al. (4).

360 **Data availability**

361 The simulated phylogenies used to test the method are available at <https://github.com/OdileMaliet/ClaDS/tree/>
362 in the file named trees.zip. All the empirical data used for the analysis were obtained from Jetz et
363 al. (2012) study, and are available on <https://www.nature.com/articles/nature11631>.

364 **Code availability**

365 The R functions used to simulate and fit the model are available in the RPANDA R-package.
366 All the codes used to test our method are available on the github repository [https://github.com/](https://github.com/OdileMaliet/ClaDS.git)
367 [OdileMaliet/ClaDS.git](https://github.com/OdileMaliet/ClaDS.git).

368 **Competing interests**

369 The authors declare no competing interests.

370 **Acknowledgements**

371 The authors are very grateful to Leandro Aristide, Julien Clavel, Jonathan Drury, Carmelo
372 Fruciano, Sophia Lambert, Eric Lewitus, Marc Manceau, Olivier Missa, Benoît Perez, Ana Cata-
373 rina Silva and Guilhem Sommeria-Klein for their helpful comments on an earlier version of this
374 manuscript. This work was supported by an AMX grant (from Ecole Polytechnique) and the Labex
375 MemoLife to OM, PROCOPE mobility grant 57134817 to FH and HM, and the European Research
376 Council [ERC 616419-PANDA] to HM.

377 **References**

- 378 [1] Alfaro, M. E., Santini, F., Brock, C., Alamillo, H., Dornburg, A., Rabosky, D. L., Carnevale,
379 G., and Harmon, L. J. Nine exceptional radiations plus high turnover explain species diversity
380 in jawed vertebrates. *Proceedings of the National Academy of Sciences* **106**(32), 13410–13414
381 (2009).
- 382 [2] Chan, K. M. and Moore, B. R. SymmeTREE: whole-tree analysis of differential diversification
383 rates. *Bioinformatics* **21**(8), 1709–1710 (2004).
- 384 [3] Morlon, H., Parsons, T. L., and Plotkin, J. B. Reconciling molecular phylogenies with the fossil
385 record. *Proceedings of the National Academy of Sciences* **108**(39), 16327–16332 (2011).
- 386 [4] Jetz, W., Thomas, G., Joy, J., Hartmann, K., and Mooers, A. The global diversity of birds in
387 space and time. *Nature* **491**(7424), 444 (2012).
- 388 [5] Rabosky, D. L. Automatic detection of key innovations, rate shifts, and diversity-dependence
389 on phylogenetic trees. *PloS one* **9**(2), e89543 (2014).
- 390 [6] Barido-Sottani, J., Vaughan, T. G., and Stadler, T. A Multi-State Birth-Death model for
391 Bayesian inference of lineage-specific birth and death rates. *bioRxiv* (2018).
- 392 [7] Sanderson, M. J. and Wojciechowski, M. F. Diversification rates in a temperate legume clade:
393 are there so many species of *Astragalus* (Fabaceae)? *American Journal of Botany* **83**(11),
394 1488–1502 (1996).
- 395 [8] Miller, A. H. Some ecologic and morphologic considerations in the evolution of higher taxonomic
396 categories. *Ornithologie als biologische Wissenschaft* , 84–88 (1949).
- 397 [9] Hunter, J. P. Key innovations and the ecology of macroevolution. *Trends in ecology & evolution*
398 **13**(1), 31–36 (1998).
- 399 [10] Benton, M. J. The Red Queen and the Court Jester: species diversity and the role of biotic
400 and abiotic factors through time. *Science* **323**(5915), 728–732 (2009).
- 401 [11] Goldberg, E. E., Kohn, J. R., Lande, R., Robertson, K. A., Smith, S. A., and Igić, B. Species
402 selection maintains self-incompatibility. *Science* **330**(6003), 493–495 (2010).

- 403 [12] Onstein, R. E., Baker, W. J., Couvreur, T. L., Faurby, S., Svenning, J.-C., and Kissling, W. D.
404 Frugivory-related traits promote speciation of tropical palms. *Nature ecology & evolution* **1**(12),
405 1903 (2017).
- 406 [13] FitzJohn, R. G. Diversitree: comparative phylogenetic analyses of diversification in R. *Methods*
407 *in Ecology and Evolution* **3**(6), 1084–1092 (2012).
- 408 [14] Title, P. O. and Rabosky, D. L. Tip rates, phylogenies, and diversification: what are we
409 estimating, and how good are the estimates? *Methods in Ecology and Evolution* (2018).
- 410 [15] Emerson, B. C. and Kolm, N. Species diversity can drive speciation. *Nature* **434**(7036), 1015
411 (2005).
- 412 [16] Rabosky, D. L. and Lovette, I. J. Explosive evolutionary radiations: decreasing speciation or
413 increasing extinction through time? *Evolution* **62**(8), 1866–1875 (2008).
- 414 [17] Phillimore, A. B. and Price, T. D. Density-dependent cladogenesis in birds. *PLoS biology* **6**(3),
415 e71 (2008).
- 416 [18] Moen, D. and Morlon, H. Why does diversification slow down? *Trends in Ecology & Evolution*
417 **29**(4), 190–197 (2014).
- 418 [19] Rosenzweig, M. L. Species diversity gradients: we know more and less than we thought. *Journal*
419 *of mammalogy* **73**(4), 715–730 (1992).
- 420 [20] Rabosky, D. L. and Huang, H. A robust semi-parametric test for detecting trait-dependent
421 diversification. *Systematic Biology* **65**(2), 181–193 (2015).
- 422 [21] May, M. R. and Moore, B. R. How well can we detect lineage-specific diversification-rate shifts?
423 A simulation study of sequential AIC methods. *Systematic biology* **65**(6), 1076–1084 (2016).
- 424 [22] Moore, B. R., Höhna, S., May, M. R., Rannala, B., and Huelsenbeck, J. P. Critically evaluating
425 the theory and performance of Bayesian analysis of macroevolutionary mixtures. *Proceedings*
426 *of the National Academy of Sciences* **113**(34), 9569–9574 (2016).
- 427 [23] Rabosky, D. L., Mitchell, J. S., and Chang, J. Is BAMM flawed? Theoretical and practical
428 concerns in the analysis of multi-rate diversification models. *Systematic biology* **66**(4), 477–498
429 (2017).
- 430 [24] Rabosky, D. L. How to make any method "fail": BAMM at the kangaroo court of false equiv-
431 alency. *arXiv preprint arXiv:1711.03253* (2017).
- 432 [25] Rabosky, D. L. Extinction rates should not be estimated from molecular phylogenies. *Evolution*
433 **64**(6), 1816–1824 (2010).
- 434 [26] Mitchell, J. S. and Rabosky, D. L. Bayesian model selection with BAMM: effects of the model
435 prior on the inferred number of diversification shifts. *Methods in Ecology and Evolution* **8**(1),
436 37–46 (2017).
- 437 [27] Freckleton, R. P., Phillimore, A. B., and Pagel, M. Relating traits to diversification: a simple
438 test. *The American Naturalist* **172**(1), 102–115 (2008).

- 439 [28] Rabosky, D. L. and Goldberg, E. E. FiSSE: A simple nonparametric test for the effects of a
440 binary character on lineage diversification rates. *Evolution* **71**(6), 1432–1442 (2017).
- 441 [29] Harvey, M. G. and Rabosky, D. L. Continuous traits and speciation rates: Alternatives to
442 state-dependent diversification models. *Methods in Ecology and Evolution* (2017).
- 443 [30] Bromham, L., Hua, X., and Cardillo, M. Detecting macroevolutionary self-destruction from
444 phylogenies. *Systematic biology* **65**(1), 109–127 (2015).
- 445 [31] Hua, X. and Bromham, L. Phylometrics: an R package for detecting macroevolutionary pat-
446 terns, using phylogenetic metrics and backward tree simulation. *Methods in Ecology and Evo-*
447 *lution* **7**(7), 806–810 (2016).
- 448 [32] Maddison, W. P., Midford, P. E., and Otto, S. P. Estimating a binary character’s effect on
449 speciation and extinction. *Systematic biology* **56**(5), 701–710 (2007).
- 450 [33] Redding, D. W. and Mooers, A. Ø. Incorporating evolutionary measures into conservation
451 prioritization. *Conservation Biology* **20**(6), 1670–1678 (2006).
- 452 [34] Beaulieu, J. M. and O’Meara, B. C. Detecting hidden diversification shifts in models of trait-
453 dependent speciation and extinction. *Systematic biology* **65**(4), 583–601 (2016).
- 454 [35] Nee, S., May, R. M., and Harvey, P. H. The reconstructed evolutionary process. *Phil. Trans.*
455 *R. Soc. Lond. B* **344**(1309), 305–311 (1994).
- 456 [36] Stadler, T. Mammalian phylogeny reveals recent diversification rate shifts. *Proceedings of the*
457 *National Academy of Sciences* **108**(15), 6187–6192 (2011).
- 458 [37] Condamine, F. L., Rolland, J., and Morlon, H. Macroevolutionary perspectives to environmen-
459 tal change. *Ecology letters* **16**(s1), 72–85 (2013).
- 460 [38] Lewitus, E. and Morlon, H. Detecting environment-dependent diversification from phylogenies:
461 a simulation study and some empirical illustrations. *Systematic biology* (2017).
- 462 [39] Hedges, S. B., Marin, J., Suleski, M., Paymer, M., and Kumar, S. Tree of life reveals clock-like
463 speciation and diversification. *Molecular biology and evolution* **32**(4), 835–845 (2015).
- 464 [40] Thorne, J. L., Kishino, H., and Painter, I. S. Estimating the rate of evolution of the rate of
465 molecular evolution. *Molecular biology and evolution* **15**(12), 1647–1657 (1998).
- 466 [41] Huelsenbeck, J. P., Larget, B., and Swofford, D. A compound Poisson process for relaxing the
467 molecular clock. *Genetics* **154**(4), 1879–1892 (2000).
- 468 [42] Lartillot, N., Phillips, M. J., and Ronquist, F. A mixed relaxed clock model. *Phil. Trans. R.*
469 *Soc. B* **371**(1699), 20150132 (2016).
- 470 [43] Morlon, H., Lewitus, E., Condamine, F. L., Manceau, M., Clavel, J., and Drury, J. RPANDA:
471 an R package for macroevolutionary analyses on phylogenetic trees. *Methods in Ecology and*
472 *Evolution* **7**(5), 589–597 (2016).
- 473 [44] Yang, Z. and Rodríguez, C. E. Searching for efficient Markov chain Monte Carlo proposal
474 kernels. *Proceedings of the National Academy of Sciences* **110**(48), 19307–19312 (2013).

- 475 [45] Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. Bayesian
476 data analysis. *Bayesian data analysis*, volume 2. CRC press Boca Raton, FL, (2014).
- 477 [46] ter Braak, C. J. and Vrugt, J. A. Differential evolution Markov chain with snooker updater
478 and fewer chains. *Statistics and Computing* **18**(4), 435–446 (2008).
- 479 [47] Grafen, A. The phylogenetic regression. *Phil. Trans. R. Soc. Lond. B* **326**(1233), 119–157
480 (1989).

481 Author contribution

482 OM, FH and HM designed the study and performed research. OM contributed new analytical
483 tools and analysed data. OM, FH and HM wrote the paper.

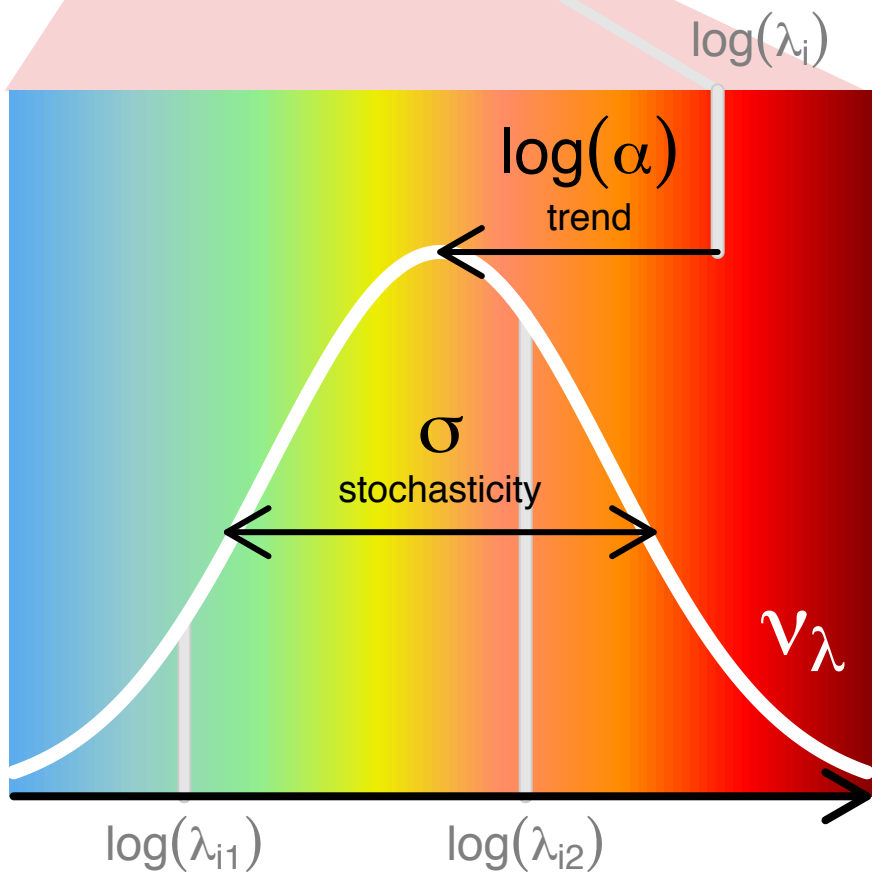
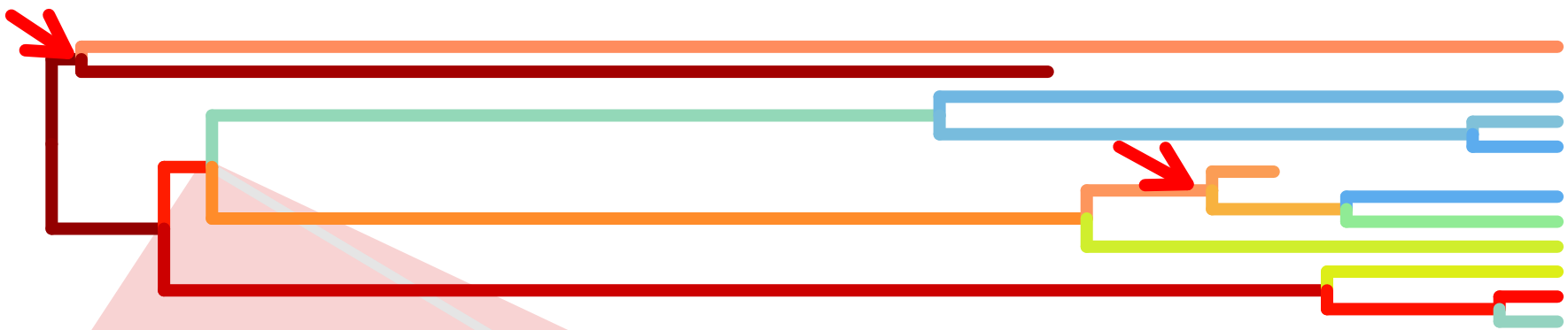
484 Figure legends

485 **Figure 1** Illustration of the cladogenetic diversification rate shift model (ClaDS). Upper panel:
486 cartoon phylogeny simulated under ClaDS, with branches colored according to their speciation
487 rate (red: high rate, blue: low rate). Speciation rates are inherited at speciation with a shift
488 determined by the probability distribution ν_λ (here taken to be a lognormal distribution, insert).
489 Red arrows indicate speciation events (and associated diversification rate shifts) that are hidden in
490 the reconstructed phylogeny as a result of extinction.

491 **Figure 2** Recovery of ClaDS parameters. Estimated λ_0 (a), α (b), and σ (c) inferred with ClaDS,
492 and (d) resulting estimation of $m = \alpha * \exp(\sigma^2/2)$. Violin plots: distribution of estimated param-
493 eters; yellow cross: median; thick black line: quartiles; red lines: values used in the simulations.
494 Different shades of brown correspond to: in a and c, the values of α used in the simulations (1.2
495 (light), 1, 0.9, 0.7 (dark)) ; in b and d, the values of σ used in the simulations (0 (light), 0.1, 0.26,
496 0.41 (dark)). Results corresponding to simulated trees of size 200; 20 trees were simulated and
497 analysed for each parameter set; results for other tree sizes are shown in Fig. S11 to S14.

498 **Figure 3** ClaDS performs well in recovering branch-specific speciation rates a) tree simulated
499 under the ClaDS model ($\lambda_0 = 0.1$, $\sigma = 0.18$, $\alpha = 1$, $\varepsilon = 0$, size $N = 200$), with branches
500 colored according to their realized speciation rate b) same tree with branches colored according
501 to inferred speciation rates c) Inferred versus simulated branch-specific speciation rates (on a log
502 scale) for 20 trees simulated with the same parameters and size as the tree from panel a; the darker
503 points highlight rates for the tree shown in panel a. Each regression line (light gray) corresponds
504 to one of the 20 trees, and the black line corresponds to the regression across all trees. The red
505 line displays the 1:1 relationship. Values in the bottom right corner correspond to the mean and
506 standard deviation of the slope and correlation coefficient across the 20 regressions, and those of the
507 relative error in branch-specific speciation rates estimates ($\lambda_{\text{estimated}}/\lambda_{\text{simulated}}$) across all branches
508 from the 20 trees.

509 **Figure 4** Patterns of diversification across 42 bird clades. Distributions across clades of (a) σ , (b)
510 α , (c) $m = \alpha \exp(\sigma^2/2)$, and (d) ε values estimated with ClaDS. e: Distributions of branch specific
511 speciation rates for each specific clade (grey and colored lines) and all clades pooled together (thick
512 black line). Red: *Accipitridae*; Orange: *Muscicapidae* & *Turdidae*, *Tyrannidae* and *Parulidae*; Dark
513 blue: *Charadrii*; Medium blue: *Ramphastides*; Light blue : *Alcedinidae* and *Phasianidae*; Brown:
514 *Scolopaci*; Green: *Anatinae* f: Exemplar phylogenies colored according to their inferred branch-
515 specific speciation rates, in Myr^{-1} , and plotted on the same time scale. Top panel: the *Accipitridae*
516 phylogeny subtends very variable rates that tend to decrease through time (inferred parameters:
517 $\sigma = 0.67$, $\alpha = 0.61$, $m = 0.76$ and $\varepsilon = 0.02$). Bottom panel: the *Ramphastides* phylogeny subtends
518 rather homogeneous rates ($\sigma = 0.16$, $\alpha = 0.97$, $m = 0.98$ and $\varepsilon = 0.05$).



- λ_0 initial speciation rate
- λ_i ancestral rate
- $\lambda_{i1}, \lambda_{i2}$ daughter rates
- v_λ distribution of daughter speciation rates
- log-normal distribution of parameters $\sigma, \alpha\lambda_i$
- $m = \alpha e^{\sigma^2/2}$ mean relative daughter rate
- μ extinction rate
- $\varepsilon = \mu/\lambda$ turnover rate

