METHOD

Global Ecology and Biogeography | A Journal of Macroecology

WILEY

Challenges in estimating species' age from phylogenetic trees

Carlos Calderón del Cid[1,2] | Torsten Hauffe[2] | Juan D. Carrillo[2] |
Michael R. May[3] | Rachel C. M. Warnock[4] | Daniele Silvestro[2,5]

[1]Laboratório de Ecologia Espacial, Instituto de Biologia, Universidade Federal da Bahia, Salvador, Brazil

[2]Department of Biology, University of Fribourg, Switzerland and Swiss Institute of Bioinformatics, Fribourg, Switzerland

[3]Department of Evolution and Ecology, University of California Davis, Davis, California, USA

[4]GeoZentrum Nordbayern, Friedrich-Alexander Universität Erlangen-Nürnberg, Erlangen, Germany

[5]Department of Biological and Environmental Sciences, Gothenburg Global Biodiversity Centre, University of Gothenburg, Gothenburg, Sweden

Correspondence
Carlos Calderón del Cid, Laboratório de Ecologia Espacial, Instituto de Biologia, Universidade Federal da Bahia, Salvador CEP 40170 110, Bahia, Brazil.
Email: caldecid@gmail.com

Abstract

Aim: Species age, the elapsed time since origination, can give insight into how species longevity might influence eco-evolutionary dynamics, which has been hypothesized to influence extinction risk. Traditionally, species' ages have been estimated from fossil records. However, numerous studies have recently used the branch lengths of time-calibrated phylogenies as estimates of the ages of extant species. This approach poses problems because phylogenetic trees only contain direct information about species identity at the tips and not along the branches. Here, we show that incomplete taxon sampling, extinction and different assumptions about speciation modes can significantly alter the relationship between true species age and phylogenetic branch lengths, leading to high error rates. We found that these biases can lead to erroneous interpretations of eco-evolutionary patterns derived from comparing phylogenetic age and other traits, such as extinction risk.

Innovation: For bifurcating speciation, the default assumption in most analyses of species age, we propose a probabilistic approach based on the properties of a birth–death process to improve the estimation of species ages. Our approach can reduce the error by one order of magnitude under cases of high extinction and a high percentage of unsampled extant species.

Main conclusion: Our results call for caution in interpreting the relationship between phylogenetic ages and eco-evolutionary traits, as this can lead to biased and erroneous conclusions. We show that, under the assumption of bifurcating speciation, we can obtain unbiased approximations of species age by combining information from branch lengths with the expectations of a birth–death process.

KEYWORDS
evolutionary history, extinction rates, incomplete sampling, phylogeny, simulations, speciation modes

---

1 | INTRODUCTION

The estimation of species age, or the elapsed time since species origin, is important for evaluating mechanisms that link species longevity with eco-evolutionary processes (Benton, 2013; Swenson, 2019). For instance, age-dependent extinction hypotheses test the relationship between species age and extinction probability, assessing whether extinction rates differ between young and old species (Balmford, 1996; Calderón del Cid et al., 2024; Eldredge et al., 2005; Pearson, 1995). Likewise, species age could be a measure of colonization time, especially in island systems (Tanentzap et al., 2015) or during biotic invasions triggered by geological events, such as the formation of the Central American Isthmus for the Great American Interchange (Carrillo et al., 2015, 2020). Species age is measured in the fossil record through different statistical and probabilistic approaches, based mostly on a taxon's stratigraphic duration (i.e., the time between the first and last appearance of a taxon in the fossil record) (Foote, 1996; Foote & Raup, 1996). Several of these approaches consider differences in fossil sampling and temporal resolution (e.g., Alroy et al., 2001; Silvestro et al., 2019). Species ages estimated from palaeobiological data offer a reliable measure of species' temporal duration, which can be used in macroevolutionary studies (Benton, 2016; Silvestro et al., 2020; Van Valen, 1973). More recently, several studies have used the length of terminal branches in time-calibrated phylogenies as a proxy for the age of extant species, an approximation that we hereafter refer to as "phylogenetic age" (Alzate et al., 2023; Davies et al., 2011; Gaston & Blackburn, 1997; Johnson et al., 2002; Pie & Caron, 2023; Sonne et al., 2022; Tanentzap et al., 2020; Verde Arregoitia et al., 2013). These phylogenetic ages have then been used as the basis to test for links between species age and current extinction risks (Tanentzap et al., 2020; Verde Arregoitia et al., 2013) and to assess various correlations with evolutionary, biogeographical and ecological patterns in living species (Alzate et al., 2023; Freer et al., 2022; Guo et al., 2024; Kennedy et al., 2022; Pie & Caron, 2023). In a similar approach at larger scales, clade ages, or the age of the most recent common ancestor of all species within a clade, have been used in correlation with latitude, elevational gradients and diversity to explain the tropical niche conservatism hypothesis (Hawkins et al., 2011; Kennedy et al., 2014; Qian & Ricklefs, 2016; Wiens & Donoghue, 2004).

While several studies have used phylogenetic age at face value for species age (e.g., Johnson et al., 2002; Tanentzap et al., 2020; Verde Arregoitia et al., 2013), their potential deviation from the true species ages remains unclear. Specifically, we identify three non-mutually exclusive shortfalls that can lead to over- or underestimation of species ages. First, incomplete sampling of extant species, either due to incomplete species sampling or linked to species still unknown to science, can bias phylogenetic age estimation by artificially increasing the length of terminal branches (Heath et al., 2008; Mynard et al., 2023).

Second, extinction events will mask branching events in phylogenetic trees of extant species (Harvey et al., 1994; Nee & May, 1997). Even in phylogenetic trees that include extinct taxa, the incompleteness of the fossil record will inevitably lead to missing lineages and incorrect topologies. Unsampled extant and extinct species in the phylogeny result in an inflation of the length of terminal branches leading to sampled species (i.e., the tips of the tree), thus altering phylogenetic species ages. For instance, if the extinct species of the *Homo* genus are not included in a phylogeny, the phylogenetic age of *Homo sapiens* is approximately 10 million years, that is, the age of the last common ancestor with its sister species, the chimpanzee (Rivas-Gonzáles et al., 2023). This estimate exceeds the age of the oldest known fossil of modern humans (i.e., *Homo sapiens*) by two orders of magnitude (Callaway, 2017).

The third shortfall is that the tree alone does not contain information about the underlying speciation mode and does not include species labels along its branches, meaning only the tips can be unequivocally assigned to a named species (Losos & Glor, 2003). Alternative speciation modes have been discussed in the literature, reflecting different biological processes and species concepts, including bifurcating, budding and anagenetic speciation (Foote, 1996; Silvestro et al., 2018). These modes define the relationship between the ancestral species and its descendants, thus contributing to determining species ages (Rosenblum et al., 2012; Wagner et al., 1995) (Figure 1). Most phylogenetic trees are depicted in a rectangular shape, where the two descending lineages split symmetrically from an ancestral lineage, suggesting a bifurcating speciation mode where two new species replace the ancestral lineage (Baum et al., 2005; Caetano & Quental, 2023). However, the often-unstated assumption of virtually all birth–death processes used to model phylogenetic branching times is that speciation occurs as a budding process, with a speciation event leading to a single new species and the survival of the parent species, even though we cannot determine which descendant branch is the new species (Gernhard, 2008; Nee et al., 1994; Stadler, 2013; but see Stadler et al., 2018). Anagenetic speciation, in contrast, does not lead to a branching event and is, therefore, not ordinarily visible on a phylogenetic tree.

All speciation modes may reflect plausible biological processes and evidence for each mode has been found in the fossil record (Foote, 1996; Simpson, 1984) and among extant species (Skeels & Cardillo, 2019). Bifurcating speciation can be linked with vicariance or allopatric speciation (Willmann & Meier, 2000) and the Hennigian species concept considers all speciation events as bifurcating (Hennig, 1999). Under this speciation mode, phylogenetic age equals species ages when there is no extinction and all extant species are sampled. Budding speciation can be interpreted as the result of parapatric, peripatric, or founder-event speciation (Anacker & Strauss, 2014; Caetano & Quental, 2023) and is related to the evolutionary species concept, which considers any species as a descendant of an ancestral lineage with its own evolutionary identity to be a valid species (Simpson, 1951: but see Simpson, 1961). In a phylogenetic tree, we typically lack information about which of the two descendent branches is the new species after a budding speciation event (but see Aze et al., 2011). Thus, in the absence of any extinction or missing species, the phylogenetic age of one sister species will equal its species age, while the other will be older, but without

FIGURE 1 Impact of speciation mode and extinction on species age. For the same hypothetical time-calibrated phylogeny of extant species, the continuation of the same colour indicates the same species, solid and dashed lines represent extant and extinct species, respectively, under different speciation modes and extinction scenarios. Numbers display the resulting true age of the respective species in million years (myr).
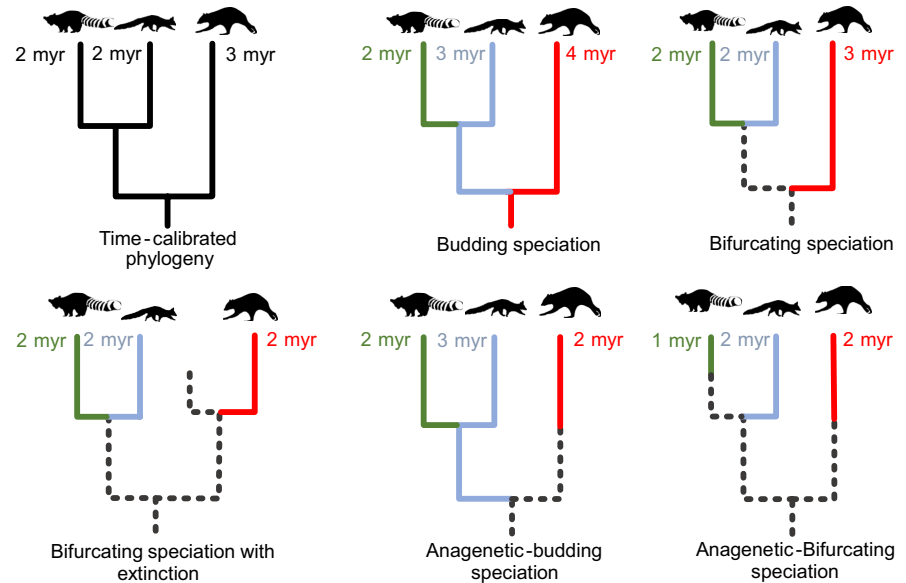
the possibility to identify which one is which. Anagenetic speciation is not associated with a branching event but can be used to describe different species or morphospecies (Emerson & Patiño, 2018) delimited by substantial phenotypic change occurring along a lineage (Roopnarine et al., 1999). This process will result in a higher phylogenetic age than the genuine species age.

Here, we use simulations to quantify the predictability of species age from phylogenetic trees of extant taxa under different diversification scenarios. Specifically, we performed simulations where we know the true age of species to: (1) quantify the error in phylogenetic ages under various scenarios combining different speciation modes with a range of speciation and extinction rates and incomplete sampling; (2) examine whether this error affects our ability to make qualitative age comparisons between species; and (3) explore whether the signal of age-correlated extinction risk is preserved in the phylogenetic age of species. Finally, we propose a new approach to more accurately estimate species age under the assumption of bifurcating speciation while correcting for incomplete species sampling and assessing its ability to improve our interpretation of age-dependent extinction risks.

2 | METHODS

2.1 | Simulating species ages

We generated complete phylogenies of extant and extinct species under a stochastic birth–death process using the package TreeSim 2.4 (Stadler, 2010) for the R 4.3.0 statistical programming environment (R Core Team, 2019). We used the R package FossilSim 2.3.1 (Barido-Sottani et al., 2019) to map species onto the complete phylogenies under different speciation modes, thus assigning species labels across all tree branches. Subsequently, we used the labels assigned to terminal extant taxa to determine the true species ages. We then dropped all extinct species from the tree and obtained the

length of terminal branches to quantify the phylogenetic age of extant species. Finally, we rescaled all phylogenetic trees to a root age of one, ensuring that the absolute errors in species ages are comparable in plots and compared the relative true and phylogenetic ages among different simulation scenarios.

2.2 | Error in equating phylogenetic and species age

To explore whether there is a consistent over- or underestimation of species ages and to quantify the error in approximating species ages with phylogenetic ages, we simulated a range of datasets with different speciation modes and diversification rates. First, we simulated three sets of 100 phylogenetic trees with 100 extant species based on speciation rates equal to 0.1, 0.5 and 1, combined with 100 extinction rates ranging from 0 to 0.99 in equal increments. Second, on each of these phylogenies, we mapped species according to different scenarios of speciation: (1) budding speciation, (2) bifurcating speciation, (3) a combination of budding speciation and anagenetic speciation with the rate of anagenesis set to half of the speciation rate and (4) bifurcating speciation combined with anagenetic speciation with the rate of anagenesis set to half of the speciation rate.

Across all trees, we obtained a total of 120,000 extant species, 30,000 for each speciation scenario. For each speciation mode and extinction fraction (defined as the extinction rate divided by the speciation rate), we calculated the mean absolute percentage error (MAPE) across all species for each tree as a measure of the deviation between the phylogenetic ages from the true age:

$$\text{MAPE} = \frac{100}{n} \times \sum_{i=1}^{n} \frac{|v_i - s_i|}{s_i},$$

where s is the true species age, v is the phylogenetic age and n is the number of tips in the tree (pruned of the extinct species). Next, for

each speciation mode, we plotted the MAPE against the simulated extinction fraction.

2.3 | Impact of age error in comparing species ages

To explore whether the error introduced by approximating species age with phylogenetic age impacts our ability to make qualitative judgements, such as which of two extant species is the younger one, we simulated 1000 phylogenetic trees with values of extinction fractions of 0.9, 0.5 and 0.0, combined with a fixed speciation rate of 1 (totalling 3000 trees). Second, on each of these phylogenies, we mapped species according to budding and bifurcating speciation. Thus, we simulated 300,000 extant species for each speciation mode. Next, we calculated the proportion of cases where the younger of two species, according to its phylogenetic age, is, in fact, the older one given the true age of the two species. We performed this comparison from the perspective of an empirical researcher who can only obtain the phylogenetic ages. We made two types of comparisons for each phylogeny: (1) between the youngest and oldest species in the phylogeny and (2) between two randomly selected species.

2.4 | Error in the phylogenetic age due to uniform incomplete sampling

To evaluate the error of equating phylogenetic and true age introduced by uniform incomplete sampling under bifurcating speciation, we simulated 1000 phylogenetic trees with a speciation rate of 0.3 and extinction rates of 0.05, 0.15 and 0.25. In addition to fully sampled phylogenies (where all extant species are included), we also simulated trees with incomplete taxon sampling, where 25% or 50% of the tips were randomly removed. We conditioned the simulation on the number of tips, ensuring that each tree included 100 sampled tips even after dropping the unsampled ones (134 or 200 initial tips). We calculated the MAPE for each tree and compared the incomplete sampling percentages for each extinction scenario.

2.5 | A probabilistic approach to infer species age

We developed a new approach to infer the age of a species based on its phylogenetic age in a tree of extant taxa and on the properties of a reconstructed birth–death process. Under the assumption of bifurcating speciation, the phylogenetic age represents the upper boundary of plausible species ages and corresponds to the true age in the absence of extinction and complete taxon sampling. However, the true age could be younger if extinction led to the disappearance of recent cladogenetic events from the phylogeny and/or if incomplete sampling of extant species led to unobserved branching events. We assume a species phylogenetic age v to be the result of a reconstructed birth–death process (Nee et al., 1994) with time-homogenous speciation rate λ and extinction rate μ and with uniform incomplete sampling of the

extant species determined by the sampling fraction ρ (i.e., $\rho = 1$ if all living species are included in the tree).

Based on the properties of the birth–death process, we can derive the probability that a lineage leaves at least one sampled descendant, without which the node with age v would not be observed, as defined by Yang and Rannala (1997):

$$P(0, v) \frac{\rho(\lambda - \mu)}{\rho\lambda + (\lambda(1 - \rho) - \mu)e^{(\mu - \lambda)v}} \tag{1}$$

Thus, the probability that a birth–death process leaves no descendants after time t is

$$\Phi(t) = 1 - P(0, t) \tag{2}$$

The true age of a species is equal to its phylogenetic age only if no other unobserved events of speciation occurred between the time v and the present ($t = 0$), assuming bifurcating speciation. The rate at which an unobserved speciation event can occur at time t along the branch leading to a single sampled species at the present is equal to the product of the speciation rate λ and the probability that one of the two new species left no sampled descendants at time 0:

$$\phi_t = \lambda \times 2\Phi(t) \tag{3}$$

where we multiply by 2 to account for the possibility that either of the two species might leave no descendants. Given a Poisson process with rate $\phi_{(t)}$, we can calculate the probability that the true species age s is equal to its phylogenetic age v as the probability that no unobserved speciation events occur during this time:

$$P(s = v) = \exp\left(- \int_0^v \phi(t)dt \right) \tag{4}$$

This is computed based on the rate of unobserved speciation events integrated from time $t_1 = 0$ to time $t_2 = v$, that can be computed as

$$\int_{t_1}^{t_2} \phi_s ds = 2\left[(\mu s - \log(\lambda\rho - (\lambda(\rho - 1) + \mu)\exp((\mu - \lambda)s)))\right]\Big|_{t_1}^{t_2} (t_2 > t_1) \tag{5}$$

Similarly, we can compute the probability that the true age of a species is instead younger than its phylogenetic age due to a hidden speciation event occurring at time t. This is equal to the rate of a hidden event at time t multiplied by the probability that no events occurred between t and the present:

$$P(s = t) = \phi(t)\exp\left(- \int_0^t \phi(t)dt \right), \text{with } v < t < 0 \tag{6}$$

Thus, the probability distribution describing the true species age is a function of the phylogenetic age, the speciation and extinction rates and the taxon sampling. It is a mixture between the probability mass assigned to $P(s = v)$ and the density for $P(0 < s < v)$. The function can be reproduced accurately through stochastic birth–death simulations (as we show in Figure S5).

We then estimate the expected age from the probability distribution as its mean and median values, which can be obtained through numerical integration and optimization:

$$\text{mean}(P(s)) = \nu P(s = \nu) + \int_0^\nu tP(s = t)dt$$

and

$$\text{median}(P(s)) = \nu - t \text{ with } t \in (0, \nu) \text{ such that } P(s = \nu) + \int_\nu^t P(s = u)du = 0.5,$$

noting that the median is equal to the phylogenetic age when $P(s = \nu) \geq 0.5$.

To evaluate the accuracy of our probabilistic species age estimator, we applied it to the simulations described in the "Error in the phylogenetic age due to uniform incomplete sampling" section, using the true speciation rate, extinction rate and sampling fraction. We then calculated the MAPE for each tree as a measure of the deviation between the function's point estimates and the true age. Then, we calculated the ΔMAPE as the difference between the MAPE of the estimated mean and median species ages and the MAPE of the corresponding phylogenetic ages.

2.6 | Assessing the probabilistic species age estimator

To assess the performance and limitations of our probabilistic species age estimator on an empirical dataset, we calculated the phylogenetic age of all living mammals and estimated the mean and median species ages based on a complete phylogenetic tree from Upham et al. (2019). While speciation and extinction rates (required along with taxon sampling for the probabilistic estimation of species age) can be in inferred from the phylogeny in principle (Gernhard, 2008; Nee, 2001; Stadler, 2013), their accurate estimation can be difficult (Louca & Pennell, 2020; Rabosky, 2010; Stadler & Bokma, 2013). Therefore, we evaluated three scenarios of low (0.15), intermediate (0.50) and high (0.95) extinction fractions. We used the R package diversitree 0.10.0 (FitzJohn, 2012) to estimate speciation and extinction rates from the phylogenetic tree constrained on the different extinction fractions.

For *Homo sapiens*, we used the more recent phylogenetic tree by Rivas-Gonzáles et al. (2023) to estimate species age. Then, for a selection of four species (*Balaena mysticetus* (bowhead whale), *Homo sapiens* (human), *Acinonyx jubatus* (cheetah) and *Ursus arctos* (brown bear)), we also compared phylogenetic age and estimated species age with the age of the oldest fossil record based on the datasets compiled by Silvestro et al. (2018). We plotted phylogenetic ages against the estimated mean species ages under different extinction fractions to evaluate the impact of our probabilistic approach on the distribution.

2.7 | Simulation of age-dependent extinction risks

To evaluate the impact of the erroneous estimation of species age due to incomplete sampling and extinction on macroevolutionary analyses, we explored whether a relationship between species age and contemporary extinction risk (e.g., Johnson et al., 2002; Tanentzap et al., 2020; Verde Arregoitia et al., 2013) is correctly preserved in the phylogenetic ages. We performed two experiments: (i) different extinction scenarios and fully sampled trees and (ii) different levels of nonrandom incomplete sampling (where older species were less prone to be sampled than younger ones) under intermediate extinction. For the first experiment, we simulated 1000 phylogenetic trees with a speciation rate of 0.3 and extinction rates of 0.05, 0.15 and 0.25. For the second experiment, we simulated 1000 phylogenetic trees with a speciation rate of 0.3, extinction rate of 0.15 and two levels of nonrandom incomplete taxon sampling, where 25% or 50% of the tips were removed. Then, in each experiment, we binned the same number of extant species according to their age and assigned them to five categories simulating an increase in extinction risks with age, encapsulated by the IUCN Red List categories (International Union for the Conservation of Nature, 2016): Least Concern (LC), Near Threatened (NT), Vulnerable (VU), Endangered (EN) and Critically Endangered (CR). This enabled us to generate a positive effect, with older species being at higher extinction risk regarding the IUCN categories, assuming bifurcating speciation. Finally, we quantified the share of the 1000 datasets where the order of the mean age per IUCN category did not match with the simulated monotonic increase when utilizing (a) phylogenetic ages and (b) the mean and median estimated species ages. We presented on the correspondent figures (Figures 7 and S7) only 10% of the results (the correct and incorrect estimations).

3 | RESULTS

3.1 | Error in equating phylogenetic and species age

Under the assumption of bifurcating speciation and in the absence of extinction events, phylogenetic ages matched the true age of extant species (Figure 2a). At low extinction fractions (<0.25), 94% of the phylogenetic age estimations were congruent with the true age. At higher extinction fractions (>0.75), this was also the case for most species (73%; Figure 2b). However, age overestimation increased with extinction fraction, occasionally suggesting that the species is as old as the root age. While under bifurcating speciation, the phylogenetic age never underestimated the true species age, both over- and underestimation occurred in the case of budding speciation (Figure 2a). Moreover, the proportion of cases where the phylogenetic ages equal the species age was lower than in the bifurcating scenario (Figure 2b). Overestimated ages were more frequent with high extinction, while underestimations occurred with low extinction, but in principle, both scenarios happened under the complete range of extinction rates (Figure 2b). Even at low extinction fractions, ~50% of phylogenetic ages did not match the true ages.

In datasets simulated under a mixture of anagenetic and bifurcating speciation, phylogenetic ages deviated more strongly from the true ages than under a pure bifurcating process, as anagenetic
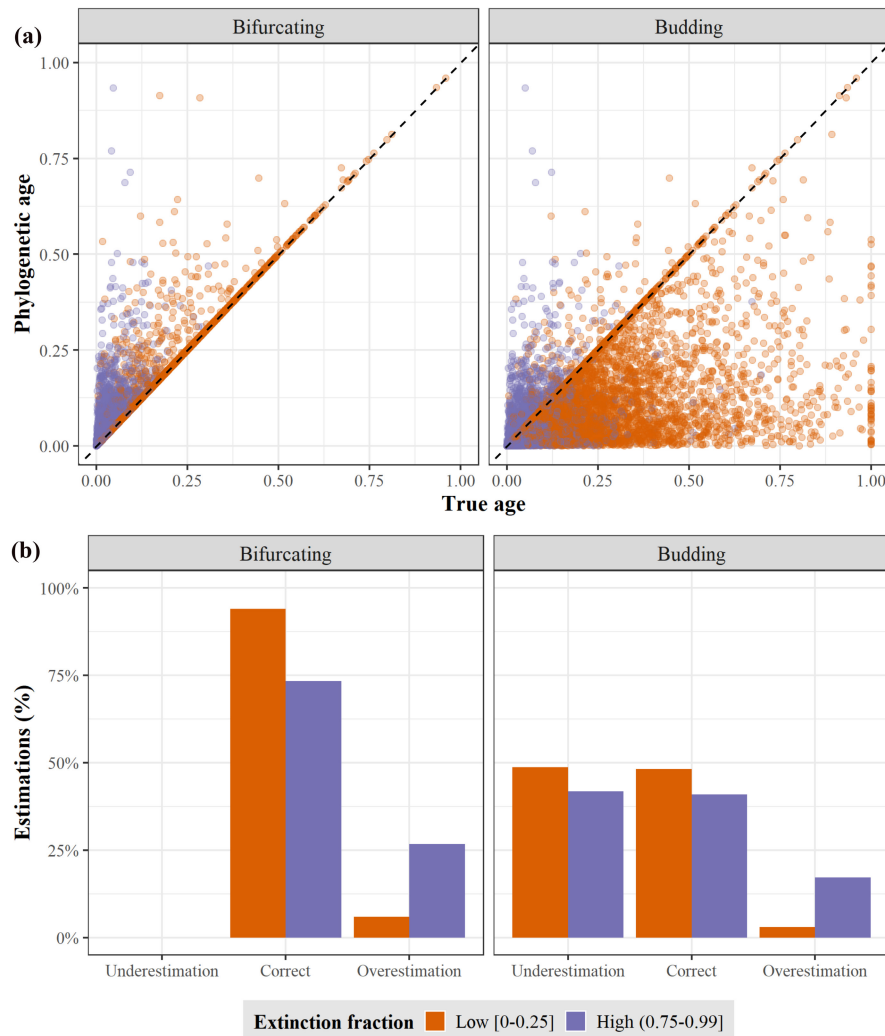
FIGURE 2 (a) True age versus phylogenetic age at low and high extinction fractions for bifurcating and budding speciation. Each point represents a species and both true and phylogenetic ages are scaled relative to the root age of the correspondent phylogenetic tree. (b) Percentage of underestimations, correct estimations and overestimations at low and high extinction fractions for bifurcating and budding speciation.

events are unobserved in the phylogeny (Figure S1a). With a low extinction fraction, phylogenetic ages were congruent with the true species ages in 75% of cases instead of 94% and with high extinction, the share decreased from 73% to 62% (Figure S1b). In datasets with mixed anagenetic and budding speciation, phylogenetic ages also deviated more than under a pure budding process; with a low extinction fraction, phylogenetic ages were congruent at 39% instead of 50% of cases and with high extinction, accuracy decreased to 34%.

While a budding speciation mode led to a higher baseline error than bifurcation, the latter showed a stronger increase with extinction (Figure 3). Under either mode of speciation, speciation rates did not substantially impact error in age. For strictly bifurcating speciation, the error was negligible in the absence of extinction, but the MAPE increased to up to 150% with extinction fractions exceeding 0.75. In contrast, under budding speciation, the MAPE was around 25% in the absence of extinction, increasing to 30%–120% with extinction fractions exceeding 0.75. In datasets incorporating anagenetic speciation, the MAPE reached as high as 500% in some simulations (Figure S2).

3.2 | Impact of age error on comparing species ages

For the combination of strictly bifurcating speciation and all extinction scenarios, selecting the phylogenetically youngest and oldest species never resulted in a case where the presumed older species turned out to be the younger of the two species according to their simulated age (Figure S3). Thus, in this speciation mode, the risk of a qualitative error when comparing species at the extremes of the age range is minimal. However, the age ranking of two randomly selected species was found to be incorrect in 4.3% and 6.9% of cases for intermediate and high extinction, respectively (Figure S4). Thus, qualitative errors in comparing species ages are non-negligible even under the assumption of bifurcating speciation.

For budding speciation, the age rank of the oldest and youngest species was erroneously determined in 2.2% of the simulations in the absence of extinction, increasing to 7.5% and 12.2% for intermediate and high extinction, respectively (Figure S3). Thus, under the assumption of budding speciation, there is a substantial risk of mistaking the oldest and youngest species in the clade. The error in age

FIGURE 3 Error in equating phylogenetic age with true age. The error was quantified as mean absolute percentage error (MAPE) between the true and phylogenetic ages across all species for each tree simulated under bifurcating and budding speciation. Each dot represents one of the 300 trees for each speciation mode using different rates of speciation and extinction fraction. The dashed lines represent the MAPE trend along the extinction fraction for each speciation rate.
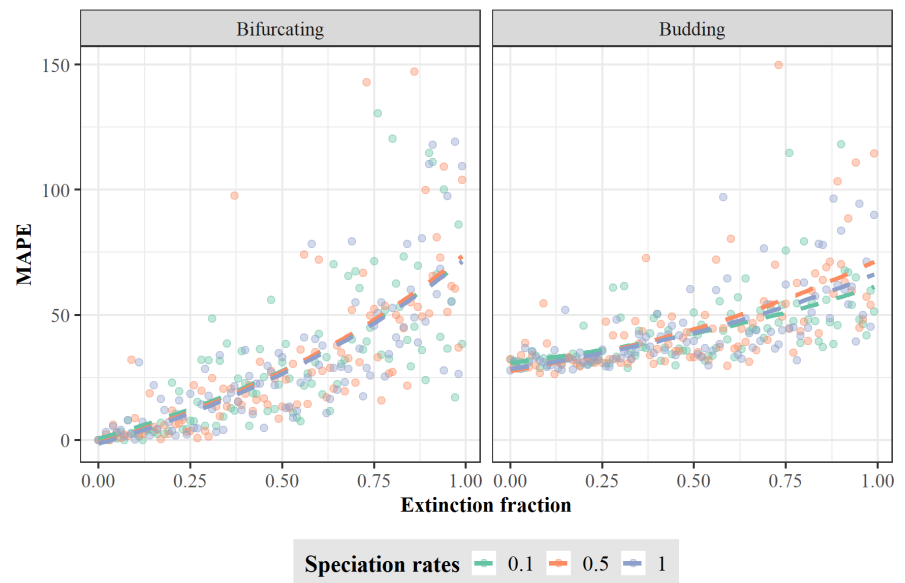

FIGURE 4 Effect of incomplete taxon sampling and extinction on error in species ages. Error in equating the phylogenetic age with true species age for the fully sampled phylogeny, with 25% and 50% of missing extant species, in the three extinction scenarios (no extinction, intermediate and high; from left to right). The error was quantified as mean absolute percentage error (MAPE) between the true phylogenetic ages across 100 species for each of 1000 trees for each missing species scenario simulated under bifurcating speciation.
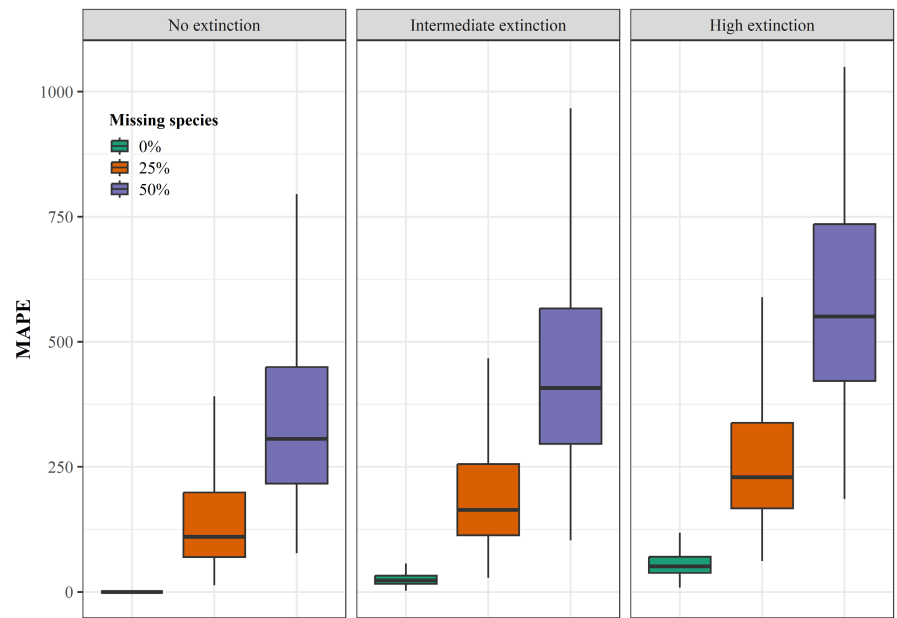
ranking of two randomly selected species was even higher, around 25%, irrespective of the extinction level (Figure S4).

3.3 | Error on equating phylogenetic and species age given uniform incomplete sampling

In the absence of extinction and assuming bifurcating speciation, the MAPE for fully sampled trees was ~0%. However, it increased to ~100% and ~300% for 25% missing species and 50% missing species, respectively (Figure 4). With intermediate extinction, the MAPE for fully sampled trees was $25 \pm 20\%$, but increased 15-fold for trees missing 25% of the extant species; and for trees missing 50% of the extant species, the error increased 85-fold. Under high extinction,

the MAPE for fully sampled trees was $60 \pm 38\%$; while for trees missing 25% and 50% of extant species, the error increased 8-fold and 23-fold, respectively.

3.4 | Assessing the probabilistic estimation of species age

Our probabilistic approach of estimating species age accurately predicted the probability distribution of species ages conditional on branching time, speciation and extinction rates and taxon sampling (Figure S5). The estimation of species age based on our probabilistic approach resulted in improved accuracy compared with phylogenetic age across most simulation scenarios, except

for simulation cases with low extinction and fully sampled trees, where both estimates converged to the true age (MAPE = 0%). With increasing extinction and fractions of missing species, our probabilistic estimates of species age resulted in a substantially lower error compared with phylogenetic ages (Figures 5 and S6). For instance, with high extinction and a taxon sampling of 50% species, the estimated species age reduced the error by around 400% compared to the phylogenetic age (ΔMAPE = −408% for the mean estimated ages and −411% for the median across estimates; Figures 5 and S6). Mean and median estimates of species age led to similar levels of accuracy.

3.5 | Species age in the mammalian tree

The mean phylogenetic age of 5987 mammalian species was 4.38 myr, with a 95% interval ranging from 0.07 to 14.88 myr. The mean estimated species ages across all species were between 1.38 and 2.40 myr, assuming low or high extinction fractions with 95% intervals ranging from 0.07 to 14.88 myr for low extinction and from 0.07 to 6.41 myr for high extinction (Figure 6a). As predicted by our simulations, the difference between phylogenetic and estimated species ages was largest for species with a comparatively old phylogenetic age and under the assumption of a high extinction fraction. A comparison for four species among phylogenetic and estimated ages and the age of their oldest fossil record attributed to the same species showed different patterns. For the bowhead whale (*Balena mysticetus*) and *Homo sapiens*, the ages of the oldest fossils were substantially younger than the phylogenetic ages. In this case, the ages estimated under a high extinction fraction reduced the discrepancy between fossil and phylogenetic evidence. In contrast, the fossil age of the cheetah (*Acinonyx jubatus*) and brown bear (*Ursus arctos*) was older than their respective phylogenetic ages.

3.6 | Detecting simulated age-dependent extinction risk

The use of phylogenetic age as an approximation of species age led to error rates of 1.2%, 5.9% and 19.2% in detecting the correct monotonic correlation between species ages and extinction risk, with fully sampled trees, and for scenarios with low, intermediate and high extinction rates, respectively (Figure S7). Thus, even under intermediate extinction, the true relationship between age and extinction risk was incorrectly estimated in a significant fraction of the simulations and higher extinction rates led to a further substantial drop in the reliability of this approach. In contrast, estimating species ages based on our probabilistic method led to 6 to 8-fold lower error rates, which dropped to 1% and 2.3% with intermediate and high extinction, respectively. Under incomplete taxon sampling, in which the sampling probability was negatively correlated with species age and under intermediate extinction, the use of phylogenetic age as an approximation of species age led to error rates of 13%

and 56% for scenarios with 25% and 50% of missing extant species, respectively (Figure 7). For 25% missing species, the function's point estimates reduced the error rates 8 to 26-fold, decreasing to 1.5% for mean age and 0.5% for median age. For 50% missing species, the point estimates reduced the error rates 5 to 21-fold, decreasing to 11.8% for mean age and 2.7% for median age.

4 | DISCUSSION

The use of branch lengths of phylogenetic trees as an approximation of species ages is becoming central to an increasing number of studies that use them to evaluate the relationship with macroecological or evolutionary patterns, such as current extinction risks (Guo et al., 2024; Pie & Caron, 2023; Sonne et al., 2022; Tanentzap et al., 2020; Verde Arregoitia et al., 2013). Here, we showed that this approximation leads to substantial errors and that its accuracy is hampered by three shortfalls: incomplete sampling of extant species, unobserved extinction events and unknown speciation modes. The only instance in which phylogenetic ages correctly predict species age is under the assumption of a bifurcating speciation process in the absence of extinction and with all living species included in the phylogenetic tree. While the prevalence of true speciation modes remains difficult to access (Bapst & Hopkins, 2017; Silvestro et al., 2018; Wagner et al., 1995), the fossil record unequivocally illustrates extinction across all clades in the tree of life (Bambach, 2006; Benton, 2023; Pimm et al., 2014) and there is substantial evidence that many living species remain unknown to science and are therefore absent from empirical phylogenetic trees (Blackwell, 2011; Mora et al., 2011). Thus, the scenario under which phylogenetic age correctly predicts species age is highly improbable.

Some authors acknowledged the challenges associated with measuring species age from phylogenies (Swenson, 2019) and have proposed approaches to account for them. For example, Sonne et al. (2022) determined young and old Andean hummingbirds and assessed the sensitivity of their results to incomplete taxon sampling by generating 1000 trees with randomly missing species. Pie and Caron (2023) accounted for taxonomic incompleteness by pruning an additional 1%–5% of species and evaluated if their conclusions changed and found that they did not. Yet, the magnitude of the error associated with the direct use of the length of phylogenetic branches as estimators of species ages remains underappreciated, as shown by the prevalence of this approach in many studies.

We showed that the largest error in estimating species ages from phylogenetic trees is linked with incomplete sampling of extant species. This is a problem that in principle can be solved by extending the scope of sampling to include all species in the phylogenetic inference. Yet, despite the advances in the scalability of DNA sequencing, this remains impractical for large clades, including some of the best-sampled ones such as vertebrate groups, in which many species still lack genetic data (Jetz & Pyron, 2018; Tonini et al., 2016; Upham et al., 2019). In addition, a substantial proportion of species might be missing from phylogenetic trees because they are yet to be

FIGURE 5 Performance of the probabilistic age estimator. Error in equating the phylogenetic age and the point estimates (mean and median) with the true species age for three sampling scenarios (full, 25% and 50% of missing species; from up to down) and three extinction scenarios (no extinction, intermediate and high; from left to right). The error was quantified as mean absolute percentage error (MAPE) between the true and point estimates or phylogenetic ages across 100 species for all 1000 trees for each sampling and extinction scenario simulated under bifurcating speciation.
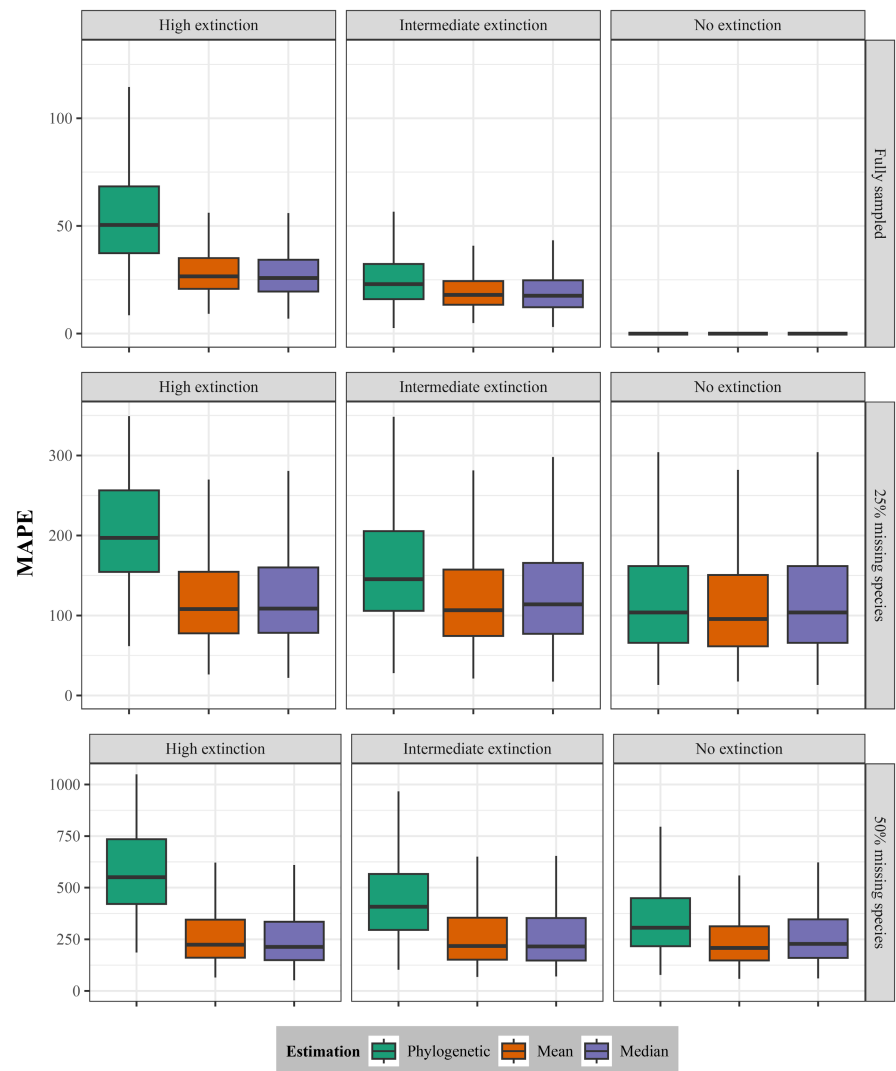
discovered, a problem often termed as the Linnean shortfall (Diniz Filho et al., 2023; Hortal et al., 2015). The magnitude of the Linnean shortfall is unknown, but available estimates show that it affects some clades significantly more than others (Moura & Jetz, 2021; Ondo et al., 2023), with the diversity of vastly diverse groups, such as bacteria, insects and fungi, likely to be highly underestimated (Blackwell, 2011; Mora et al., 2011; Wiens, 2023).

Our simulations have revealed that the accurate estimation of species ages from phylogenetic trees is essentially impossible under some scenarios using current approaches. Under the assumption of budding speciation, the error is high even without extinction and with complete sampling. This is because phylogenetic trees are agnostic about parent and descendant species following a branching event (Figures 1 and 3). Phylogenetic ages are, by construction, identical for sister species, which is necessarily wrong within a budding speciation scenario. Similarly, anagenetic speciation also leads to high error, which did not vary substantially with extinction. However, anagenetic speciation might be impossible to quantify, except perhaps in cases of high-resolution fossil time series (Aze et al., 2011), resulting in a general debate on the use of the term anagenesis in evolutionary biology (Vaux et al., 2016) and biogeography

(Emerson & Patiño, 2018; Meiri et al., 2018). Thus, species age is unidentifiable under the assumption of speciation modes that deviate from a strictly bifurcating scenario. For some groups, such as vertebrates, budding speciation appears more plausible (Skeels & Cardillo, 2019).

The lowest error in species age estimation was observed under scenarios of bifurcating speciation. This is the implicit assumption of most studies using approximations of species ages (Alzate et al., 2023; Freer et al., 2022; Kennedy et al., 2022), even though it is at odds with the assumption of all birth–death models, which are based on budding speciation, used in the molecular clock analyses that estimate the phylogenetic trees in the first place (Gernhard, 2008; Nee et al., 1994; Stadler, 2013). Despite the lower error, our simulations showed that both extinction and missing lineages can yield a substantial decrease in accuracy (Figures 3–5, S3, and S4), potentially leading to qualitative misinterpretations of general patterns such as age-dependent extinction risks (Figures 7 and S7).

Our probabilistic approach to estimate species ages, based on the properties of the reconstructed birth–death process, effectively mitigated biases associated with incomplete sampling and extinction. It substantially improved the accuracy of species age
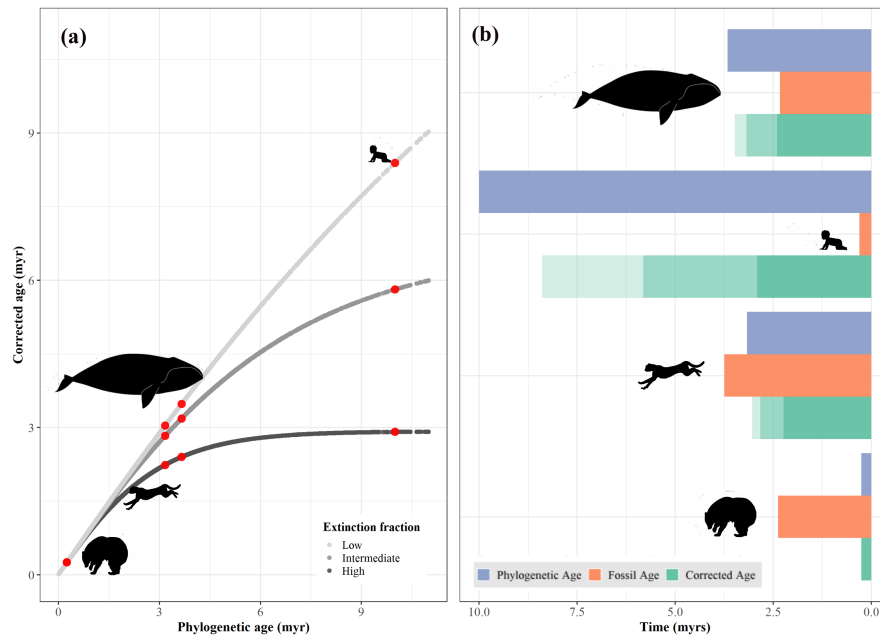
FIGURE 6 Phylogenetic ages and estimated species ages inferred for 5987 species of mammals. (a) Phylogenetic age versus estimated age obtained from three extinction fractions (low (0.15), intermediate (0.50) and high fraction (0.95)). The *x*-axis is truncated at 11 myr for clarity, but phylogenetic ages extended to a maximum of 299.21 myr. Red circles correspond to the phylogenetic-estimated ages' relationship across the extinction fractions for *Ursus arctos*, *Acinonyx jubatus*, *Balaena mysticetus* and *Homo sapiens* (from bottom to top). (b) Comparison of phylogenetic age, mean estimated age (based on the three extinction fractions) and age of the oldest fossil record for four mammals with fossil record (*Balaena mysticetus*, *Homo sapiens*, *Acinonyx jubatus* and *Ursus arctos*; from top to bottom). The transparency of the mean corrected age corresponds to the different levels of extinction fraction (low, intermediate and high fraction; from transparent to solid green).

estimation, leading to lower error rates, especially under scenarios of high extinction and a high percentage of missing extant species (Figure 5). Similarly, the use of estimated ages enabled the correct identification of simulated extinction–risk relationships under different incomplete sampling and extinction scenarios (Figures 7 and S7), reducing the error by more than one order of magnitude compared to the use of the phylogenetic ages.

The application of our approach to a large empirical dataset of more than 5000 mammalian species revealed that the estimated species ages can strongly diverge from phylogenetic ages. This pattern was most pronounced for species sitting on long branches, such as humans, who likely separated from their most recent common ancestor with the chimpanzee around 10 Ma. The probability that no other speciation event occurred along that lineage is low based on our probabilistic function, which indeed estimates a younger species age. This brings the estimated species age closer to our knowledge based on the fossil record of *Homo sapiens* and its extinct relatives. In other instances, the apparent species age observed in the fossil record is older than its phylogenetic age and therefore its estimated age. This pattern could be driven by errors in the data, for example, in the estimated branching times, the tree topology or species identification in the fossil record. However, the same pattern can result from a violation of the assumption of bifurcating speciation, as seen under budding speciation where a species can be older than its phylogenetic age (Figure 1).

The main difference between phylogenetic and estimated species ages in our empirical analysis lies in the resulting range of ages, which spans to much larger values when taking branching times at face value. Previous analyses of subsets of the mammalian fossil record that explicitly corrected for sampling biases (Hagen et al., 2018) estimated the mean species longevity at 2.02 myr (95% credible interval 1.76–2.27 myr) and similar estimates were obtained from a large compendium of fossil North American mammalian occurrences (Prothero, 2014). Interestingly, while these estimates are at odds with the mean phylogenetic age (>4 myr), they match closely with our estimated mean age of 1.38 and 2.40 myr assuming low or high extinction, respectively. Thus, our probabilistic estimation of species age brings the estimated distribution of species ages closer to our understanding of species longevity based on paleontological evidence.

The performance of our estimator is, however, contingent on the ability of birth–death models to correctly estimate speciation and extinction rates from phylogenies of extant species. While these methods have demonstrated high accuracy across several simulation settings (Silvestro et al., 2011; Stadler, 2011), their performance has been questioned with empirical datasets and under complex models of rate variation (Louca & Pennell, 2020; Rabosky, 2010). A commonly observed pattern is the estimation of a 0-extinction rate from empirical phylogenies (Louca & Pennell, 2021). This would make estimated ages identical to the
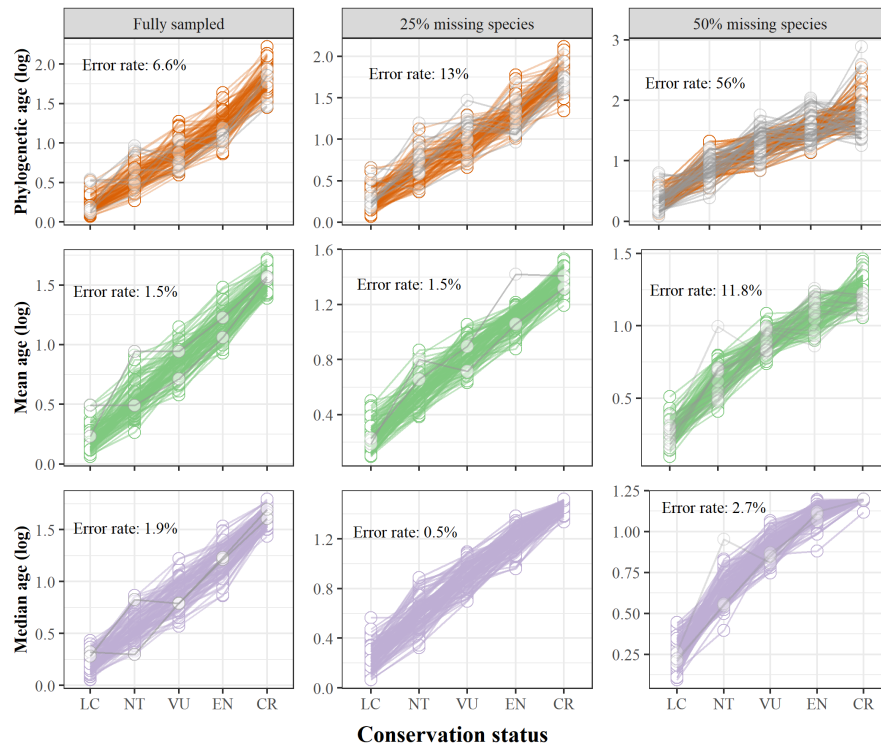
FIGURE 7 Power to recover an age extinction-risk relationship under different incomplete sampling scenarios. Simulated species ages under an intermediate extinction scenario and assuming bifurcating speciation were binned into conservation status categories, which represents an increase in extinction risk by true age (CR, critically endangered; EN, endangered; LC, least concern; NT, near threatened; VU, vulnerable). We used the phylogenetic age (orange), the mean age (green) and the median age (purple) obtained from our function's point estimates to calculate the mean age per conservation status category and assess if every mean age increases in comparison with the previous category with lower extinction risk. We evaluated this relationship in three sampling scenarios (fully sampled, 25% and 50% of missing species; from left to right). The error rate is the percentage across all 1000 phylogenies where the relationship between the mean ages and the conservation status categories is not always increasing (shown by grey lines).

phylogenetic ages in completely sampled phylogenies. Yet, the fossil record decisively shows variation in extinction and speciation rates within the same order of magnitude across virtually all clades (Parry, 2021). The robustness of estimated extinction rates can increase by incorporating fossil data into the analyses (Heath et al., 2014; Silvestro et al., 2018; Warnock et al., 2020). As a way to evaluate the robustness of conclusions drawn from patterns of estimated species ages, our probabilistic approach can be applied across a range of plausible values of speciation rates, extinction rates (as demonstrated in our analyses of mammal species) and potentially across fractions of incomplete sampling if the true species richness of a clade is uncertain.

5 | CONCLUSION

In this study, we quantified the deviation between true species age and phylogenetic age due to incomplete taxon sampling, extinction and unknown speciation modes. We found that phylogenetic age serves as a reliable proxy of species age only in a rather unlikely case in which (1) all species within a clade are known to science and included in the phylogenetic tree, (2) speciation occurs as a strictly bifurcating process and (3) there is either no or low extinction. Using simulations, we identified that incomplete taxon sampling and budding and anagenetic speciation cause the highest mismatch between phylogenetic age and true species age, and can only be accounted for with additional information, for example, fossil data. We proposed a probabilistic approach based on the properties of the birth–death process and under the assumption of bifurcating speciation that can drastically improve the accuracy of estimated species ages by mitigating biases associated with extinction and incomplete taxon sampling. We note that, even in this case, the robustness of such estimates will be contingent on the accuracy of estimated speciation and extinction rates, the quantification of missing species and the underlying phylogenetic tree and dating of the branching times. In light of our findings, we caution against the direct use of branch lengths as proxies for species ages and advocate for applying our probabilistic approach to correct species age estimates. We suggest exploring different values of diversification rates and incomplete sampling fractions. We hope our results will spur a discussion about the use of phylogenetic trees in inferring species age and prompt a critical evaluation of the robustness of inferences linking species age with traits, ecological variables and extinction risks.

AUTHOR CONTRIBUTIONS

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST STATEMENT

We declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

All the codes utilized in this study are available as Supplementary materials and in this repository: https://github.com/caldecid/Evolutionary-age-discrepancies. Our corrective species age approach, for its implementation in R, can be downloaded directly from this GitHub repository: https://github.com/thauffe/SpeciesAge.

ORCID

Carlos Calderón del Cid https://orcid.org/0000-0002-3897-4615
Torsten Hauffe https://orcid.org/0000-0001-5711-9457
Juan D. Carrillo https://orcid.org/0000-0003-2475-3341
Michael R. May https://orcid.org/0000-0002-5031-4820
Rachel C. M. Warnock https://orcid.org/0000-0002-9151-4642
Daniele Silvestro https://orcid.org/0000-0003-0100-0961

REFERENCES

Alroy, J., Marshall, C. R., Bambach, R. K., Bezusko, K., Foote, M., Fürsich, F. T., Hansen, T. A., Holland, S. M., Ivany, L. C., Jablonski, D., & Jacobs, D. K. (2001). Effects of sampling standardization on estimates of phanerozoic marine diversification. *Proceedings of the National Academy of Sciences of the United States of America*, 98(11), 6261–6266. https://doi.org/10.1073/pnas.111144698

Alzate, A., Rozzi, R., Velasco, J. A., Robertson, D. R., Zizka, A., Tobias, J. A., Hill, A., Bacon, C. D., Janzen, T., Pellissier, L., van der Plas, F., Rosindell, J., & Onstein, R. E. (2023). The evolutionary age-range size relationship is modulated by insularity and dispersal in plants and animals. bioRxiv, 2023-11 https://doi.org/10.1101/2023.11.11.566377

Anacker, B. L., & Strauss, S. Y. (2014). The geography and ecology of plant speciation: Range overlap and niche divergence in sister species. *Proceedings of the Royal Society B: Biological Sciences*, 281(1778), 20132980. https://doi.org/10.1098/rspb.2013.2980

Aze, T., Ezard, T. H. G., Purvis, A., Coxall, H. K., Stewart, D. R. M., Wade, B. S., & Pearson, P. N. (2011). A phylogeny of Cenozoic macroperforate planktonic foraminifera from fossil data. *Biological Reviews*, 86(4), 900–927. https://doi.org/10.1111/j.1469-185X.2011.00178.x

Balmford, A. (1996). Extinction filters and current resilience: The significance of past selection pressures for conservation biology. *Trends in Ecology and Evolution*, 11, 193–196. https://doi.org/10.1016/0169-5347(96)10026-4

Bambach, R. K. (2006). Phanerozoic biodiversity mass extinctions. *Annual Review of Earth and Planetary Sciences*, 34, 127–155. https://doi.org/10.1146/annurev.earth.33.092203.122654

Bapst, D. W., & Hopkins, M. J. (2017). Comparing cal3 and other a posteriori time-scaling approaches in a case study with the pterocephaliid trilobites. *Paleobiology*, 43(1), 49–67. https://doi.org/10.1017/pab.2016.34

Barido-Sottani, J., Pett, W., O'Reilly, J. E., & Warnock, R. C. M. (2019). FossilSim: An R package for simulating fossil occurrence data under mechanistic models of preservation and recovery. *Methods in Ecology and Evolution*, 10(6), 835–840. https://doi.org/10.1111/2041-210X.13170

Baum, D. A., Smith, S. D. W., & Donovan, S. S. S. (2005). The tree-thinking challenge. *Science*, 310(5750), 979–980. https://doi.org/10.1126/science.1117727

Benton, M. J. (2013). Origins of biodiversity. *Palaeontology*, 56, 1–7. https://doi.org/10.1111/pala.12012

Benton, M. J. (2016). Origins of biodiversity. *PLoS Biology*, 14(11), 1–7. https://doi.org/10.1371/journal.pbio.2000724

Benton, M. J. (2023). *Extinctions: How life survives, adapts and evolves*. Thames & Hudson.

Blackwell, M. (2011). The fungi: 1, 2, 3... 5.1 million species? *American Journal of Botany*, 98(3), 426–438. https://doi.org/10.3732/ajb.1000298

Caetano, D. S., & Quental, T. B. (2023). How important is budding speciation for comparative studies? *Systematic Biology*, 72(6), 1443–1453. https://doi.org/10.1093/sysbio/syad050

Calderón del Cid, C., Villalobos, F., Dobrovolski, R., Carrillo, J. D., Silvestro, D., & Vilela, B. (2024). The clade replacement theory: A framework to study age-dependent extinction. *Journal of Evolutionary Biology*, 37(3), 290–301. https://doi.org/10.1093/jeb/voae012

Callaway, E. (2017). Oldest *Homo sapiens* fossil claim rewrites our species' history. *Nature*, 546, 289–293. https://doi.org/10.1038/nature.2017.22114

Carrillo, J. D., Faurby, S., Silvestro, D., Zizka, A., Jaramillo, C., Bacon, C. D., & Antonelli, A. (2020). Disproportionate extinction of south American mammals drove the asymmetry of the Great American Biotic Interchange. *Proceedings of the National Academy of Sciences of the United States of America*, 117(42), 26281–26287. https://doi.org/10.1073/pnas.2009397117

Carrillo, J. D., Forasiepi, A., Jaramillo, C., Sánchez-Villagra, M. R., & Richardson, J. E. (2015). Neotropical mammal diversity and the great American Biotic Interchange: Spatial and temporal variation in South America's fossil record. *Frontiers in Genetics*, 5, 451. https://doi.org/10.3389/fgene.2014.00451

Davies, T. J., Smith, G. F., Bellstedt, D. U., Boatwright, J. S., Bytebier, B., Cowling, R. M., Forest, F., Harmon, L. J., Muasya, A. M., Schrire, B. D., Steenkamp, Y., van der Bank, M., & Savolainen, V. (2011). Extinction risk and diversification are linked in a plant biodiversity hotspot. *PLoS Biology*, *9*(5), e1000620. https://doi.org/10.1371/journal.pbio.1000620

Diniz Filho, J. A. F., Jardim, L., Guedes, J. J., Meyer, L., Stropp, J., Frateles, L. E. F., Pinto, R. B., Lohmann, L. G., Tessarolo, G., de Carvalho, C. J. B., Ladle, R. J., & Hortal, J. (2023). Macroecological links between the Linnean, Wallacean, and Darwinian shortfalls. *Frontiers of Biogeography*, *15*(2). https://doi.org/10.21425/F5FBG59566

Eldredge, N., Thompson, J. N., Brakefield, P. M., Gavrilets, S., Jablonski, D., Jackson, J. B. C., Lenski, R. E., Lieberman, B. S., McPeek, M. A., & Miller, W., III. (2005). The dynamics of evolutionary stasis. *Paleobiology*, *31*(S2), 133–145. https://doi.org/10.1666/0094-8373(2005)031

Emerson, B. C., & Patiño, J. (2018). Anagenesis, cladogenesis, and speciation on islands. *Trends in Ecology and Evolution*, *33*(7), 488–491. https://doi.org/10.1016/j.tree.2018.04.006

FitzJohn, R. G. (2012). Diversitree: Comparative phylogenetic analyses of diversification in R. *Methods in Ecology and Evolution*, *3*(6), 1084–1092. https://doi.org/10.1111/j.2041-210X.2012.00234.x

Foote, M. (1996). On the probability of ancestors in the fossil record. *Paleobiology*, *22*(2), 141–151. https://doi.org/10.1017/S0094837300016146

Foote, M., & Raup, D. M. (1996). Fossil preservation and the stratigraphic ranges of taxa. *Paleobiology*, *22*(2), 121–140. https://doi.org/10.1017/S0094837300016134

Freer, J. J., Collins, R. A., Tarling, G. A., Collins, M. A., Partridge, J. C., & Genner, M. J. (2022). Global phylogeography of hyperdiverse lanternfishes indicates sympatric speciation in the deep sea. *Global Ecology and Biogeography*, *31*(11), 2353–2367. https://doi.org/10.1111/geb.13586

Gaston, K. J., & Blackburn, T. M. (1997). Evolutionary age and risk of extinction in the global avifauna. *Evolutionary Ecology*, *11*, 557–565. https://doi.org/10.1007/s10682-997-1511-4

Gernhard, T. (2008). The conditioned reconstructed process. *Journal of Theoretical Biology*, *253*(4), 769–778. https://doi.org/10.1016/j.jtbi.2008.04.005

Guo, Q., Qian, H., Zhang, J., & Liu, P. (2024). The relationships between species age and range size. *Journal of Biogeography*, *51*, 1095–1103. https://doi.org/10.1111/jbi.14809

Hagen, O., Andermann, T., Quental, T. B., Antonelli, A., & Silvestro, D. (2018). Estimating age-dependent extinction: Contrasting evidence from fossils and phylogenies. *Systematic Biology*, *67*(3), 458–474. https://doi.org/10.1093/sysbio/syx082

Harvey, P. H., May, R. M., & Nee, S. (1994). Phylogenies without fossils. *Evolution*, *48*(3), 523–529. https://doi.org/10.1111/j.1558-5646.1994.tb01341.x

Hawkins, B. A., Rodriguez, M. A., & Weller, S. G. (2011). Global angiosperm family richness revisited: Linking ecology and evolution to climate. *Journal of Biogeography*, *38*(7), 1253–1266. https://doi.org/10.1111/j.1365-2699.2011.02490.x

Heath, T. A., Hedtke, S. M., & Hillis, D. M. (2008). Taxon sampling and the accuracy of phylogenetic analyses. *Journal of Systematics and Evolution*, *46*(3), 239–257. https://doi.org/10.3724/SP.J.1002.2008.08016

Heath, T. A., Huelsenbeck, J. P., & Stadler, T. (2014). The fossilized birth–death process for coherent calibration of divergence-time estimates. *Proceedings of the National Academy of Sciences*, *111*(29), E2957–E2966. https://doi.org/10.1073/pnas.1319091111

Hennig, W. (1999). *Phylogenetic systematics*. University of Illinois Press.

Hortal, J., de Bello, F., Diniz-Filho, J. A. F., Lewinsohn, T. M., Lobo, J. M., & Ladle, R. J. (2015). Seven shortfalls that beset large-scale knowledge of biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, *46*, 523–549. https://doi.org/10.1146/annurev-ecolsys-112414-054400

IUCN. (2016). *IUCN Red list of threatened species. Version 2016-2*. Fourth Quarter.

Jetz, W., & Pyron, R. A. (2018). The interplay of past diversification and evolutionary isolation with present imperilment across the amphibian tree of life. *Nature Ecology & Evolution*, *2*(5), 850–858. https://doi.org/10.1038/s41559-018-0515-5

Johnson, C. N., Delean, S., & Balmford, A. (2002). Phylogeny and the selectivity of extinction in Australian marsupials. *Animal Conservation*, *5*, 135–142. https://doi.org/10.1017/S1367943002002196

Kennedy, J. D., Marki, P. Z., Reeve, A. H., Blom, M. P., Prawiradilaga, D. M., Haryoko, T., Koane, B., Kamminga, P., Irestedt, M., & Jønsson, K. A. (2022). Diversification and community assembly of the world's largest tropical Island. *Global Ecology and Biogeography*, *31*(6), 1078–1089. https://doi.org/10.1111/geb.13484

Kennedy, J. D., Wang, Z., Weir, J. T., Rahbek, C., Fjeldså, J., & Price, T. D. (2014). Into and out of the tropics: The generation of the latitudinal gradient among New World passerine birds. *Journal of Biogeography*, *41*(9), 1746–1757. https://doi.org/10.1111/jbi.12346

Losos, J. B., & Glor, R. E. (2003). Phylogenetic comparative methods and the geography of speciation. *Trends in Ecology and Evolution*, *18*(5), 220–227. https://doi.org/10.1016/S0169-5347(03)00037-5

Louca, S., & Pennell, M. W. (2020). Extant timetrees are consistent with a myriad of diversification histories. *Nature*, *580*(7804), 502–505. https://doi.org/10.1038/s41586-020-2176-1

Louca, S., & Pennell, M. W. (2021). Why extinction estimates from extant phylogenies are so often zero. *Current Biology*, *31*(14), 3168–3173. https://doi.org/10.1016/j.cub.2021.04.066

Meiri, S., Raia, P., & Santos, A. M. (2018). Anagenesis and cladogenesis are useful Island biogeography terms. *Trends in Ecology & Evolution*, *33*(12), 895–896. https://doi.org/10.1016/j.tree.2018.09.005

Mora, C., Tittensor, D. P., Adl, S., Simpson, A. G., & Worm, B. (2011). How many species are there on Earth and in the ocean? *PLoS Biology*, *9*(8), e1001127. https://doi.org/10.1371/journal.pbio.1001127

Moura, M. R., & Jetz, W. (2021). Shortfalls and opportunities in terrestrial vertebrate species discovery. *Nature Ecology & Evolution*, *5*(5), 631–639. https://doi.org/10.1038/s41559-021-01411-5

Mynard, P., Algar, A., Lancaster, L., Bocedi, G., Fahri, F., Gubry-Rangin, C., Lupiyaningdyah, P., Nangoy, M., Osborne, O. G., Papadopulos, A. S., & Sudiana, I. M. (2023). Impact of phylogenetic tree completeness and misspecification of sampling fractions on trait dependent diversification models. *Systematic Biology*, *72*(1), 106–119. https://doi.org/10.1093/sysbio/syad001

Nee, S. (2001). Inferring speciation rates from phylogenies. *Evolution*, *55*(4), 661–668. https://doi.org/10.1111/j.0014-3820.2001.tb00801.x

Nee, S., & May, R. M. (1997). Extinction and the loss of evolutionary history. *Science*, *278*(5338), 692–694. https://doi.org/10.1126/science.278.5338.692

Nee, S., May, R. M., & Harvey, P. H. (1994). The reconstructed evolutionary process. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *344*(1309), 305–311. https://doi.org/10.1098/rstb.1994.0068

Ondo, I., Dhanjal-Adams, K., Pironon, S., Silvestro, D., Deklerck, V., Grace, O., Monro, A. K., Nicolson, N., Walker, B., & Antonelli, A. (2023). Plant diversity darkspots for global collection priorities. bioRxiv, 2023-09 https://doi.org/10.1101/2023.09.12.557387

Parry, L. A. (2021). Evolution: No extinction? No way! *Current Biology*, *31*(14), R907–R909. https://doi.org/10.1016/j.cub.2021.05.044

Pearson, P. N. (1995). Investigating age-dependency of species extinction rates using dynamic survivorship analysis. *Historical Biology*, *10*(2), 119–136. https://doi.org/10.1080/10292389509380516

Pie, M. R., & Caron, F. (2023). Substantial variation in species ages among vertebrate clades. *Evolutionary Journal*, 2(1), kzad006. https://doi.org/10.1093/evolinnean/kzad006

Pimm, S. L., Jenkins, C. N., Abell, R., Brooks, T. M., Gittleman, J. L., Joppa, L. N., Raven, P. H., Roberts, C. M., & Sexton, J. O. (2014). The biodiversity of species and their rates of extinction, distribution, and protection. *Science*, 344(6187), 1246752. https://doi.org/10.1126/science.1246752

Prothero, D. R. (2014). Species longevity in North American fossil mammals. *Integrative Zoology*, 9(4), 383–393. https://doi.org/10.1111/1749-4877.12054

Qian, H., & Ricklefs, R. E. (2016). Out of the tropical lowlands: Latitude versus elevation. *Trends in Ecology & Evolution*, 31(10), 738–741. https://doi.org/10.1016/j.tree.2016.07.012

R Core Team. (2019). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.r-project.org/

Rabosky, D. L. (2010). Extinction rates should not be estimated from molecular phylogenies. *Evolution*, 64(6), 1816–1824. https://doi.org/10.1111/j.1558-5646.2009.00926.x

Rivas-Gonzáles, I., Rousselle, M., Li, F., Zhou, L., Dutheil, J., Munch, K., Shao, Y., Wu, D., Schierup, M. H., & Zhang, G. (2023). Pervasive incomplete lineage sorting illuminates speciation and selection in primates. *Science*, 380(6648), eabn4409. https://doi.org/10.1126/science.abn4409

Roopnarine, P. D., Byars, G., Fitzgerald, P., Paleobiology, S., & Winter, N. (1999). Anagenetic evolution, stratophenetic patterns, and random walk models. *Paleobiology*, 25(1), 41–57. https://doi.org/10.1666/0094-8373(1999)0252.3.CO;2

Rosenblum, E. B., Sarver, B. A. J., Brown, J. W., Des Roches, S., Hardwick, K. M., Hether, T. D., Eastman, J. M., Pennell, M. W., & Harmon, L. J. (2012). Goldilocks meets Santa Rosalia: An ephemeral speciation model explains patterns of diversification across time scales. *Evolutionary Biology*, 39(2), 255–261. https://doi.org/10.1007/s11692-012-9171-x

Silvestro, D., Castiglione, S., Mondanaro, A., Serio, C., Melchionna, M., Piras, P., di Febbraro, M., Carotenuto, F., Rook, L., & Raia, P. (2020). A 450 million years long latitudinal gradient in age-dependent extinction. *Ecology Letters*, 23(3), 439–446. https://doi.org/10.1111/ele.13441

Silvestro, D., Salamin, N., Antonelli, A., & Meyer, X. (2019). Improved estimation of macroevolutionary rates from fossil data using a Bayesian framework. *Paleobiology*, 45(4), 546–570. https://doi.org/10.1017/pab.2019.23

Silvestro, D., Schnitzler, J., & Zizka, G. (2011). A Bayesian framework to estimate diversification rates and their variation through time and space. *BMC Evolutionary Biology*, 11, 1–15. https://doi.org/10.1186/1471-2148-11-311

Silvestro, D., Warnock, R. C. M., Gavryushkina, A., & Stadler, T. (2018). Closing the gap between palaeontological and neontological speciation and extinction rate estimates. *Nature Communications*, 9(1), 5237. https://doi.org/10.1038/s41467-018-07622-y

Simpson, G. G. (1961). *Principles of animal taxonomy*. Columbia University Press.

Simpson, G. G. (1984). *Tempo and mode in evolution*. Columbia University Press.

Simpson, G. G. (1951). The species concept. *Evolution*, 5(4), 285–298.

Skeels, A., & Cardillo, M. (2019). Reconstructing the geography of speciation from contemporary biodiversity data. *American Naturalist*, 193(2), 240–255. https://doi.org/10.1086/701125

Sonne, J., Dalsgaard, B., Borregaard, M. K., Kennedy, J., Fjeldså, J., & Rahbek, C. (2022). Biodiversity cradles and museums segregating within hotspots of endemism. *Proceedings of the Royal Society B: Biological Sciences*, 289(1981), 20221102. https://doi.org/10.1098/rspb.2022.1102

Stadler, T. (2011). Mammalian phylogeny reveals recent diversification rate shifts. *Proceedings of the National Academy of Sciences*, 108(15), 6187–6192. https://doi.org/10.1073/pnas.1016876108

Stadler, T. (2013). Recovering speciation and extinction dynamics based on phylogenies. *Journal of Evolutionary Biology*, 26(6), 1203–1219. https://doi.org/10.1111/jeb.12139

Stadler, T. (2010). TreeSim: Simulating Phylogenetic Trees. R package version 2.4. https://CRAN.R-project.org/package=TreeSim

Stadler, T., & Bokma, F. (2013). Estimating speciation and extinction rates for phylogenies of higher taxa. *Systematic Biology*, 62(2), 220–230. https://doi.org/10.1093/sysbio/sys087

Stadler, T., Gavryushkina, A., Warnock, R. C., Drummond, A. J., & Heath, T. A. (2018). The fossilized birth-death model for the analysis of stratigraphic range data under different speciation modes. *Journal of Theoretical Biology*, 447, 41–55. https://doi.org/10.1016/j.jtbi.2018.03.005

Swenson, N. G. (2019). *Phylogenetic ecology: A history, critique, and remodeling*. University of Chicago Press.

Tanentzap, A. J., Brandt, A. J., Smissen, R. D., Heenan, P. B., Fukami, T., & Lee, W. G. (2015). When do plant radiations influence community assembly? The importance of historical contingency in the race for niche space. *New Phytologist*, 207(2), 468–479. https://doi.org/10.1111/nph.13362

Tanentzap, A. J., Igea, J., Johnston, M. G., & Larcombe, M. J. (2020). Does evolutionary history correlate with contemporary extinction risk by influencing range size dynamics? *American Naturalist*, 195(3), 569–576. https://doi.org/10.1086/707207

Tonini, J. F. R., Beard, K. H., Ferreira, R. B., Jetz, W., & Pyron, R. A. (2016). Fully-sampled phylogenies of squamates reveal evolutionary patterns in threat status. *Biological Conservation*, 204, 23–31. https://doi.org/10.1016/j.biocon.2016.03.039

Upham, N. S., Esselstyn, J. A., & Jetz, W. (2019). Inferring the mammal tree: Species-level sets of phylogenies for questions in ecology, evolution, and conservation. *PLoS Biology*, 17(12), e3000494. https://doi.org/10.1371/journal.pbio.3000494

Van Valen, L. (1973). A new evolutionary law. *Evolutionary Theory*, 1, 1–30.

Vaux, F., Trewick, S. A., & Morgan-Richards, M. (2016). Lineages, splits and divergence challenge whether the terms anagenesis and cladogenesis are necessary. *Biological Journal of the Linnean Society*, 117(2), 165–176. https://doi.org/10.1111/bij.12665

Verde Arregoitia, L. D., Blomberg, S. P., & Fisher, D. O. (2013). Phylogenetic correlates of extinction risk in mammals: Species in older lineages are not at greater risk. *Proceedings of the Royal Society B: Biological Sciences*, 280(1765), 20131092. https://doi.org/10.1098/rspb.2013.1092

Wagner, P. J., Erwin, D. H., & Anstey, R. L. (1995). Phylogenetic patterns as tests of speciation models. In *New approaches to speciaiton in the fossil record* (pp. 87–122). Columbia University Press.

Warnock, R. C., Heath, T. A., & Stadler, T. (2020). Assessing the impact of incomplete species sampling on estimates of speciation and extinction rates. *Paleobiology*, 46(2), 137–157. https://doi.org/10.1017/pab.2020.12

Wiens, J. J. (2023). How many species are there on Earth? Progress and problems. *PLoS Biology*, 21(11), e3002388. https://doi.org/10.1371/journal.pbio.3002388

Wiens, J. J., & Donoghue, M. J. (2004). Historical biogeography, ecology and species richness. *Trends in Ecology & Evolution*, 19(12), 639–644. https://doi.org/10.1016/j.tree.2004.09.011

Willmann, R., & Meier, R. (2000). A critique from the Hennigian species concept perspective. In *Species concepts and phylogenetic theory: A debate* (pp. 101–118). Columbia University Press.

Yang, Z., & Rannala, B. (1997). Bayesian phylogenetic inference using DNA sequences: A Markov Chain Monte Carlo method. *Molecular*

Biology and Evolution, 14(7), 717–724. https://doi.org/10.1093/oxfordjournals.molbev.a025811

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.