Article

Extant timetrees are consistent with a myriad of diversification histories

https://doi.org/10.1038/s41586-020-2176-1

Received: 14 September 2019

Accepted: 10 March 2020

Published online: 15 April 2020

Check for updates

Stilianos Louca[1,2]✉ & Matthew W. Pennell[3,4]✉

Time-calibrated phylogenies of extant species (referred to here as 'extant timetrees') are widely used for estimating diversification dynamics[1]. However, there has been considerable debate surrounding the reliability of these inferences[2–5] and, to date, this critical question remains unresolved. Here we clarify the precise information that can be extracted from extant timetrees under the generalized birth–death model, which underlies most existing methods of estimation. We prove that, for any diversification scenario, there exists an infinite number of alternative diversification scenarios that are equally likely to have generated any given extant timetree. These 'congruent' scenarios cannot possibly be distinguished using extant timetrees alone, even in the presence of infinite data. Importantly, congruent diversification scenarios can exhibit markedly different and yet similarly plausible dynamics, which suggests that many previous studies may have over-interpreted phylogenetic evidence. We introduce identifiable and easily interpretable variables that contain all available information about past diversification dynamics, and demonstrate that these can be estimated from extant timetrees. We suggest that measuring and modelling these identifiable variables offers a more robust way to study historical diversification dynamics. Our findings also make it clear that palaeontological data will continue to be crucial for answering some macroevolutionary questions.

A central challenge in evolutionary biology is to reconstruct rates of speciation and extinction over time[5]. Unfortunately, the majority of taxa that have ever lived have not left much trace in the fossil record, and the primary source of information on their past diversification dynamics therefore comes from extant timetrees. Many methods have been developed for extracting this information; most methods fit variants of a birth–death process[1,6]. Despite the popularity of these methods, which collectively have been used in thousands of studies[7–9], their reliability has been called into question by comparisons with fossil-based estimates[1,3,5,6,10]. The reasoning behind these critiques is that there may be insufficient information in extant timetrees to fully reconstruct historical diversification dynamics. However, this critical issue has remained unresolved; it is unknown precisely what information on speciation and extinction rates is contained in extant timetrees.

Here we present a definite answer to this question for the general stochastic birth–death process with homogeneous (that is, lineage-independent) rates, in which speciation ('birth') rates (λ) and extinction ('death') rates (μ) can vary over time, that underlies the majority of existing methods for reconstructing diversification dynamics from phylogenies[1]. We mathematically show that, for any given candidate birth–death model, there exists an infinite number of alternative birth–death models that can explain any extant timetree equally as well as can the candidate model. These alternative models may appear to be similarly plausible and yet exhibit markedly different features, such as different trends through time in both λ and μ. This severe ambiguity persists for arbitrarily large trees and cannot be resolved even with an infinite amount of data; it is thus impossible to design asymptotically consistent estimators for λ and μ. Using simulated and real timetrees as examples, we demonstrate how failing to recognize this issue can seriously mislead our inferences about past diversification dynamics. We present appropriately modified variables that are asymptotically identifiable and that contain all available information on historical diversification dynamics.

Lineages through time

An important feature of extant timetrees is the lineages-through-time curve (LTT), which counts the number of lineages at each time in the past that are represented by at least one sampled extant descending species in the tree. The likelihood of a tree under a given birth–death model, the LLT of the tree and the LTT that would be expected under the model are linked as follows. Any given combination of (potentially time-dependent) speciation and extinction rates (λ and μ, respectively) and the probability that an extant species will be included in the tree ('sampling fraction') (ρ) can be used to define a deterministic diversification process, in which the number of lineages through time no longer varies stochastically but instead according to a set of differential equations[6,11] (Supplementary Information section S.1). Given a number of extant sampled species (M_0), the LTT predicted by these

[1]Department of Biology, University of Oregon, Eugene, OR, USA. [2]Institute of Ecology and Evolution, University of Oregon, Eugene, OR, USA. [3]Biodiversity Research Centre, University of British Columbia, Vancouver, British Columbia, Canada. [4]Department of Zoology, University of British Columbia, Vancouver, British Columbia, Canada. ✉e-mail: louca.research@gmail.com; pennell@zoology.ubc.ca

differential equations (the deterministic LTT (dLTT)) corresponds to the LTT that is expected for trees generated by the original stochastic model[1,11]. The likelihood of a tree under a given birth–death model can be written purely in terms of the LTT of the tree and the dLTT of the model (see Supplementary Information section S.1.2 and ref. [12] for derivations). This means that any two models with the same dLTT (conditioned on M_o) yield identical likelihoods for the tree. We therefore call two models 'congruent' if they have the same dLTT for any given M_o. Any two models are either congruent or non-congruent, regardless of any particular data considered (Supplementary Information section S.1). The probability distribution of tree sizes generated by a model, when conditioned on the age of the stem or crown, is identical among congruent models (Supplementary Information section S.1.7). Hence, congruent models have equal probabilities of generating any given timetree, analogous to how congruent geometric objects exhibit similar properties (discussion in Supplementary Information section S.1.8). Although the mathematical relationship between the dLTT of a model and its likelihood has been known[12], its implications for macroevolutionary inference have remained unexamined and—as we show below—severely underestimated.

The breadth of congruent model sets

When seen as a random variable, extant timetrees have the same probability distribution under any two congruent models. Therefore, in the absence of further information, congruent models cannot possibly be distinguished solely on the basis of extant timetrees—neither through the likelihood nor any other test statistic. For any birth–death model, this leads to four important unresolved questions: how many alternative congruent models there are, how different these congruent models are from one another, how many of these congrument models correspond to plausible scenarios, and how these scenarios can be explored. To answer these questions, we present an alternative method for recognizing congruent models (full details are provided in Supplementary Information section S.1.1). Given a number of sampled species (M_o), the dLTT of a model is fully determined by its relative slope (hereafter, pulled speciation rate), denoted by $\lambda_p = -M^{-1}\mathrm{d}M/\mathrm{d}\tau$ (in which M is the dLTT, τ is time before present (or age) and p is a label (for 'pulled')). It can be shown that $\lambda_p = \lambda P$, in which $P(\tau)$ is the probability that a lineage extant at age τ survives until the present day and is included in the timetree. In the absence of extinction ($\mu = 0$) and under complete species sampling ($\rho = 1$), λ_p is identical to λ; however, in the presence of extinction λ_p is pulled downwards relative to λ at older ages, whereas under incomplete sampling λ_p is pulled downwards relative to λ near the present. Because the dLTT of a model is fully determined by λ_p and vice versa, two models are congruent if and only if they have the same λ_p at all ages. In a similar way, it can be shown that two models are congruent if and only if they have the same product $\rho\lambda_o$ (in which $\lambda_o = \lambda(0)$) and the same 'pulled diversification rate'[13], which is another composite variable and is defined as

$$r_p = \lambda - \mu + \frac{1}{\lambda}\frac{\mathrm{d}\lambda}{\mathrm{d}\tau} \quad (1)$$

The r_p is equal to the net diversification rate ($r = \lambda - \mu$) whenever λ is constant in time ($\mathrm{d}\lambda/\mathrm{d}\tau = 0$), but differs from r when λ varies with time.

We are now ready to examine the breadth of congruent model sets. We begin with a model with speciation rate $\lambda > 0$, extinction rate $\mu \geq 0$ and sampling fraction $\rho \in (0, 1]$. If we denote $\eta_o = \rho\lambda_o$, then for any alternative chosen extinction rate function $\mu^* \geq 0$ and any alternative assumed sampling fraction $\rho^* \in (0, 1]$, there exists a speciation rate function $\lambda^* > 0$ such that the alternative model (λ^*, μ^* and ρ^*) is congruent to the original model (λ, μ and ρ). In other words, regardless of the chosen μ^* and ρ^*, we can find a hypothetical λ^* that satisfies $\lambda^* - \mu^* + (1/\lambda^*)\mathrm{d}\lambda^*/\mathrm{d}\tau = r_p$ and $\rho^*\lambda^*(0) = \eta_o$. Indeed, to construct such a λ^* one merely needs to solve the following differential equation:

$$\frac{\mathrm{d}\lambda^*}{\mathrm{d}\tau} = \lambda^* \cdot (r_p - \lambda^* + \mu^*) \quad (2)$$

with initial condition $\lambda^*(0) = \eta_o/\rho^*$ (solution in Supplementary Information section S.1.4). The above observation implies that—starting from almost any birth–death model—we can generate an infinite number of alternative congruent models simply by modifying the extinction rate (μ) and/or the assumed sampling fraction (ρ). Alternatively, congruent models can be constructed by assuming various ratios of μ/λ (Supplementary Information section S.1.5). This set of congruent models (hereafter, the congruence class) is thus infinitely large. The congruence class can have an arbitrary number of dimensions (depending on restrictions imposed a priori on λ^* and μ^*), as μ^* could depend on an arbitrarily high number of free parameters.

As an illustration of these principles, the simulations in Fig. 1 show four markedly distinct and yet congruent models (pulled rates are shown in Extended Data Fig. 1). The first scenario exhibits a constant λ and a temporary spike in μ (that is, a mass extinction event), the second scenario exhibits a constant μ and a temporary drop in λ around the same time, the third scenario exhibits a mass extinction event at a completely different time and a fluctuating λ, and the fourth scenario exhibits an exponentially decaying μ and a fluctuating λ. These congruent scenarios were obtained simply by assuming alternative extinction rates, and a myriad of other congruent scenarios exist. Figure 2 shows a model with exponentially varying speciation and extinction rates, $\lambda = \alpha e^{\beta\tau}$ and $\mu = \gamma e^{\delta\tau}$, with α, β, γ and δ fitted to a timetree of 79,874 extant seed plant species[14] via maximum-likelihood methods. Simply by modifying the coefficient δ and choosing λ according to equation (2), one can obtain a similarly complex scenario with opposite trends over time (Fig. 2b). Similar examples can also be generated using more realistic speciation and extinction rates, three of which can be seen in Extended

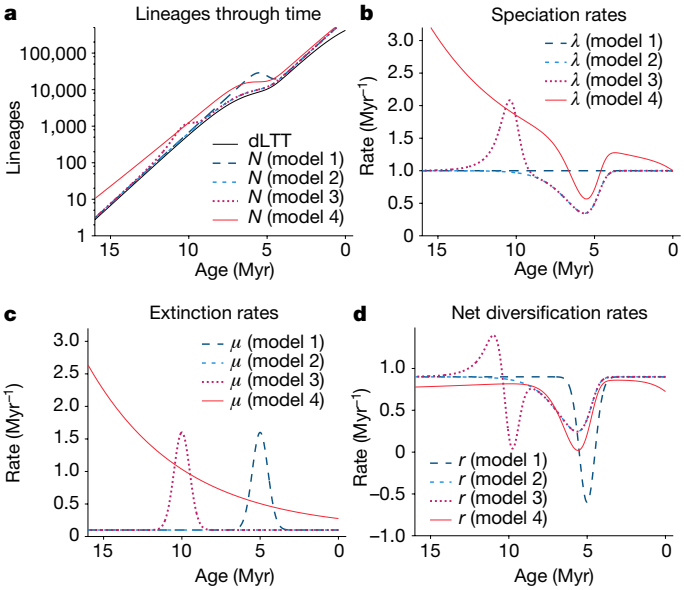


Fig. 1 | Illustration of congruent birth–death processes (simulations). Example of four hypothetical congruent—yet markedly different—birth–death models. All models exhibit the same dLTT, and would yield the same likelihood for any given extant timetree. **a**, dLTT and deterministic diversities (N) predicted by the models, plotted over age (time before present). **b**–**d**, Speciation rates (λ) (**b**), extinction rates (μ) (**c**) and net diversification rates ($r = \lambda - \mu$) (**d**) of the models. Myr, million years. For additional examples, see Extended Data Fig. 4.

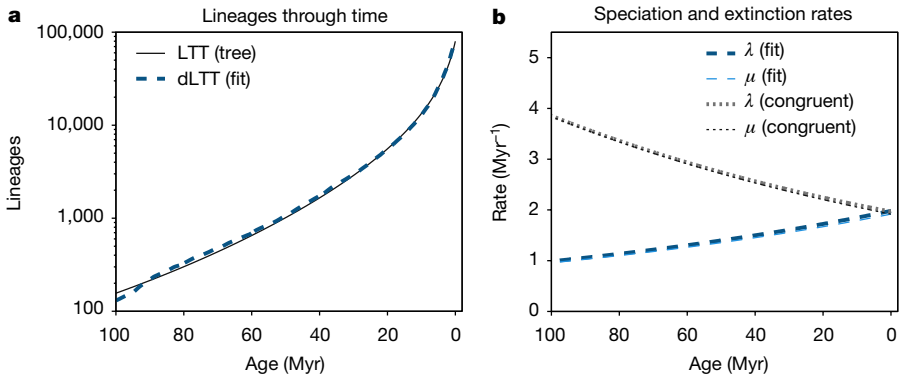


Fig. 2 | Illustration of congruent birth–death processes (real data). Birth–death model with exponentially varying λ and μ, fitted to an extant timetree of 79,874 seed plant species[14] over the past 100 Myr, compared to a congruent model obtained by simply modifying the exponential coefficient of μ. **a**, LTT of the tree, compared to the dLTT predicted by the two models. **b**, Speciation rates (λ) and extinction rates (μ) of the two models. In each model, μ is almost identical to λ. The two models cannot possibly be distinguished using extant timetrees alone. For additional examples, see Extended Data Fig. 2 and Supplementary Information sections S.10 and S.11.

Data Fig. 2 (based on data from ref. [15]), Extended Data Fig. 3 (based on data from ref. [10]) and Extended Data Fig. 4.

Such ambiguities have previously been observed in special cases[11,16]. For example, a previous study[11] recognized that a variable λ and constant μ can be exchanged for a constant λ and a variable μ to produce the same dLTT. Other work on constant-rate birth–death models has revealed that alternative combinations of time-independent λ, μ and ρ can yield the same likelihood for a tree[17,18]. Our work not only unifies these previous findings (which are all special cases of our general theory), but in fact reveals that vast (infinite-dimensional) expanses of model space are fundamentally indistinguishable even if ρ is known or all extant species have been sampled.

Implications

To estimate λ and μ, previous phylogenetic studies have imposed largely arbitrary constraints. For example, many studies assume that λ or μ vary exponentially through time[19]. However, this functional form is rarely justified biologically, and alternative functional forms of comparable simplicity and shape can be envisioned. Normally one expects that, with sufficient data, fitting any of these forms will lead to qualitatively similar trends and shapes. This expectation simply does not hold here, because the best-fitting representative within a given model set will generally only be the one closest to the congruence class of the true process, rather than closest to the true process itself (Fig. 3). Consequently, fitting alternative functional forms can result in markedly different inferences with alternative trends in λ and μ, even if each functional form used is in principle adequate for approximating the true historical λ and μ (examples are shown in Extended Data Fig. 5, Supplementary Information section S.10). This conclusion applies to almost any model set used in practice, including models in which λ and μ change at discrete time points[20]. Because any given true diversification history (even a relatively simple one) is unlikely to exactly match the particular functional form considered, fitting the latter may not even approximately yield the true diversification history. The existence of congruent scenarios can thus seriously alter macroevolutionary conclusions—for example, when assessing the influence of environmental factors on diversification dynamics (example shown in Supplementary Information section S.4, and further discussion in Supplementary Information section S.5). Our findings thus shed doubt over previous work on diversification dynamics that is based solely on extant timetrees, including some of the conclusions from work that we have coauthored[9,13]. Previous studies have underestimated this issue because they typically consider only a limited set of candidate models at a time, both when analysing real datasets and when assessing parameter identifiability via simulations; as a result, previous studies have been (un)lucky enough to not compare two models in the same congruence class (see Supplementary Information sections S.3 and S.7 for reasoning). We stress that common model selection methods that are based on parsimony or 'Occam's razor' (such as the Akaike information criterion[21]) generally cannot resolve these issues (Extended Data Fig. 6, details in Supplementary Information sections S.2 and S.10).

Ways forward

Our findings are analogous to classic results from coalescent theory in population genetics, in which many alternative models can give rise to the

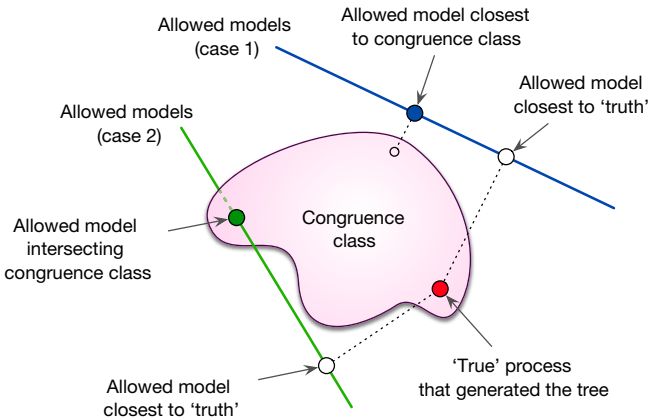


Fig. 3 | Conceptual implications. Conceptual illustration of the limited identifiability of a diversification process, assumed to be adequately described by some unknown birth–death model (red circle; hereafter 'true process'). The congruence class of the true process is shown as a sub-space comprising a continuum of alternative models (pink area). In practice, maximum-likelihood model selection is performed among a parameterized low-dimensional set of allowed models, the precise nature of which can vary from case to case (for example, depending on assumed functional forms for λ and μ, or the number of allowed rate shifts)[20]. The two continuous lines shown here represent two alternative cases of allowed model sets (for example, considered in two alternative studies), from within each the model closest to the truth is (ideally) sought. However, in each case likelihood-based model selection will converge towards the allowed model closest to the congruence class (blue and green filled circles), which in general is not the allowed model that is actually closest to the true process (white circles). This identifiability issue persists even for infinitely large datasets.

same drift process as the idealized Wright–Fisher model[22,23]. This realization was particularly important for the field: it focused the attention of researchers on the dynamics of the effective population size, an identifiable parameter, rather than on actual (but non-identifiable) historical demography. Similarly, congruent birth–death models can be defined in terms of λ_p or—equivalently—in terms of r_p and $\rho\lambda_o$, all of which are identifiable provided sufficient data[13]. Each congruence class contains exactly one model with $\mu = 0$ and $\rho = 1$, which is also the model in which $\lambda = \lambda_p$; hence, the pulled speciation rate can be interpreted as the speciation rate that generates the dLTT of the congruence class in the absence of extinctions and under complete species sampling. In other words, λ_p can be seen as the 'effective' speciation rate that fully explains the shape of the LTT of the tree. Similarly, each congruence class contains models with time-independent λ, and for these models $r_p = r$; therefore, the pulled diversification rate can be interpreted as the effective net diversification rate if λ was time-independent.

Fossil data could help to resolve the ambiguities highlighted here[24,25], and biological knowledge could, in principle, also help to reduce ambiguities. For example, if ρ and μ are somehow known from other sources, the congruence class collapses to a unique diversification scenario (Extended Data Fig. 7). Nevertheless, for many taxa the fossil record remains scarce and ambiguous, and our general understanding of what constitutes a plausible diversification scenario is poorly developed. Rather than attempting to estimate λ and μ, one can estimate λ_p, r_p and $\rho\lambda_o$ (and λ_o, if ρ is known)—for example, by using likelihood methods[26] (Extended Data Fig. 8, Supplementary Information section S.9). Previous work[13] has shown that r_p can indeed yield insight into diversification dynamics and help to detect major transitions over time (Supplementary Information section S.8), as changes in r_p necessarily imply changes in λ and/or μ. Through λ_p, it also becomes possible to simulate and analyse diversification models with substantially simplified mathematical tools, as any model is congruent to a model with speciation rate λ_p, zero extinction and complete species sampling[27]. Reciprocally, many existing estimation tools can be used to estimate λ_p and r_p by constraining μ to be zero or λ to be time-independent, respectively. Depending on the situation, other invariants of congruence classes may also have advantages, such as the 'coalescent density' introduced by ref. [12], which permits an elegant description of the distribution of branching ages (see Supplementary Information section S.8 for further details).

Conclusions

Without further information or biologically well-justified constraints, in general extant timetrees alone cannot be used to reliably infer speciation rates (except for the present day), extinction rates or net diversification rates. Consequently, correlations between λ, μ or r and fluctuating environmental factors (such as temperature) also cannot be reliably inferred, neither when λ, μ or r are first estimated and then related to the environmental factors nor if λ, μ and r are expressed as parameterized functions of the environmental factors and then fitted to the timetree (Supplementary Information section S.5), because different parameterizations can lead to completely different inferences. Our findings could explain why diversification dynamics observed in the fossil record often contradict inferences based on phylogenetics[1,3,5,6,10], although other explanations have also been proposed[28,29]. It is possible that similar major identifiability issues may also be hiding in other evolutionary reconstruction methods based on extant organisms alone, but this remains to be examined. On a more positive note, we have resolved a long-standing debate and precisely clarified what information can be extracted from extant timetrees alone, formulated in terms of easily interpretable and identifiable variables.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2176-1.

1. Morlon, H. Phylogenetic approaches for studying diversification. *Ecol. Lett.* **17**, 508–525 (2014).
2. Quental, T. B. & Marshall, C. R. Extinction during evolutionary radiations: reconciling the fossil record with molecular phylogenies. *Evolution* **63**, 3158–3167 (2009).
3. Quental, T. B. & Marshall, C. R. Diversity dynamics: molecular phylogenies need the fossil record. *Trends Ecol. Evol.* **25**, 434–441 (2010).
4. Liow, L. H., Quental, T. B. & Marshall, C. R. When can decreasing diversification rates be detected with molecular phylogenies and the fossil record? *Syst. Biol.* **59**, 646–659 (2010).
5. Marshall, C. R. Five palaeobiological laws needed to understand the evolution of the living biota. *Nat. Ecol. Evol.* **1**, 0165 (2017).
6. Morlon, H., Parsons, T. L. & Plotkin, J. B. Reconciling molecular phylogenies with the fossil record. *Proc. Natl Acad. Sci. USA* **108**, 16327–16332 (2011).
7. Rabosky, D. L. et al. An inverse latitudinal gradient in speciation rate for marine fishes. *Nature* **559**, 392–395 (2018).
8. Condamine, F. L., Rolland, J. & Morlon, H. Assessing the causes of diversification slowdowns: temperature-dependent and diversity-dependent models receive equivalent support. *Ecol. Lett.* **22**, 1900–1912 (2019).
9. Henao Diaz, L. F., Harmon, L. J., Sugawara, M. T. C., Miller, E. T. & Pennell, M. W. Macroevolutionary diversification rates show time dependency. *Proc. Natl Acad. Sci. USA* **116**, 7403–7408 (2019).
10. Steeman, M. E. et al. Radiation of extant cetaceans driven by restructuring of the oceans. *Syst. Biol.* **58**, 573–585 (2009).
11. Kubo, T. & Iwasa, Y. Inferring the rates of branching and extinction from molecular phylogenies. *Evolution* **49**, 694–704 (1995).
12. Lambert, A. & Stadler, T. Birth–death models and coalescent point processes: the shape and probability of reconstructed phylogenies. *Theor. Popul. Biol.* **90**, 113–128 (2013).
13. Louca, S. et al. Bacterial diversification through geological time. *Nat. Ecol. Evol.* **2**, 1458–1467 (2018).
14. Smith, S. A. & Brown, J. W. Constructing a broadly inclusive seed plant phylogeny. *Am. J. Bot.* **105**, 302–314 (2018).
15. Alroy, J. Colloquium paper: dynamics of origination and extinction in the marine fossil record. *Proc. Natl Acad. Sci. USA* **105** (Suppl 1), 11536–11542 (2008).
16. Stadler, T. Simulating trees with a fixed number of extant species. *Syst. Biol.* **60**, 676–684 (2011).
17. Stadler, T. On incomplete sampling under birth–death models and connections to the sampling-based coalescent. *J. Theor. Biol.* **261**, 58–66 (2009).
18. Stadler, T. & Steel, M. Swapping birth and death: symmetries and transformations in phylodynamic models. *Syst. Biol.* **68**, 852–858 (2019).
19. Rabosky, D. L. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS ONE* **9**, e89543 (2014).
20. Stadler, T. Mammalian phylogeny reveals recent diversification rate shifts. *Proc. Natl Acad. Sci. USA* **108**, 6187–6192 (2011).
21. Akaike, H. Likelihood of a model and information criteria. *J. Econom.* **16**, 3–14 (1981).
22. Möhle, M. Robustness results for the coalescent. *J. Appl. Probab.* **35**, 438–447 (1998).
23. Sjödin, P., Kaj, I., Krone, S., Lascoux, M. & Nordborg, M. On the meaning and existence of an effective population size. *Genetics* **169**, 1061–1070 (2005).
24. Heath, T. A., Huelsenbeck, J. P. & Stadler, T. The fossilized birth–death process for coherent calibration of divergence-time estimates. *Proc. Natl Acad. Sci. USA* **111**, E2957–E2966 (2014).
25. Stadler, T., Gavryushkina, A., Warnock, R. C. M., Drummond, A. J. & Heath, T. A. The fossilized birth-death model for the analysis of stratigraphic range data under different speciation modes. *J. Theor. Biol.* **447**, 41–55 (2018).
26. Louca, S. & Doebeli, M. Efficient comparative phylogenetics on large trees. *Bioinformatics* **34**, 1053–1055 (2018).
27. Louca, S. Simulating trees with millions of species. *Bioinformatics* https://doi.org/10.1093/bioinformatics/btaa031 (2020).
28. Rabosky, D. L. Heritability of extinction rates links diversification patterns in molecular phylogenies and fossils. *Syst. Biol.* **58**, 629–640 (2009).
29. Silvestro, D., Warnock, R. C. M., Gavryushkina, A. & Stadler, T. Closing the gap between palaeontological and neontological speciation and extinction rate estimates. *Nat. Commun.* **9**, 5237 (2018).

Methods

No statistical methods were used to predetermine sample size. Thorough mathematical derivations and computational details are provided in Supplementary Information sections S.1–S.11.

Reporting summary

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

Data availability

No new data were generated for this manuscript. All phylogenetic datasets used as examples have previously been published previously, and are cited where appropriate.

Code availability

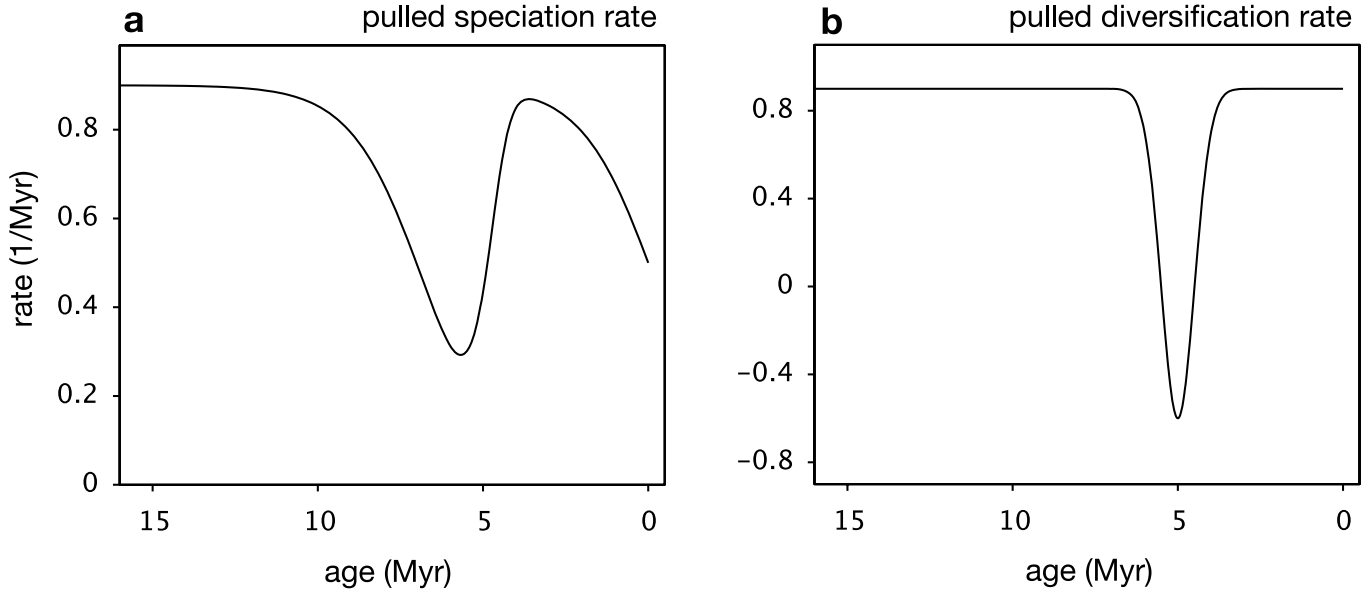
Computational methods used for this article—including functions for simulating birth–death models, for constructing models within a given congruence class, for calculating the likelihood of a congruence class and for directly fitting congruence classes (either in terms of λ_p or in terms of r_p and $\rho\lambda_o$) to extant timetrees—are implemented in the R package castor v.1.5.5, which is available from The Comprehensive R Archive Network at https://cran.r-project.org/package=castor.

Acknowledgements S.L. was supported by a start-up grant by the University of Oregon. M.W.P. was supported by an NSERC Discovery Grant. We thank L. Harmon, S. Otto, A. MacPherson, D. Schluter, T. J. Davies, M. Whitlock, L. F. Henao Diaz, K. Kaur, J. Uyeda, D. Caetano, J. Rolland, L. Parfrey and A. Mooers for insightful comments on this work.

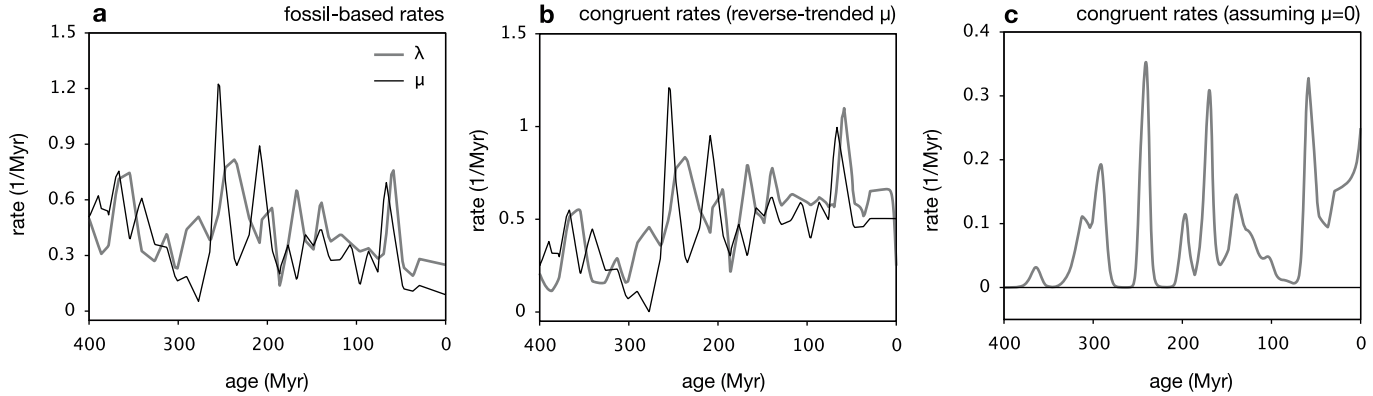
Author contributions S.L. performed the mathematical calculations and computational analyses. S.L. and M.W.P. conceived the project and contributed to the writing of the manuscript.

Competing interests The authors declare no competing interests.

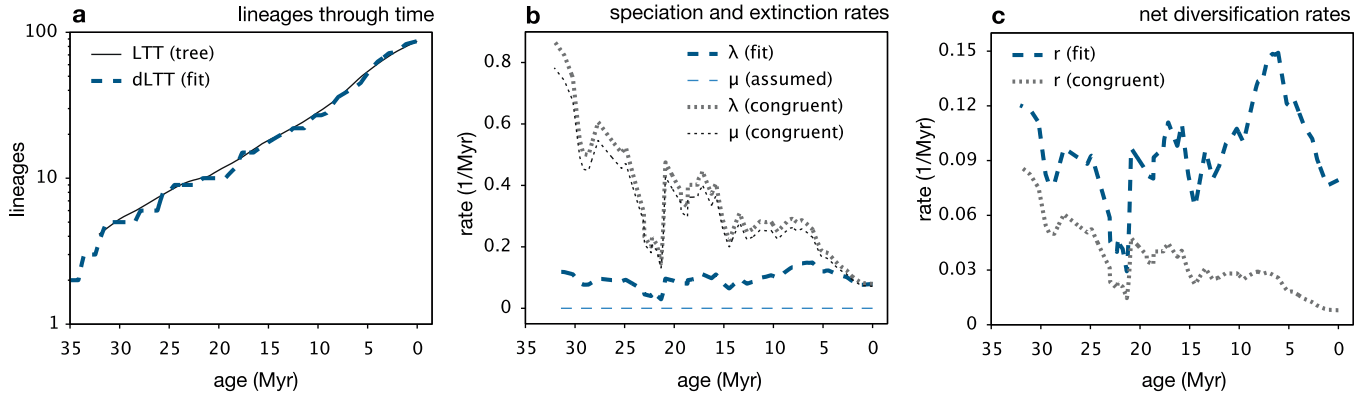
Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2176-1.
Correspondence and requests for materials should be addressed to S.L. or M.W.P.
Peer review information *Nature* thanks Lee Hsiang Liow, Antonis Rokas, Mike Steel and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.
Reprints and permissions information is available at http://www.nature.com/reprints.



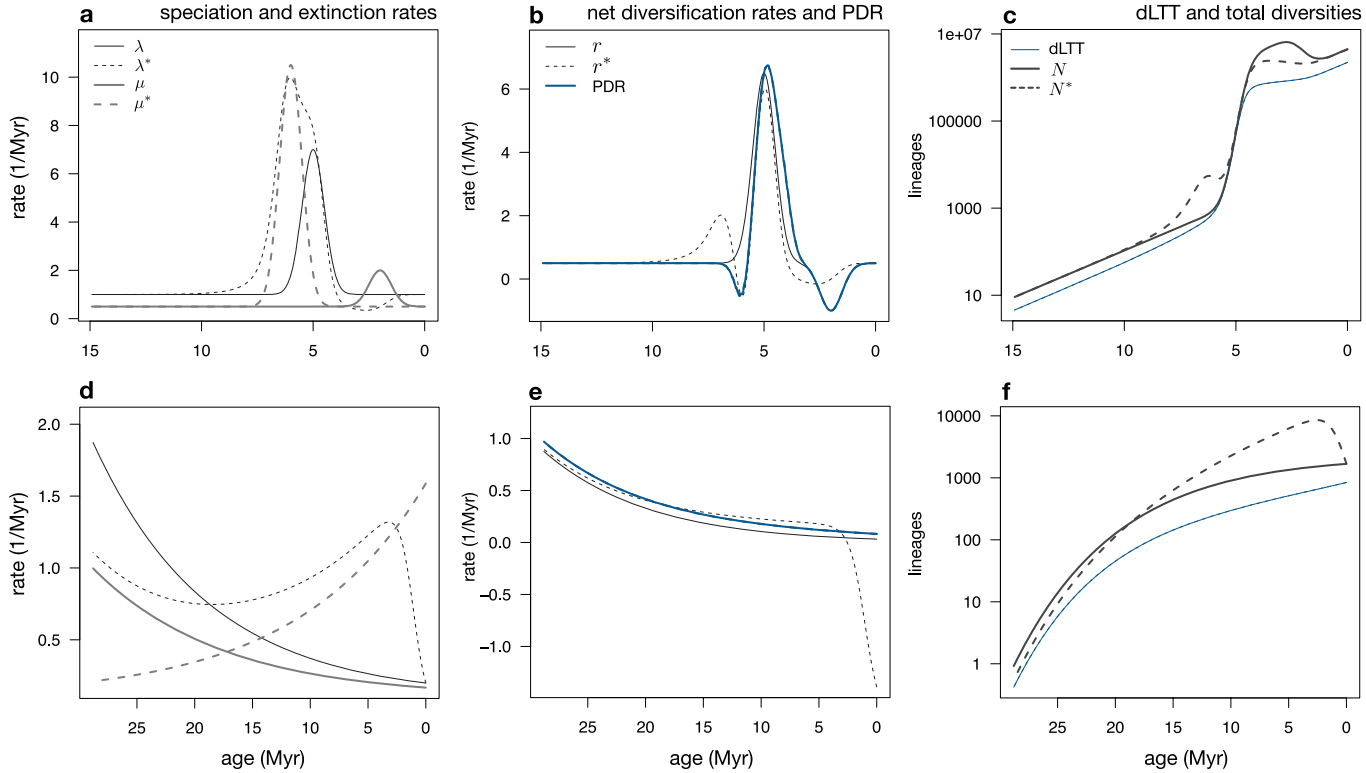
Extended Data Fig. 1 | Pulled speciation and diversification rates. a, **b**, Pulled speciation rate (**a**) and pulled diversification rate (**b**) of the four congruent models shown in Fig. 1.



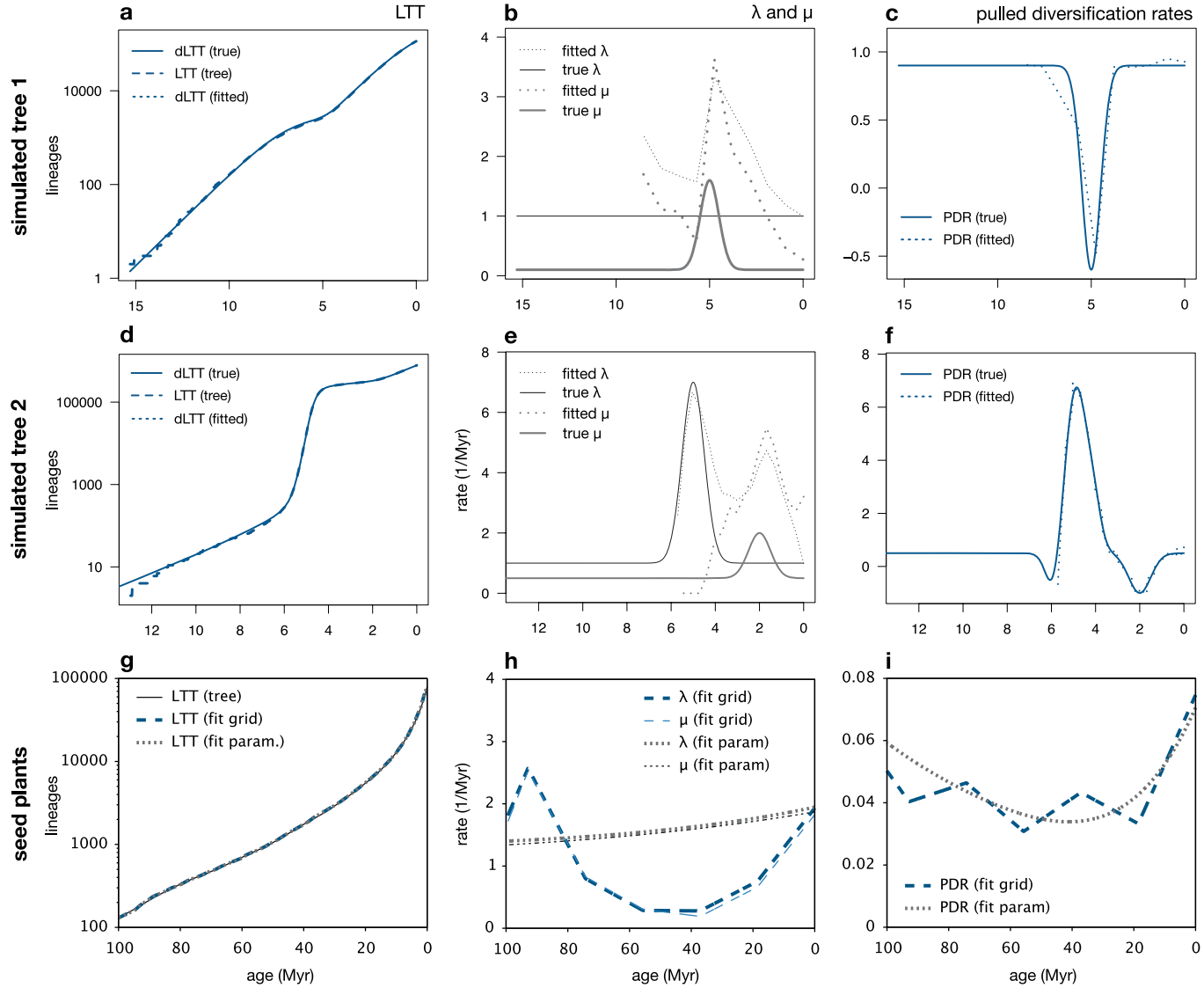
Extended Data Fig. 2 | Illustration of congruent birth–death processes (fossil data). a, Origination and extinction rates of marine invertebrate genera, estimated from fossil data. **b**, Congruent scenario to that in **a**, obtained by reversing the linear trend of μ (that is, fitting a linear curve to the original μ, and then subtracting that curve twice) and adjusting λ according to equation (2). **c**, Congruent scenario to that in **a**, assuming an extinction rate of zero. Further details are provided in Supplementary Information section S.10.



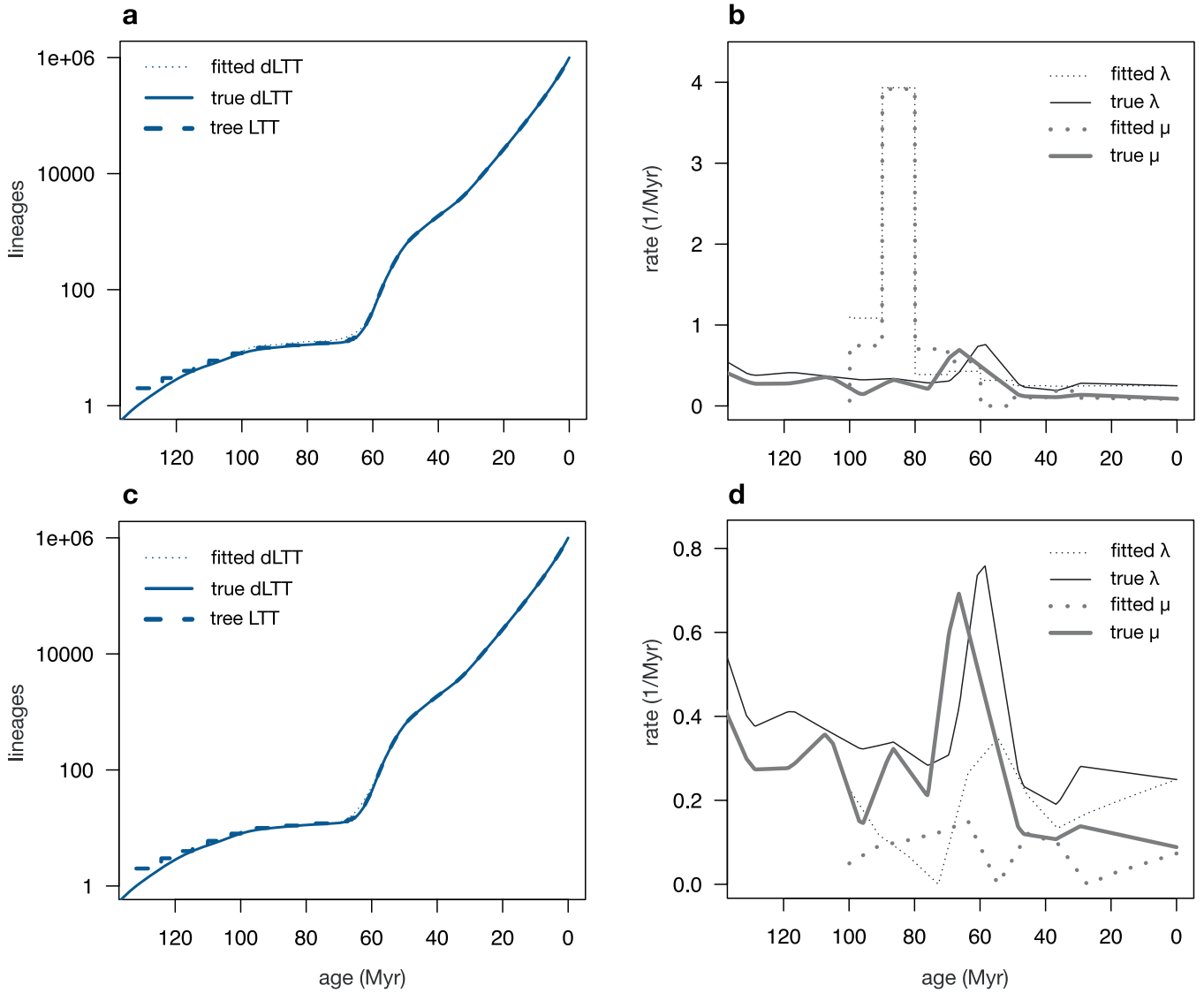
Extended Data Fig. 3 | Previous studies are likely to have over-interpreted phylogenetic data. Time-dependent birth–death model fitted to a nearly complete extant timetree of the Cetacea, under the assumption of extinction rates of zero ($\mu = 0$), compared to a congruent model in which the extinction rate is close to the speciation rate ($\mu = 0.9\lambda$). **a**, LTT of the tree, compared to the dLTT predicted by the two models. **b**, Speciation rates (λ) and extinction rates (μ) of the two models. **c**, Net diversification rates ($r = \lambda - \mu$) of the two models. Further details are provided in Supplementary Information section S.4.



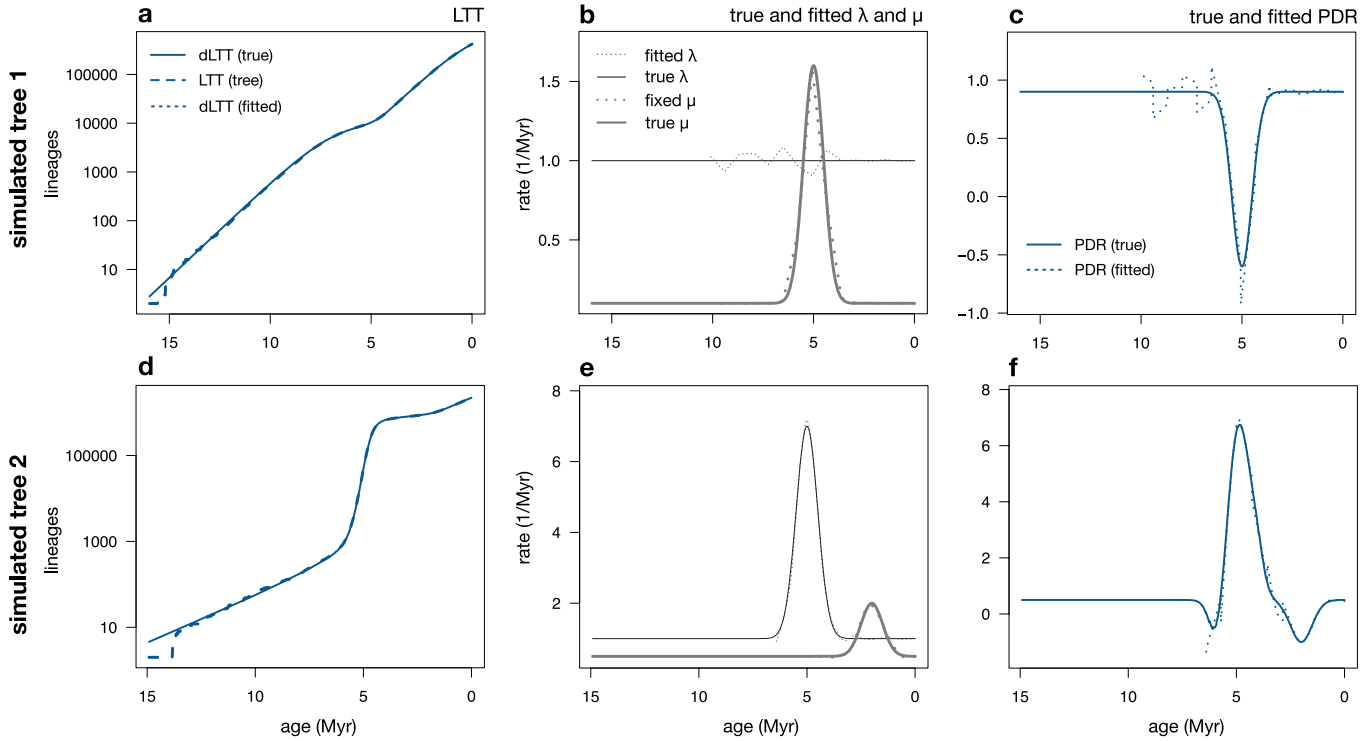
Extended Data Fig. 4 | Additional examples of congruent birth–death processes. a–**c**, Example of two congruent—yet markedly different—birth–death models. Both models exhibit a temporary spike in the extinction rate and a temporary spike in the speciation rate; however, the timings of these events differ substantially between the two models. Both models exhibit the same dLTT and the same pulled diversification rate (r_p) and would yield identical likelihoods for any given extant timetree. **a**, Speciation rates (λ and λ^*) and extinction rates (μ and μ^*) of the two models, plotted over time. Continuous curves correspond to the first model, and dashed curves correspond to the second model. **b**, Net diversification rates (r and r^*) and pulled diversification rate (r_p) of the two models. **c**, dLTT and deterministic total diversities (N and N^*) predicted by the two models. **d**–**f**, Another example of two congruent models. In the first model, the speciation and extinction rates both decrease exponentially over time, whereas in the second model the extinction rate increases exponentially over time and the speciation rate exhibits variable directions of change over time. In all models, the sampling fraction is $\rho = 0.5$.



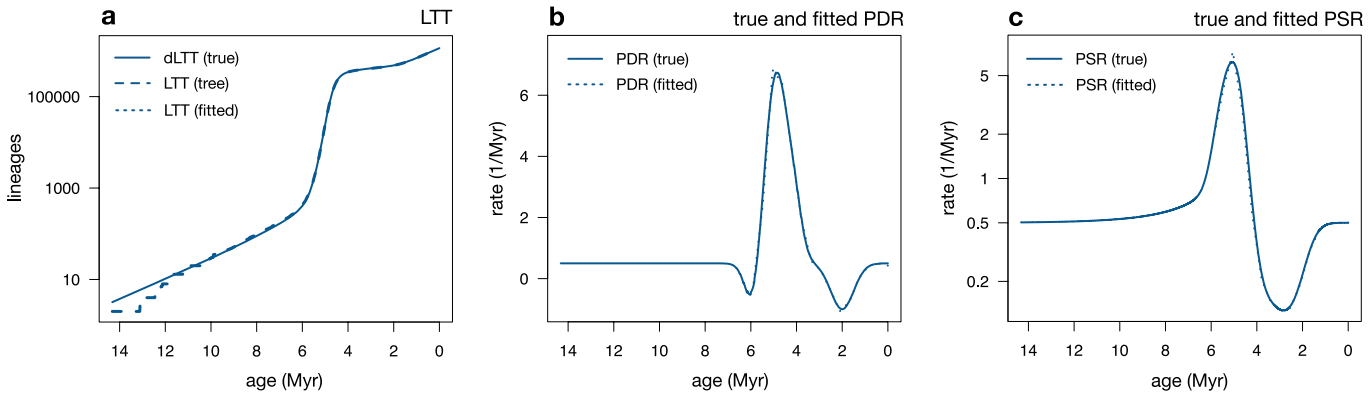
Extended Data Fig. 5 | Identifiability issues persist in large trees. a–**c**, Diversification analysis of a timetree (about 114,000 tips) simulated from a birth–death process that exhibits a mass extinction event at around 5 Myr before present. **a**, LTT of the generated tree (long-dashed curve), dLTT of the true model that generated the tree (continuous curve) and dLTT of a maximum-likelihood fitted model (short-dashed curve) are shown. The fitted dLTT is practically identical to the true dLTT and thus is covered by the latter. **b**, True speciation and extinction rates (continuous curves), compared to fitted speciation and extinction rates (dashed curves). There is considerable disagreement between the fitted and true λ and μ, despite the fact that the allowed model set could—in principle—approximate the true rates reasonably well. **c**, Pulled diversification rate (PDR) of the true model (continuous curve), compared to the pulled diversification rate of the fitted model (dashed curve). **d**–**f**, Diversification analysis of a timetree (about 785,000 tips) simulated from a birth–death process that exhibits a rapid radiation event at around 5 Myr before present and a mass extinction event at around 2 Myr before present. **d**–**f** are analogous to **a**–**c**. There is considerable disagreement between the fitted and true λ and μ, despite the fact that the allowed model set could—in principle—approximate the true rates reasonably well. Extended Data Figure 7 provides the fitting results when μ is fixed to its true value. **g**–**i**, Diversification analyses of an extant timetree of 79,874 seed plant species, performed either by fitting λ and μ on a grid of discrete time points or by fitting the parameters of generic polynomial or exponential functions for λ and μ. **g**, LTT of the tree, dLTT of the grid-fitted model and dLTT of the fitted parametric model. **h**, Speciation and extinction rates predicted by the grid-fitted model or the fitted parametric model. **i**, Pulled diversification rate predicted by the grid-fitted model and the fitted parametric model. Further details are provided in Supplementary Information sections S.10 and S.11.



Extended Data Fig. 6 | Identifiability issues cannot be resolved with the Akaike information criterion. Maximum-likelihood birth–death models fitted to a tree comprising 1,000,000 tips, simulated on the basis of the origination and extinction rates of marine invertebrate genera estimated from fossil data. Top row, maximum-likelihood-fitted piecewise constant model (also known as birth–death–shift model), with grid size ($n = 11$) chosen by minimizing the Akaike information criterion (AIC). Bottom row, maximum-likelihood-fitted piecewise linear model, with grid size ($n = 12$) chosen by minimizing the AIC. Left column, dLTTs of the fitted models compared to the true dLTT and the LTT of the tree. Right column, fitted speciation and extinction rates, compared to the true rates used to generate the tree. In both cases, the maximum-likelihood models poorly reflect the true rates despite a near-perfect match of the LTT, even when the complexity of the models was optimized on the basis of the AIC. For further details, see Supplementary Information sections S.2 and S.10.



Extended Data Fig. 7 | Estimating λ when μ and ρ are fixed or known. a–**c**, Example analysis of a simulated extant timetree (about 114,000 tips) that exhibits a mass extinction event at around 5 Myr before present. A birth–death model was fitted while fixing μ and ρ to their true values; λ was fitted at 15 discrete time points. **a**, LTT of the generated tree (long-dashed curve), dLTT of the true model that generated the tree (continuous curve) and dLTT of a maximum-likelihood fitted model (short-dashed curve). The fitted dLTT is practically identical to the true dLTT, and is thus covered by the latter. **b**, True speciation and extinction rates (continuous curves), along with the fitted speciation rate and fixed extinction rate (dashed curves). **c**, Pulled diversification rate of the true model (r_p, continuous curve), compared to the pulled diversification rate of the fitted model (dashed curve). **d**–**f**, Example analysis of a simulated extant timetree (about 785,000 tips) that exhibits a rapid radiation event at about 5 Myr before present and a mass extinction event at about 2 Myr before present. A birth–death model was fitted similarly to the example shown in **a**–**c**, and **d**–**f** are analogous to **a**–**c**. In both cases, rate estimation was restricted to ages at which the LTT included at least 500 lineages. Further details are provided in Supplementary Information section S.10.



Extended Data Fig. 8 | Fitting congruence classes instead of models. Analysis of an extant timetree generated by a birth–death model that exhibits a temporary rapid radiation event about 5 Myr before present and a mass extinction event about 2 Myr before present. A congruence class was fitted to the timetree either in terms of the pulled diversification rate (r_p) and the product $\rho\lambda_o$, or in terms of the pulled speciation rate (PSR) (λ_p). **a**, LTT of the tree (long-dashed curve), together with the dLTT of the true model (continuous curve) and the dLTT of the fitted congruence classes (short-dashed curve); in both cases, the fitted dLTT was almost identical to the true dLTT, and is thus completely covered by the latter. **b**, Pulled diversification rate of the true model (continuous curve), compared to the fitted pulled diversification rate (short-dashed curve). **c**, Pulled speciation rate of the true model (continuous curve), compared to the fitted pulled speciation rate (short-dashed curve). The pulled diversification rate and pulled speciation rate were fitted via maximum-likelihood methods, allowing the pulled diversification rate or pulled speciation rate to vary freely at 15 discrete equidistant time points. Further details are provided in in Supplementary Information section S.9.

nature research

Corresponding author(s): Stilianos Louca, Matthew W. Pennell

Last updated by author(s): Feb 27, 2020

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	No software was used to collect data. Simulated data was generated using the R package "castor", which is freely available on CRAN.
Data analysis	Computational methods used for this article, including functions for simulating birth-death models, for constructing models within a given congruence class, for calculating the likelihood of a congruence class, and for directly fitting congruence classes to extant timetrees, are implemented in the R package "castor", which is freely available on CRAN.

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

No raw data was generated for this manuscript. All phylogenetic datasets used have been published previously and are cited in the manuscript where appropriate.

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 use mathematical proofs, numerical simulations and analysis of previously published phylogenetic data to investigate the identifiability of past diversification dynamics from extant time-calibrated phylogenies.
Research sample	The following previously published datasets were used: Origination and extinction rates of marine invertebrate genera, estimated from the fossil record by (Alroy et al. 2008). Time-calibrated phylogeny of 79874 extant seed plant species, published by Smith and Brown (2018). Time-calibrated phylogeny of the Cetacea, published by Steeman et al. (2009).
Sampling strategy	Sample sizes (i.e., phylogeny sizes) were chosen as large as possible in our examples, to demonstrate that the identifiability issues discussed in the paper persist even for massive datasets. For simulated trees we used very large numbers of tips (larger than commonly seen in the literature), again to illustrate that our conclusions remain valid even for massive data sets.
Data collection	No new data were collected.
Timing and spatial scale	No new data were collected.
Data exclusions	No data was excluded from the analysis.
Reproducibility	Our full mathematical proofs and numerical procedures are described in detail in the manuscript and supplemental material. All new simulation code is published through the R package "castor" (version 1.5.5), which is freely available on CRAN.
Randomization	This is not relevant to our study, as no experiments were performed.
Blinding	Blinding was not relevant to our study, as no experiments were performed.

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