nature ecology & evolution

ARTICLES

https://doi.org/10.1038/s41559-019-0908-0

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

Odile Maliet[1]*, Florian Hartig[2] and Hélène Morlon[1]

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

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

Shifts in diversification rates are probably quite widespread. Speciation and extinction rates may vary across lineages as a response to the particular biotic and abiotic environment experienced by each lineage[10]; they may also vary as a response to traits that affect reproductive isolation, such as reproduction mode[11], or pollination and dispersal syndromes[12]. Such changes in diversification rates probably occur far more frequently than key innovations, resulting in heterogeneous diversification rates at much finer taxonomic scales[4]. Accounting for such finer-scale heterogeneity is crucial if we want to obtain refined estimates of lineage-specific diversification rates and better understand the processes subtending heterogeneity in the diversification of life. Methods of the state–speciation–extinction family[13] can, in principle, better account for these types of heterogeneities, but they require an assumption of trait dependency of rates (Supplementary Section 3.5). Non-model-based approaches, such as the diversification rate statistic (known as the DR statistic)[4], can also account for fine-scale heterogeneities, but they are rather ad hoc and generally do not perform as well as model-based approaches[14].

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

Results

A new model of diversification rate variation. We consider a birth–death diversification process (the ClaDS model) where diversification rates are inherited at speciation, but with a shift (Fig. 1). At the beginning of the process, the clade is composed of one lineage with speciation rate λ_0 and extinction rate μ_0. At each speciation event, the two daughter lineages inherit new diversification rates (λ_{i1}, λ_{i2}) and (μ_{i1}, μ_{i2}) sampled from a joint probability distribution ν parameterized by the parental rates λ_i and μ_i. If the changes in speciation and extinction rates are assumed to be independent, λ_i values are sampled from a distribution ν_λ, μ_i values are sampled from a distribution ν_μ, and $\nu = \nu_\lambda \times \nu_\mu$. Moreover, we allow for the possibility that some extant species are missing by assuming that each extant species is observed with probability $f \leq 1$. We derive the probability density of a reconstructed phylogeny under this general model and implement its computation in R (Methods and Supplementary Sections 2–5).

We then consider several scenarios in ClaDS where: (1) ν_λ is a log-normal distribution with parameters $\log[\alpha \times \lambda]$ and σ (the latter ensures that the relative change in rate at speciation λ_i/λ is independent of the parental rate λ, with a mean, m, given by $\alpha \times \exp(\sigma^2/2)$; σ controls how constrained daughter rates are (highly constrained for small σ values) and α controls the trend at speciation (that is, whether daughter rates tend to be higher or lower than parental rates); and (2) extinction rates are either negligible ($\mu_i = 0$ for all lineages; ClaDS0), homogeneous across all lineages in the clade

[1]Institut de Biologie de l'École Normale Supérieure, École Normale Supérieure, CNRS, INSERM, PSL Research University, Paris, France. [2]Theoretical Ecology, Faculty of Biology and Preclinical Medicine, University of Regensburg, Regensburg, Germany. *e-mail: odile.maliet@orange.fr

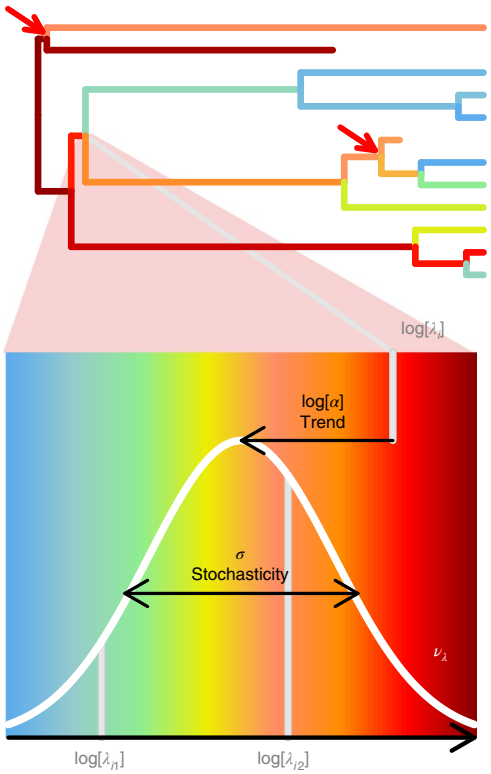

Fig. 1 | Illustration of the ClaDS model. Top, phylogeny simulated under ClaDS, with branches coloured according to their speciation rate λ_i (red, high; blue, low). Speciation rates are inherited at speciation, with a shift determined by the probability distribution ν_λ (here, taken to be a log-normal distribution with parameters $\log[\alpha \times \lambda_i]$ and σ, where α is the deterministic trend in speciation rates changes and σ the stochasticity in the rate inheritance). Red arrows indicate speciation events (and associated diversification rate shifts) that are hidden in the reconstructed phylogeny as a result of extinction.

($\mu_i = \mu_0$ for all lineages; ClaDS1) or vary across lineages, but with a constant turnover ε (that is, $\mu_i/\lambda_i = \varepsilon$ for all lineages; ClaDS2). We use Monte Carlo simulations under ClaDS1 and ClaDS2 (Methods and Supplementary Section 6) to verify that our likelihood expression is correct (Supplementary Section 6 and Supplementary Figs. 6–8). Finally, we implement a Monte Carlo Markov chain (MCMC) sampler that, given a reconstructed phylogeny, simultaneously estimates both the parameters of ClaDS (λ_0, α, σ, and either μ_0 or ε) and the speciation rates λ_i at the origin of each branch i of the phylogeny (Methods and Supplementary Section 7; see also Supplementary Section 8 and Supplementary Figs. 9 and 10 for a test of the sampler). Branch-specific extinction rates μ_i at the origin of each branch i of the phylogeny are given by μ_0 for ClaDS1 and by $\varepsilon \times \lambda_i$ for ClaDS2. In what follows, for simplicity, we refer to λ_i and μ_i as 'branch-specific rates' instead of the more accurate 'rates at the origin of each branch'.

Under these scenarios of the ClaDS process, heterogeneity in speciation rates across lineages is determined on the one hand by a stochastic component (controlled by σ) and on the other hand by a trend component (controlled by m). When the expected daughter rate is equal to the parental rate ($m = 1$), the resulting trees are relatively imbalanced and tippy (Supplementary Section 1 and Supplementary Figs. 1 and 2): lineages that by chance have high speciation rates early in clade's history spread, leading to rates that are heterogeneous across lineages, and average rates that increase through time. This sorting effect is exacerbated when the expected daughter rate is higher than the parental rate ($m > 1$; Supplementary Figs. 1 and 2), corresponding to a 'niche-piling' scenario where diversity begets diversity[15]. In contrast, when the expected daughter rate is lower than the parental rate ($m < 1$), corresponding to a 'niche-filling' scenario where diversification gets harder as new species arise[16–18], the heterogeneity in speciation rates across lineages is reduced, and with a low enough m, the average rate is constant or even decreases through time (Supplementary Figs. 1 and 2). Importantly, ClaDS is able to produce the combination of stemmy and imbalanced tree shapes observed in nature, and under a wider set of parameter values for the scenario with constant turnover (ClaDS2) than the scenario with constant extinction rate (ClaDS1) (Supplementary Figs. 1 and 2).

Performance of ClaDS. We begin by testing the performance of ClaDS under frequent rate changes and in the absence of extinction (ClaDS0) (Methods). We find that the approach provides unbiased estimates of all of the model's parameters for large enough trees (200 tips; Fig. 2); the relative change in rate at speciation m is also well estimated (Fig. 2d). As expected, bias and variability around parameter estimates increase for smaller trees (Supplementary Figs. 11–14).

ClaDS provides reliable estimates of branch-specific speciation rates on average: while low rates tend to be slightly overestimated and large rates slightly underestimated, ClaDS can detect regions of the tree with relatively high or low rates (Fig. 3 and Supplementary Figs. 15 and 16).

When also considering extinctions, focusing on the scenario with constant turnover (ClaDS2), as it generally produced tree shapes closer to those observed in nature, we found that estimates remain accurate at low levels of extinction ($\varepsilon = 0.1$), for both model parameters (Supplementary Fig. 20) and branch-specific speciation rates (Supplementary Fig. 21). At high levels of extinction ($\varepsilon = 0.9$), σ and, when the mean change in rate at speciation m approaches 1, branch-specific speciation rates, remain well estimated. However, this is not the case of the turnover rate ε, α and branch-specific speciation rates when $m < 1$, although accounting for extinction does improve inferences over ignoring it (Supplementary Figs. 20 and 21). When extinction is not accounted for, estimated branch-specific speciation rates are generally lower than realized ones, but higher than realized net diversification rates (Supplementary Fig. 21c,d).

If there are a small number of major rate shifts during the evolution of clades, rather than many small changes (tested here with a single rate shift; Methods), ClaDS is still able to provide reliable estimates of branch-specific rates (Supplementary Figs. 17 and 19). The model is also able to detect when two branches in the tree belong to distinct speciation regimes as soon as the difference in rates between the two regimes is large enough (a twofold increase or decrease in our simulations) and both regimes are represented by a large enough number of branches in the phylogeny (Supplementary Fig. 19, left). The false detection rate associated with this test is low (Supplementary Fig. 19, right).

Finally, when comparing the performance of ClaDS with that of two other popular methods for estimating branch-specific rates (the diversification rate (DR) statistic and Bayesian analysis of macroevolutionary mixtures (BAMM)[4,5]) under various simulation schemes (Supplementary Section 9), we find, overall, that ClaDS outperforms the other methods for trees simulated with both many small shifts at speciation (Supplementary Figs. 22 and 26–28) and gradual changes along branches (Supplementary Fig. 24), and that it performs as well as other methods for trees simulated with few large shifts (Supplementary Fig. 23) and variations in extinction rates (Supplementary Fig. 25). Importantly, ClaDS provides reliable estimates of the variance in rates under both the many small scenario and the few large shifts scenario (Supplementary Figs. 30

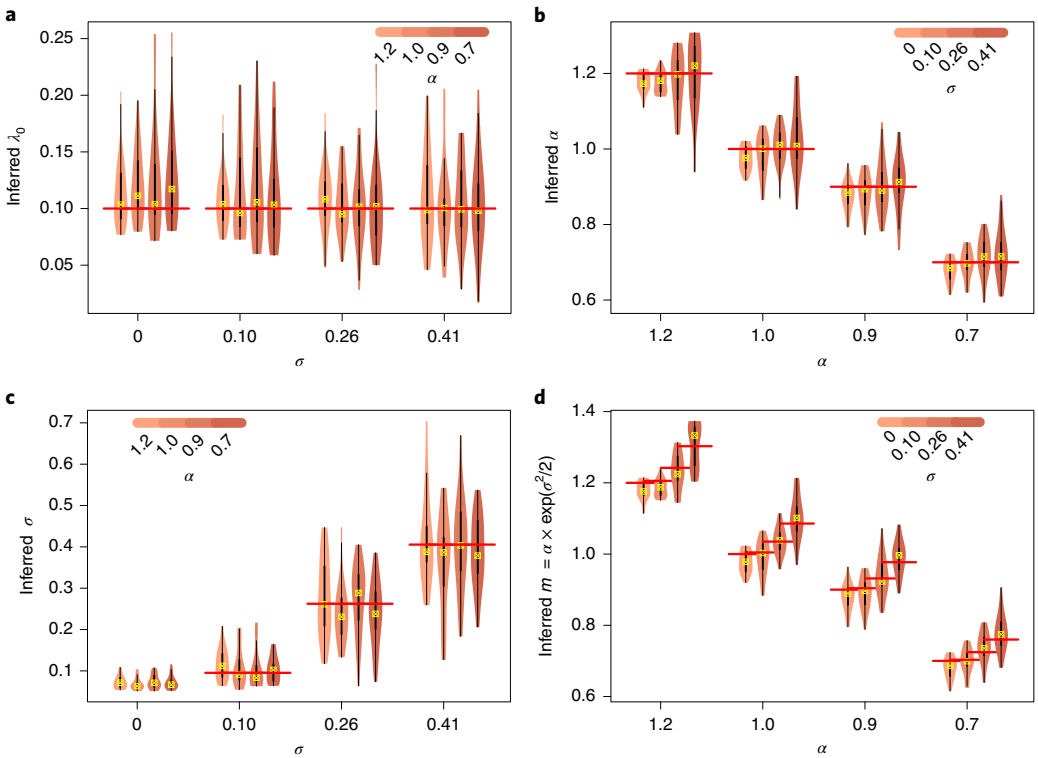

Fig. 2 | Recovery of ClaDS parameters. a–d, Estimated λ_0 (**a**), α (**b**) and σ (**c**) inferred with ClaDS, and the resulting estimation of $m = \alpha \times \exp(\sigma^2/2)$ (**d**). Violin plots show the distribution of estimated parameters. Yellow crosses represent the medians, thick black lines the quartiles, and red lines the values used in the simulations. Different shades of brown in **a** and **c** correspond to the values of α used in the simulations (1.2, 1.0, 0.9 and 0.7). In **b** and **d**, they correspond to the values of σ used in the simulations (0, 0.10, 0.26 and 0.41). The results correspond to simulated trees with 200 tips; 20 trees were simulated and analysed for each parameter set. The results for other tree sizes are shown in Supplementary Figs. 11–14.

and 32), while BAMM underestimates the variance in rates under the many small shifts scenario (Supplementary Fig. 31). ClaDS and BAMM provide low (and similar) estimates of the variance in rates for trees simulated under constant rates (Supplementary Fig. 29); in the presence of rate heterogeneity, they tend to underestimate rather than overestimate rate variance (Supplementary Figs. 30–33), and BAMM more so than ClaDS.

Diversification across the avian radiation. When applying ClaDS to major bird clades (Methods), we found that lineage-specific speciation rates can vary by as much as two orders of magnitude within clades (Fig. 4e). For example, in Accipitridae (hawks and allies), speciation rates range from 0.013−1.200 Myr^{-1}, which almost covers the range found across the entire avian radiation (0.013−5.000 Myr^{-1}). Comparable within-clade heterogeneities occur in other clades, such as Muscicapidae, Turdidae, Tyrannidae and Parulidae (Fig. 4e, orange). Such within-clade heterogeneities are way above heterogeneities arising from estimation error (Supplementary Fig. 29). A variance partitioning of speciation rates across the bird radiation (Methods) reveals that intraclade variance accounts for 76% of the total variance. In comparison, BAMM would have estimated far fewer within-clade heterogeneities, with an intraclade variance accounting for only 46% of the total variance (Supplementary Fig. 34). Given our simulation results, this suggests that BAMM underestimates the intraclade variance, and thus that many small shifts occurred during bird diversification that BAMM cannot detect.

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

Discussion

Models of diversification applied to phylogenies of extant taxa are increasingly used to understand the long-term evolution of biodiversity. These approaches have highlighted how variable diversification rates can be across the tree of life, and the importance of these variations for explaining current patterns of diversity (the so-called 'diversification rate hypothesis'[19]). Yet, despite recent advances in phylogenetic approaches for understanding diversification, detecting diversification rate variations and the processes underlying these variations remains a challenge spurring a heated debate[20–24]. In this paper, we have developed ClaDS—a new model with frequent small

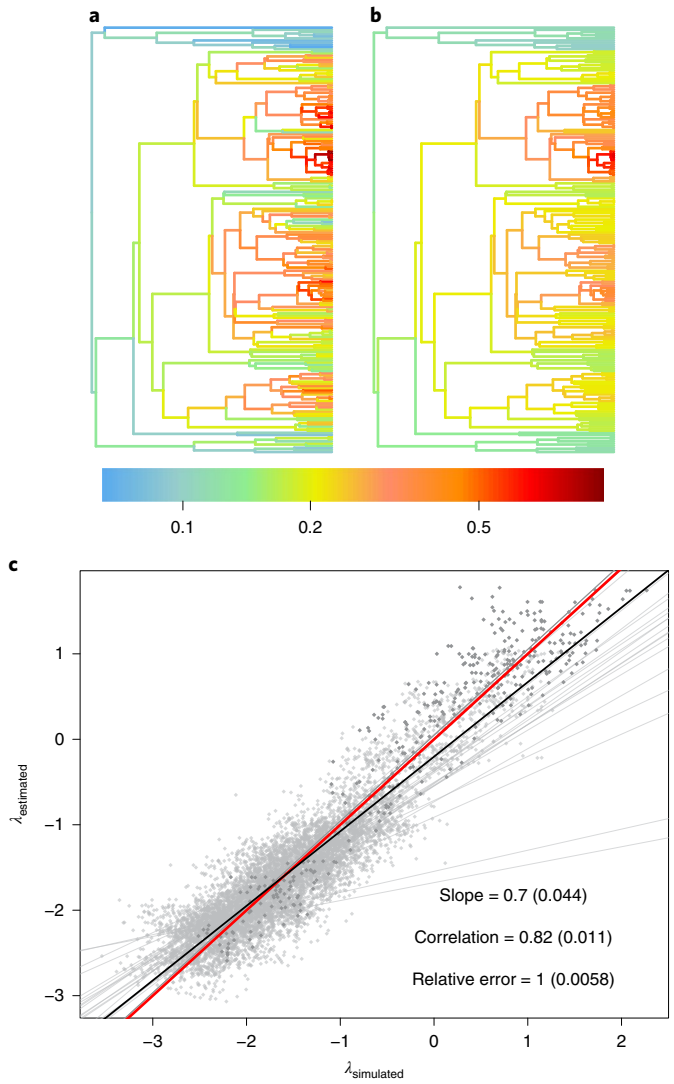

Fig. 3 | ClaDS performs well in recovering branch-specific speciation rates. a, Tree simulated under the ClaDS model ($\lambda_0 = 0.1$; $\sigma = 0.18$; $\alpha = 1$; $\varepsilon = 0$; size, $n = 200$ tips), with branches coloured according to their realized speciation rate. **b**, The same tree as in **a**, but with branches coloured according to inferred speciation rates. **c**, Inferred versus simulated branch-specific speciation rates (on a log scale) for 20 trees simulated with the same parameters and size as the tree in **a**; the darker points highlight rates for the tree shown in **a**. Each regression line (light grey) corresponds to one of the 20 trees, and the black line corresponds to the regression across all trees. The red line displays the 1:1 relationship. Values in the bottom right corner correspond to the mean (± s.d.) of the slope and correlation coefficient across the 20 regressions, and those of the relative error in branch-specific speciation rates estimates ($\lambda_{estimated}/\lambda_{simulated}$) across all branches from the 20 trees.

variations in diversification rates—together with a method to infer branch-specific diversification rates on a phylogeny. We have shown using simulations that ClaDS accurately estimates branch-specific rates. Finally, applying ClaDS to the bird phylogeny, we have shown that small but frequent changes have been instrumental in shaping global rate variation during the avian radiation.

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

Another advantage of ClaDS is to avoid using model selection to select the number and location of rate shifts, by assuming that shifts happen at each speciation event. In the frequently used MEDUSA method[1], the stepwise Akaike information criterion is used to perform this selection, with associated statistical limitations[21]. In the approach of Morlon et al.[3], likelihood ratio tests are performed to select the number of shifts, but the location of these shifts needs to be fixed a priori. Finally, in the popular BAMM[5], reversible jump MCMC is used, with a prior on the number and location of shifts that may influence the results[22,26]. ClaDS avoids these limitations, while still performing well in the presence of rare rate shifts with large effects.

Maybe more important than these technical aspects, ClaDS represents a view of evolution distinct from that of previous models: existing models focus on a small number of discrete diversification shift events spread across the tree—an idea that fits well with the concept of key innovations driving major diversification shifts[1,3,5]. In contrast, ClaDS allows for frequent variations linked, for example, to changes in environmental conditions or associations with continuously evolving heritable traits. Accordingly, ClaDS does not aim to identify specific nodes in a phylogeny subtending major diversification rate shifts. Rather, it assumes that rate shifts happen at each speciation event, and focuses on estimating branch-specific diversification rates. In nature, both many shifts with small effects and few shifts with large effects are likely to occur, so it is reassuring to see that ClaDS can properly estimate branch-specific rates under these two evolutionary processes.

Accurately estimating branch-specific diversification rates is a critical step for understanding the processes that lead some species groups to diversify faster than others. For example, species' traits can modulate their propensity to diversify, and tests based on assessing the correlation between trait values at a phylogeny's tips and metrics capturing the diversification rate of the corresponding lineages ('tip-rate correlation' tests) have been developed to detect such effects[27]. These types of tests have regained interest lately (see, for example, STRAPP[20], FiSSE[28], ES-sim[29] and pNoTO[30,31]), as an alternative or complement to state-dependent speciation–extinction methods that jointly model diversification dynamics and trait evolution[13,32]. However, current metrics of species-level diversification rates have limitations. Some are derived from BAMM[5] and thus reflect a limited set of diversification rate regimes rather than lineage-specific rates per se. Others are summary statistics describing phylogenetic branching patterns, such as the 'node density'[27], 'equal split'[33] or 'diversification rate'[4] statistics; they are not rigorously derived from speciation-extinction models, and

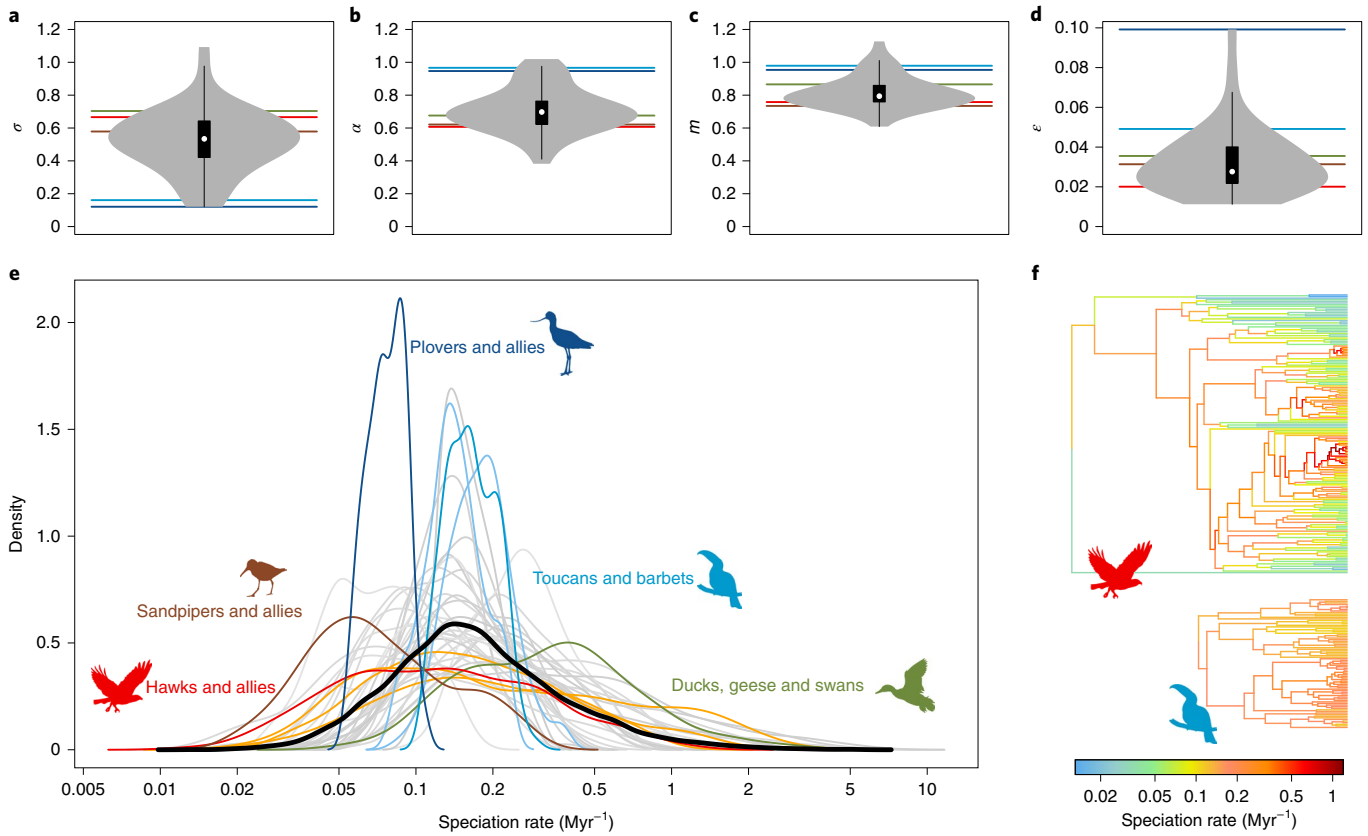

Fig. 4 | Patterns of diversification across 42 bird clades. a–d, Distributions across clades (grey area) and values for specific clades (coloured lines) of σ (**a**), α (**b**), $m = \alpha \times \exp(\sigma^2/2)$ (**c**) and ε values (**d**) estimated with ClaDS. The white dots represent the median of each distribution, the black boxes the first and third quartiles, and the vertical black lines the most extreme values that are less than (1.5 × box size) awayfrom the box. **e**, Distributions of branch-specific speciation rates for each specific clade (grey and coloured lines) and all clades pooled together (thick black line). Red, Accipitridae; orange, Muscicapidae, Turdidae, Tyrannidae and Parulidae; dark blue, Charadrii; medium blue, Ramphastides; light blue, Alcedinidae and Phasianidae; brown, Scolopaci; green, Anatinae. **f**, Exemplar phylogenies coloured according to their inferred branch-specific speciation rates (see coloured bar) and plotted on the same time scale. Top, the Accipitridae phylogeny subtends very variable rates that tend to decrease through time (inferred parameters: $\sigma = 0.67$; $\alpha = 0.61$; $m = 0.76$; $\varepsilon = 0.02$). Bottom, the Ramphastides phylogeny subtends rather homogeneous rates ($\sigma = 0.16$; $\alpha = 0.97$; $m = 0.98$; $\varepsilon = 0.05$).

they generally perform worse than model-based approaches[14] (Supplementary Figs. 22–27). ClaDS provides tip-level estimates of diversification rates that should help identify the specific features of a species that make it more or less prone to diversify. In the future, we could imagine a hybrid between state-dependent speciation–extinction and ClaDS that would account for both trait-dependent diversification and residual rate variation not accounted for by the trait, in the spirit of hidden state models (HiSSE[34] and MSBD[6]). This could, for example, be done by imputing in ClaDS specific trend parameters, α, corresponding to trait shifts.

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

Our ClaDS analysis of the avian radiation reveals a series of compelling results. First, and even though these estimates need to be taken with caution, we find significant (non-zero) turnover rates. Second, we find a pervasive pattern of declines in speciation rates over time congruent with previous studies[16–18]. Third, we find a remarkable heterogeneity in speciation rates, with per-lineage rates that vary by two orders of magnitude (0.01–5.00 Myr^{-1}), peaking around 0.15 Myr^{-1}. Fourth, we find that variability in speciation rates can be as high within as between clades, suggesting that rate variation may be much more widespread than is currently thought and implemented in existing models. Finally, we highlight a remarkable difference across clades in terms of how constrained their diversification rates are, with plovers and allies on one extreme, and hawks and allies on the other extreme of a continuum of rates that vary between less than twofold and more than 80-fold (Fig. 4e,f). These differences in the levels of constraint of diversification rates are striking, and remain to be explained: these could be linked to differences in genetic architecture, developmental constraints or biogeographies, for example.

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

Methods

Likelihood, simulation and Bayesian implementation of ClaDS. *Likelihood.* We derived the probability density of observing a reconstructed phylogeny with

branches delimited by the times $(t_i, s_i)_{i\in[[1,n]]}$, and speciation and extinction rates λ_i and μ_i at time t_i (that is, at the origin of each branch), under the cladogenetic diversification rate shift model (Supplementary Sections 2–4). We note Θ, the parameters of the new rate distribution ν. The probability density can be derived from three main probability functions: $\Phi_{\Theta,\lambda,\mu}(t)$, the probability that a lineage alive at time t has speciation and extinction rates λ and μ and no descendant in the reconstructed phylogeny; $\chi_{\Theta,\lambda,\mu}(t)$, the probability that a lineage alive at time t has speciation and extinction rates λ and μ and exactly one descendant species sampled in the reconstructed 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 and extinction rates λ and μ and gives birth at time s to two daughter lineages that respectively have speciation rates λ_1 and λ_2 and extinction rates μ_1 and μ_2. We obtained ODEs to solve for Φ, χ and ξ by considering the different events that can happen during a short time interval Δ_t and making Δ_t tend to 0 (Supplementary Sections 3.1–3.3). Under a pure birth model and for a completely sampled phylogeny, the ODEs can be solved analytically (Supplementary Section 4). In the presence of extinction and/or if there are missing taxa in the phylogeny, Φ, χ and ξ are computed by integrating the ODEs numerically, which is more computationally intensive (Supplementary Section 5).

Simulation. We implemented a simulation algorithm of ClaDS in the R package RPANDA[43] function sim_ClaDS (Supplementary Section 1). In this implementation, the speciation rates of daughter lineages are drawn independently from a distribution ν_λ. Their extinction rates are drawn from a distribution ν_μ, given by either μ_0 (constant extinction rate scenario; ClaDS1) or $\varepsilon \times \lambda_{s_i,1}$ and $\varepsilon \times \lambda_{s_i,2}$ (constant turnover scenario; ClaDS2). ν_λ and ν_μ can be normal, log-normal or uniform distributions. The simulations are continued until a stopping criterion is met: either a fixed time or a fixed number of species. In addition, sim_ClaDS takes as one of its arguments a parameter p controlling the probability that a shift happens at each speciation event (the default value $p=1$ corresponds to the model investigated here) and a parameter n controlling a maximum number of shifts (the default value $n=+\infty$ corresponds to the model investigated here; if n takes a finite value, p switches to 0 as soon as n switches have occurred).

Bayesian implementation. We implemented a Bayesian inference approach for fitting ClaDS to reconstructed phylogenies in the R package RPANDA[43] function fit_ClaDS (Supplementary Section 7). To fit ClaDS0 (no extinction), we use a Metropolis within Gibbs MCMC sampler with a Bactrian proposal[44], and convergence is monitored by running three MCMC chains in parallel and computing Gelman statistics[45]. To fit ClaDS1 and ClaDS2 (that is, in the presence of extinction) and/or if there are missing taxa in the phylogeny, we use the faster-blocked differential evolution MCMC sampler, with sampling from the past of the chains[46]. We also run three chains. For both with and without extinction, we use an inverse gamma prior with shape parameter 1 and rate parameter 0.1 for σ, and a flat prior for all other parameters. Each estimate is computed as the mean over the iterations and the three chains.

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

Many small rate shifts. For each combination of the following parameter values, we simulated 20 pure birth trees, stopping the simulation when a target tip number of 50, 100 or 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 {1.2, 1.1, 1.0, 0.95, 0.9, 0.7}. We recorded the realized speciation rate on each branch in each of these simulations. We then ran ClaDS on each simulated tree using our run_ClaDS0 function. Lastly, we compared the retrieved estimates of λ_0, σ and α with their simulated values; we also compared the retrieved estimates of branch-specific speciation rates for each tree with their realized values by performing linear regressions and computing relative errors (ratio of estimated versus realized rates).

To explore the model accounting for extinction, we simulated 5 trees with 100 tips under 4 scenarios with a constant turnover rate (ClaDS2), and for each condition either low ($\varepsilon=0.1$) or high ($\varepsilon=0.9$) turnover (8 scenarios in total). We focused on the scenario with constant turnover because this scenario produced tree shapes similar to those of empirical trees under a wider set of parameter values than the alternative scenario with a constant extinction rate (Supplementary Fig. 2.1 versus Supplementary Fig. 2.2). Maintaining a balance where extinction is neither negligible nor driving clades to extinction is also easier under ClaDS2. The four scenarios were as follows: (1) high heterogeneity and decreasing rates: $\lambda_0=0.1$; $\sigma=0.7$; $\alpha=0.7$ (mean relative change: $m=0.9$); (2) no heterogeneity and constant rates (equivalent to constant-rate birth–death trees): $\lambda_0=0.1$; $\sigma=0$; $\alpha=1$ ($m=1$); (3) low heterogeneity and no average change in rate at speciation: $\lambda_0=0.1$; $\sigma=0.2$; $\alpha=0.98$ ($m=1$); and (4) low heterogeneity and decreasing rates: $\lambda_0=0.1$; $\sigma=0.2$; $\alpha=0.88$ ($m=0.9$). We recorded the realized speciation rate at the beginning of each branch in each of these simulations. We then ran ClaDS on each simulated tree using our R function, both accounting (run_ClaDS) and not accounting for extinction (run_ClaDS0), the latter to evaluate the bias resulting from not accounting for extinction when it occurs. Lastly, we compared the retrieved estimates of σ, α, m and ε for each tree with their simulated values. We did not compare the retrieved estimates of λ_0 with the simulated values because the estimates correspond to the speciation rate at the crown while the simulated values correspond to the speciation rate at the stem. These two rates can be very different in the presence of extinction. We also compared the retrieved estimates of branch-specific speciation rates and net diversification rates (speciation minus extinction) for each tree with their realized values by performing linear regressions and computing relative errors.

Few large rate shifts. We also tested the behaviour of ClaDS under a 'key innovation' scenario with only a single, large rate shift during the history of the clade. To simulate this scenario, we used our sim_ClaDS function with λ_0 (the background rate in this case) fixed at 0.1, P (the probability that a rate shift happens at each speciation event) fixed at 0.02, and n (the maximum number of shifts) fixed at 1. The new speciation rate took a series of values from lower (uniformly 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], [0.2,0.3], [0.3,0.4], [0.4,1]) than the background rate. For each of these rate values, we simulated phylogenies 200 tips until we had a good coverage of subclade new rate/size combination (from 300–500 phylogenies per parameter set). In such simulations, there are only two distinct rates across the tree: the background rate and the new rate. We then ran ClaDS on each simulated tree using our run_ClaDS0 function, and compared the retrieved estimates of branch-specific speciation rates for each tree with their simulated values by performing linear regressions and computing relative errors. Finally, we tested whether the model is able to detect whether two branches in the tree belong to the same or distinct speciation regime(s): two branches were considered to have significantly different rates (distinct regimes) if the difference in the estimated speciation rates between the two branches was of a constant sign on at least 95% of the MCMC chains. We assessed the significance of speciation rate differences (and the corresponding sign) for all pairs of branches in the simulated trees. Finally, we quantified the 'proper detection' rate as the proportion of pairs for which a significant difference was inferred when the two branches indeed belonged to distinct speciation regimes (that is, one had the background speciation rate and the other had the new rate), and the 'false detection' rate as the proportion of pairs for which a significant difference was inferred, while the two branches actually belonged to the same speciation regime (that is, both had either the background speciation rate or the new rate).

Diversification of the avian radiation. We applied ClaDS, accounting for extinction (ClaDS2; model with constant turnover) and incomplete sampling, to bird phylogenies. We used the maximum clade credibility trees from Jetz et al.[4] with only the species for which there was molecular data, along with the associated sampling fractions provided by the authors. Most of these are family-level phylogenies, with some spawning two or a few more families. We ran the model on the 42 bird phylogenies with more than 50 species. We report the distribution of branch-specific speciation rates across the 42 clades, as well as individual distributions for each clade. We partitioned the total variance of the logarithm of the branch-specific speciation rates $\left(\sum_i (\ln[\lambda_i]-\overline{\ln[\lambda]})^2\right.$, where $\overline{\ln[\lambda]}$ is the mean of the log of the speciation rates for all branches in all clades) between the intraclade $\left(\sum_i (\ln[\lambda_i]-\overline{\ln[\lambda_{c_i}]})^2\right.$, where c_i is the clade to which branch i belongs, and $\overline{\ln[\lambda_c]}$ is the mean of the log of the speciation rates for all branches in clade c) and interclade variance $\left(\sum_i (\overline{\ln[\lambda_{c_i}]}-\overline{\ln[\lambda]})^2\right)$. We also tested for a potential correlation between the variance in rates and the size (number of tips) and age (crown age) of clades using PGLS[47] (two-sided test) on the Hackett backbone phylogeny provided in Jetz et al.[4].

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The simulated phylogenies used to test the method are available at https://github.com/OdileMaliet/ClaDS/tree/master/Simulations in a file named trees.zip. All of the empirical data used for the analysis were obtained from the Jetz et al.[4] study, and are available from https://www.nature.com/articles/nature11631.

Code availability

The R functions used to simulate and fit the model are available in the RPANDA R package. All of the codes used to test our method are available from the GitHub repository at https://github.com/OdileMaliet/ClaDS.git.

Received: 4 September 2018; Accepted: 26 April 2019;
Published online: 03 June 2019

References

1. Alfaro, M. E. et al. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proc. Natl Acad. Sci. USA* **106**, 13410–13414 (2009).
2. Chan, K. M. & Moore, B. R. SymmeTREE: whole-tree analysis of differential diversification rates. *Bioinformatics* **21**, 1709–1710 (2004).
3. Morlon, H., Parsons, T. L. & Plotkin, J. B. Reconciling molecular phylogenies with the fossil record. *Proc. Natl Acad. Sci. USA* **108**, 16327–16332 (2011).
4. Jetz, W., Thomas, G., Joy, J., Hartmann, K. & Mooers, A. The global diversity of birds in space and time. *Nature* **491**, 444–448 (2012).
5. Rabosky, D. L. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS ONE* **9**, e89543 (2014).
6. Barido-Sottani, J., Vaughan, T. G. & Stadler, T. A multi-state birth–death model for Bayesian inference of lineage-specific birth and death rates. Preprint at https://www.biorxiv.org/content/10.1101/440982v1 (2018).
7. Sanderson, M. J. & Wojciechowski, M. F. Diversification rates in a temperate legume clade: are there so many species of *Astragalus* (Fabaceae)? *Am. J. Bot.* **83**, 1488–1502 (1996).
8. Miller, A. H. In *Ornithologie als Biologische Wissenschaft* (eds Mayr, E. and Schüz, E.) 84–88 (Carl Winter, 1949).
9. Hunter, J. P. Key innovations and the ecology of macroevolution. *Trends Ecol. Evol.* **13**, 31–36 (1998).
10. Benton, M. J. The Red Queen and the court jester: species diversity and the role of biotic and abiotic factors through time. *Science* **323**, 728–732 (2009).
11. Goldberg, E. E. et al. Species selection maintains self-incompatibility. *Science* **330**, 493–495 (2010).
12. Onstein, R. E. et al. Frugivory-related traits promote speciation of tropical palms. *Nat. Ecol. Evol.* **1**, 1903–1911 (2017).
13. FitzJohn, R. G. Diversitree: comparative phylogenetic analyses of diversification in R. *Methods Ecol. Evol.* **3**, 1084–1092 (2012).
14. Title, P. O. & Rabosky, D. L. Tip rates, phylogenies, and diversification: what are we estimating, and how good are the estimates? *Methods Ecol. Evol.* https://doi.org/10.1111/2041-210X.13153 (2018).
15. Emerson, B. C. & Kolm, N. Species diversity can drive speciation. *Nature* **434**, 1015–1017 (2005).
16. Rabosky, D. L. & Lovette, I. J. Explosive evolutionary radiations: decreasing speciation or increasing extinction through time? *Evolution* **62**, 1866–1875 (2008).
17. Phillimore, A. B. & Price, T. D. Density-dependent cladogenesis in birds. *PLoS Biol.* **6**, e71 (2008).
18. Moen, D. & Morlon, H. Why does diversification slow down? *Trends Ecol. Evol.* **29**, 190–197 (2014).
19. Rosenzweig, M. L. Species diversity gradients: we know more and less than we thought. *J. Mammal.* **73**, 715–730 (1992).
20. Rabosky, D. L. & Huang, H. A robust semi-parametric test for detecting trait-dependent diversification. *Syst. Biol.* **65**, 181–193 (2015).
21. May, M. R. & Moore, B. R. How well can we detect lineage-specific diversification-rate shifts? A simulation study of sequential AIC methods. *Syst. Biol.* **65**, 1076–1084 (2016).
22. Moore, B. R., Höhna, S., May, M. R., Rannala, B. & Huelsenbeck, J. P. Critically evaluating the theory and performance of Bayesian analysis of macroevolutionary mixtures. *Proc. Natl Acad. Sci. USA* **113**, 9569–9574 (2016).
23. Rabosky, D. L., Mitchell, J. S. & Chang, J. Is BAMM flawed? Theoretical and practical concerns in the analysis of multi-rate diversification models. *Syst. Biol.* **66**, 477–498 (2017).
24. Rabosky, D. L. How to make any method "fail": BAMM at the kangaroo court of false equivalency. Preprint at https://arxiv.org/abs/1711.03253 (2017).
25. Rabosky, D. L. Extinction rates should not be estimated from molecular phylogenies. *Evolution* **64**, 1816–1824 (2010).
26. Mitchell, J. S. & Rabosky, D. L. Bayesian model selection with BAMM: effects of the model prior on the inferred number of diversification shifts. *Methods Ecol. Evol.* **8**, 37–46 (2017).
27. Freckleton, R. P., Phillimore, A. B. & Pagel, M. Relating traits to diversification: a simple test. *Am. Nat.* **172**, 102–115 (2008).
28. Rabosky, D. L. & Goldberg, E. E. FiSSE: a simple nonparametric test for the effects of a binary character on lineage diversification rates. *Evolution* **71**, 1432–1442 (2017).
29. Harvey, M. G. & Rabosky, D. L. Continuous traits and speciation rates: alternatives to state-dependent diversification models. *Methods Ecol. Evol.* **9**, 984–993 (2017).
30. Bromham, L., Hua, X. & Cardillo, M. Detecting macroevolutionary self-destruction from phylogenies. *Syst. Biol.* **65**, 109–127 (2015).
31. Hua, X. & Bromham, L. Phylometrics: an R package for detecting macroevolutionary patterns, using phylogenetic metrics and backward tree simulation. *Methods Ecol. Evol.* **7**, 806–810 (2016).
32. Maddison, W. P., Midford, P. E. & Otto, S. P. Estimating a binary character's effect on speciation and extinction. *Syst. Biol.* **56**, 701–710 (2007).
33. Redding, D. W. & Mooers, A. Ø. Incorporating evolutionary measures into conservation prioritization. *Conserv. Biol.* **20**, 1670–1678 (2006).
34. Beaulieu, J. M. & O'Meara, B. C. Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. *Syst. Biol.* **65**, 583–601 (2016).
35. Nee, S., May, R. M. & Harvey, P. H. The reconstructed evolutionary process. *Phil. Trans. R. Soc. Lond. B* **344**, 305–311 (1994).
36. Stadler, T. Mammalian phylogeny reveals recent diversification rate shifts. *Proc. Natl Acad. Sci. USA* **108**, 6187–6192 (2011).
37. Condamine, F. L., Rolland, J. & Morlon, H. Macroevolutionary perspectives to environmental change. *Ecol. Lett.* **16**, 72–85 (2013).
38. Lewitus, E. & Morlon, H. Detecting environment-dependent diversification from phylogenies: a simulation study and some empirical illustrations. *Syst. Biol.* **67**, 576–593 (2017).
39. Hedges, S. B., Marin, J., Suleski, M., Paymer, M. & Kumar, S. Tree of life reveals clock-like speciation and diversification. *Mol. Biol. Evol.* **32**, 835–845 (2015).
40. Thorne, J. L., Kishino, H. & Painter, I. S. Estimating the rate of evolution of the rate of molecular evolution. *Mol. Biol. Evol.* **15**, 1647–1657 (1998).
41. Huelsenbeck, J. P., Larget, B. & Swofford, D. A compound Poisson process for relaxing the molecular clock. *Genetics* **154**, 1879–1892 (2000).
42. Lartillot, N., Phillips, M. J. & Ronquist, F. A mixed relaxed clock model. *Phil. Trans. R. Soc. B* **371**, 20150132 (2016).
43. Morlon, H. et al. RPANDA: an R package for macroevolutionary analyses on phylogenetic trees. *Methods Ecol. Evol.* **7**, 589–597 (2016).
44. Yang, Z. & Rodríguez, C. E. Searching for efficient Markov chain Monte Carlo proposal kernels. *Proc. Natl Acad. Sci. USA* **110**, 19307–19312 (2013).
45. Gelman, A et al. *Bayesian Data Analysis* (CRC press, 2014).
46. Ter Braak, C. J. & Vrugt, J. A. Differential evolution Markov chain with snooker updater and fewer chains. *Stat. Comput.* **18**, 435–446 (2008).
47. Grafen, A. The phylogenetic regression. *Phil. Trans. R. Soc. Lond. B* **326**, 119–157 (1989).

Acknowledgements

The authors are very grateful to L. Arístide, J. Clavel, J. Drury, C. Fruciano, S. Lambert, E. Lewitus, M. Manceau, O. Missa, B. Perez, A. Catarina Silva and G. Sommeria-Klein for their helpful comments on an earlier version of this manuscript. This work was supported by an AMX grant (from École Polytechique) and the LabEx MemoLife (to O.M.), PROCOPE Mobility Grant 57134817 (to F.H. and H.M.) and the European Research Council (ERC 616419-PANDA to H.M.).

Author contributions

O.M., F.H. and H.M. designed the study and performed research. O.M. contributed new analytical tools and analysed the data. O.M., F.H. and H.M. wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41559-019-0908-0.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to O.M.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

nature research

Corresponding author(s): Odile Maliet

Last updated by author(s): Apr 15, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	We applied our model to bird phylogenies, using the MCC trees from Jetz et al (2012, Nature). These phylogenies are publicly available from the supplementary material of this paper, https://www.nature.com/articles/nature11631
Data analysis	The data and all the simulation results were analysed in R, using custom code that is available upon request. Analyses and plots involved using functions from the following R packages : ape, apTreeShape, coda, fields, geiger, vioplot, phytools, parallel, pracma,expoRkit, expm, Matrix.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The simulated phylogenies used to test the method are available at {https://github.com/OdileMaliet/ClaDS/tree/master/Simulations}, in the file named trees.zip. All the empirical data used for the analysis were obtained from Jetz et al. (2012) study, and are available on \url{https://www.nature.com/articles/nature11631}. The R functions used to simulate and fit are available in the RPANDA R-package. All the codes used to test our method are available on the github repository \url{https://github.com/OdileMaliet/ClaDS.git}.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	We present a new method to estimate branch-specific speciation rates on reconstructed phylogenies. We apply this method to a dataset of bird phylogenies.
Research sample	We used the MCC trees from Jetz et al. (2012) with only the species for which there was molecular data, along with the associated sampling fractions provided by the authors. Most of these are family level phylogenies, with some spawning two or a few more families.
Sampling strategy	We used all phylogenies from the dataset that included at least 50 species.
Data collection	The phylogenies on which we applied our method were obtained from the supplementary material of Jetz et al. (2012), available on https://www.nature.com/articles/nature11631
Timing and spatial scale	Not applicable to our study.
Data exclusions	We excluded phylogenies with less than 50 species, because our tests on simulated trees show that accurate estimation of our model's parameter recquires trees that are large enough.
Reproducibility	Our study do not imply any experimental data.
Randomization	No allocation into groups was performed, so our study required no randomization procedure.
Blinding	Blinding was not relevant to our study because we used all available phylogenies including at least 50 species.

Did the study involve field work? ☐ Yes ☒ No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging