

## Estimating Age-Dependent Extinction: Contrasting Evidence from Fossils and Phylogenies

data, citation and similar papers at [core.ac.uk](http://core.ac.uk)

brought

provided by Serveur

<sup>1</sup>Swiss Federal Research Institute WSL, 8903 Birmensdorf, Switzerland; <sup>2</sup>Landscape Ecology, Institute of Terrestrial Ecosystems, ETH Zurich, 8092 Zurich, Switzerland; <sup>3</sup>Department of Biological and Environmental Sciences, University of Gothenburg, SE-405 30 Göteborg, Sweden; <sup>4</sup>Gothenburg Global Biodiversity Centre, Box 461, SE-405 30 Göteborg, Sweden; <sup>5</sup>Departamento de Ecologia, Universidade de São Paulo, 05508-900 São Paulo, Brazil; <sup>6</sup>Gothenburg Botanical Garden, Carl Skottsbergs gata 22A, SE-413 19 Göteborg, Sweden; and <sup>7</sup>Department of Computational Biology, University of Lausanne, 1015 Lausanne, Switzerland  
\*Correspondence to be sent to: [D.S. daniele.silvestro@bioenv.gu.se](mailto:D.S.daniele.silvestro@bioenv.gu.se)

Received 19 December 2016; reviews returned 3 March 2017; accepted 15 October 2017

Associate Editor: Michael Alfaro

**Abstract.**—The estimation of diversification rates is one of the most vividly debated topics in modern systematics, with considerable controversy surrounding the power of phylogenetic and fossil-based approaches in estimating extinction. Van Valen’s seminal work from 1973 proposed the “Law of constant extinction,” which states that the probability of extinction of taxa is not dependent on their age. This assumption of age-independent extinction has prevailed for decades with its assessment based on survivorship curves, which, however, do not directly account for the incompleteness of the fossil record, and have rarely been applied at the species level. Here, we present a Bayesian framework to estimate extinction rates from the fossil record accounting for age-dependent extinction (ADE). Our approach, unlike previous implementations, explicitly models unobserved species and accounts for the effects of fossil preservation on the observed longevity of sampled lineages. We assess the performance and robustness of our method through extensive simulations and apply it to a fossil data set of terrestrial Carnivora spanning the past 40 myr. We find strong evidence of ADE, as we detect the extinction rate to be highest in young species and declining with increasing species age. For comparison, we apply a recently developed analogous ADE model to a dated phylogeny of extant Carnivora. Although the phylogeny-based analysis also infers ADE, it indicates that the extinction rate, instead, increases with increasing taxon age. The estimated mean species longevity also differs substantially, with the fossil-based analyses estimating 2.0 myr, in contrast to 9.8 myr derived from the phylogeny-based inference. Scrutinizing these discrepancies, we find that both fossil and phylogeny-based ADE models are prone to high error rates when speciation and extinction rates increase or decrease through time. However, analyses of simulated and empirical data show that fossil-based inferences are more robust. This study shows that an accurate estimation of ADE from incomplete fossil data is possible when the effects of preservation are jointly modeled, thus allowing for a reassessment of Van Valen’s model as a general rule in macroevolution. [Age-dependent extinction; diversification; macroevolution; PyRate; Bayesian inference.]

*“One of the first ‘rules of biodiversity’ to be investigated concerns lineage age and extinction risk: are ‘ancient’ lineages more or less likely to survive than newly emerged lineages? If this can be resolved, then there would be a means to estimate one factor in the extinction risk of living taxa.”*

(Benton 2016, p. 2).

The potential links between extinction probability and a species’ age have been debated for several decades. Van Valen (1973) proposed that the probability of extinction of a taxon is independent of the elapsed time since its origin (hereafter taxon age). Based on log-linear longevity patterns within ecologically similar taxonomic groups, Van Valen’s Red Queen hypothesis states that species constantly struggle for limited resources. Moreover, it states that all species within ecologically homogenous groups are equally likely to become extinct. Subsequently, several studies seeking evidence in favor of, or against, the idea of age-independent extinction (sometimes proposing alternative theories) flourished (Raup 1975; Van Valen 1977; Pearson 1995; Doran et al. 2006; Liow et al. 2011).

The alternative hypothesis that extinction probability is related to taxon age, i.e. age-dependent extinction

(ADE), has great biological relevance in our understanding of macroevolutionary and ecological processes (Doran et al. 2006; Benton 2016). Species age and probability of extinction have been correlated with several factors, including changes in population size, ecological traits, the nature of competition, coevolutionary processes, genetic mechanisms, and species geographic range (Muller 1964; Pearson 1995; Parker and Arnold 1997; Doran et al. 2006; Finnegan et al. 2008; Johnson et al. 2011; Pigot et al. 2012). For instance, self-pollination in plants, and the consequent accumulation of deleterious mutations, has been hypothesized to increase extinction risk as a function of species age (Muller 1964; Johnson et al. 2011). In planktonic foraminifera, trilobites, conodonts, and graptolites, increased extinction risk in older taxa has been suggested to correlate with evolutionary stasis and loss of competitive ability or with the specialization of older taxa resulting in higher sensitivity to biotic and abiotic environmental changes (Pearson 1995). Ezard et al. (2011) similarly found an increase in extinction risk with taxon age for the macroperforate clade of Cenozoic planktonic foraminifera when taking into account climate change, ecology, and diversity dependence into their model. Other studies on planktonic foraminifera

similarly found extinction risk to increase with age (Pearson 1995; Parker and Arnold 1997; Doran et al. 2006; Ezard et al. 2011). Although Finnegan et al. (2008) encountered similar patterns for foraminifera, most of the marine Phanerozoic genus-level studies (e.g., invertebrate representatives) showed extinction risk to decrease with age. Recently, Crampton et al. (2016) described that ADE, where younger species were at a higher risk, was present during the Ordovician but absent in other time periods for graptolites, a major marine zooplankton group. Taken together, the evidence accumulated so far shows that ADE processes are susceptible to shifts on extinction risk and can change directions and intensity over time.

Based on the predictions of Van Valen's theory, but also because of mathematical convenience and its simplicity as a null hypothesis, most birth–death models used in diversification analyses assume speciation and extinction rates to be age independent, while allowing for rate variation through time and across lineages (e.g., Foote 2003; Nee 2006; Rabosky 2006; Liow and Nichols 2010; Morlon et al. 2011; Stadler 2013a; Silvestro et al. 2014b; May et al. 2016). Under a birth–death process with constant speciation and extinction rates, the expected species longevity, that is, the time between its origination and extinction, is exponentially distributed. However, if speciation or extinction probabilities are dependent on taxon age, the resulting species longevity no longer follow an exponential distribution. Instead, Weibull distributions have been used to model the distribution of species longevity under age-dependent birth–death processes (Pearson 1995; Ezard et al. 2011; Jones 2011; Hagen et al. 2015). In such cases, the shape parameter determines the difference between a distribution where young species are more prone to extinction than older ones (shape < 1) and the alternative case where extinction rates increase linearly with taxon age (shape > 1, Fig. 1a, b). This model becomes equivalent to age-independent extinction when the shape parameter is equal to 1, thus conforming to exponentially distributed longevity.

Most of the statistical methods used to test for ADE are based on fossil data, given that the paleontological record represents the most direct information about extinct species and their longevity. Observed longevity of fossil taxa, that is, species or higher rank lineages, have been used to test whether their distribution conforms to expectations based on Van Valen's hypothesis (exponential distribution) or whether they follow the predictions of different ADE models (e.g. Weibull distribution) (Van Valen 1973; Raup 1975; Parker and Arnold 1997).

Methodological advancements have led to the development of more complex models that are able to combine ADE with time-variable background extinction rates and morphological covariates (Pearson 1992; Doran et al. 2006; Ezard et al. 2011; Wiltshire et al. 2014; Smits 2015). Although applications of these methods have contributed to very important discussions following Van Valen's proposal (Van Valen 1973), most studies so

far have dealt with exceptionally well-preserved fossil data, for example, planktonic foraminifera (species level) or trilobites (genus level). These methods assume the longevity of species that are calculated based on their first and last appearances to be known and the pool of sampled species longevity to represent an unbiased sample of the longevity of all species. However, it would seem crucial to explicitly account for missing lineages and the effects of preservation on the observed longevity of sampled taxa, when investigating ADE processes in organisms with lower fossilization potential and incomplete records (the vast majority of taxa). Assuming the fossil record is the result of a stochastic process of preservation, which includes fossilization and sampling, it follows that observed longevity of lineages drastically underestimate their true longevity (Foote and Miller 2007; Liow and Stenseth 2007; Silvestro et al. 2014b). Furthermore, the pool of taxa for which at least one fossil record is available represents a nonrandom sample, because lineages with short longevity are less likely to be preserved than long-lived lineages. These sampling biases can have a significant impact on the distribution of observed lineage longevity (Foote and Raup 1996), which, in turn, affects our ability to correctly recover ADE from raw fossil data.

Recently, Alexander et al. (2016) developed a method to infer ADE from phylogenies of extant taxa, building on models to infer speciation and extinction dynamics from dated phylogenies (Lambert 2010; Lambert and Stadler 2013). The approach assumes constant speciation rate and models the distribution of species longevity using a Gamma distribution, which is qualitatively similar to the Weibull distribution used in other ADE studies (see above). The method was shown to be powerful, but its applicability is limited to clades that are diverse today, that is, with at least hundreds of species. In addition, to our knowledge, the sensitivity of this method to model violations (e.g., rate variation through time) has not yet been investigated. Further, we are not aware of any empirical study that has compared the results from this phylogenetic approach with evidence from the fossil record.

In this study, we propose a new method to infer ADE from fossil occurrence data. Instead of using only the first and last appearance of a fossil lineage as a measure of its longevity, our method uses all suitable fossil data available, including singletons (here defined as taxa for which only a single fossil occurrence is known) and extant species. We use a Bayesian framework to jointly infer: 1) the true longevity of sampled lineages, while correcting for the effects of incomplete sampling and modeling unobserved lineages, 2) the preservation rate and its variation through time, and 3) the parameters of a Weibull distribution describing the distribution of taxon longevity and the relationship between extinction rate and taxon age.

Our method is able to estimate ADE and mean species longevity while explicitly accounting for missing lineages, that is, taxa that did not leave any fossil record. We assess the performance of our model through

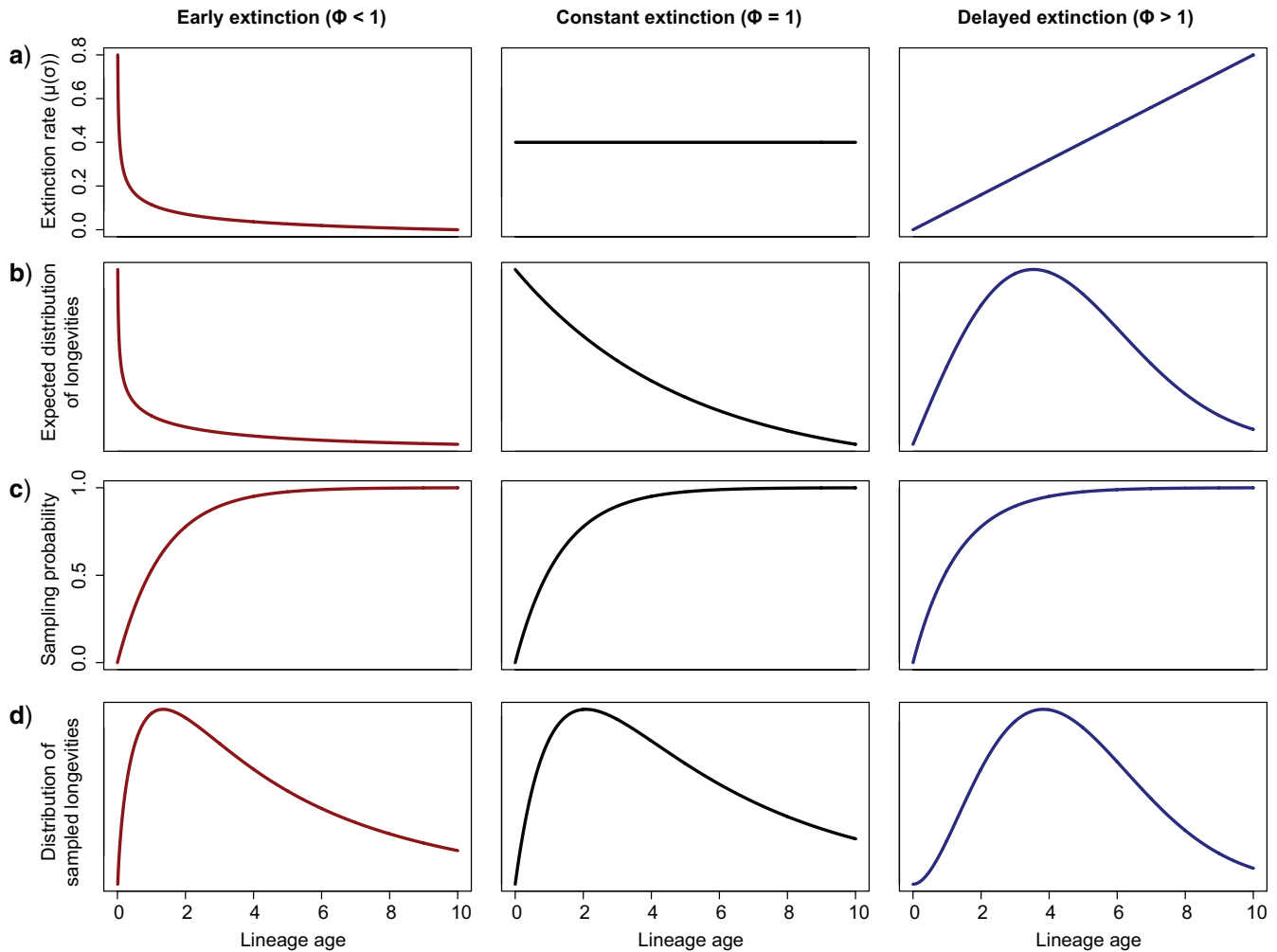


FIGURE 1. Effects of distinct extinction processes and sampling on the distribution of species longevities. Columns from left to right: ADE with higher extinction probability for younger species, age-independent extinction, and ADE with higher extinction probability for older species with a) extinction rates as a function of species age, b) expected distribution of species longevities (Weibull distributions), c) sampling probability based on a uniform Poisson process of preservation with preservation rate  $q = 0.75$ , and d) distribution of longevities resulting from the combination of the distribution of longevities and sampling probability. All plots are against lineage age, that is, time elapsed since lineage origination.

extensive simulations encompassing a wide range of diversification and preservation scenarios and test the robustness of this method to model violations. Finally, we apply our method to a recently compiled fossil data set of mammalian carnivores (Pires et al. 2015) and compare our results with an analogous ADE analysis that was run using a molecular phylogeny of extant carnivores (Nyakatura and Bininda-Emonds 2012) applying the method of Alexander et al. (2016).

## METHODS

We present here a method to estimate ADE using fossil data, consisting of a set of lineages for which one or more dated fossil records are available. The data may include extant lineages, provided that at least one fossil occurrence was sampled prior to the present. Our

method builds upon the PyRate analytical framework (Silvestro et al. 2014a,b), where fossil occurrences are used to jointly infer: 1) the true times of origination and extinction of all sampled lineages, 2) the parameters of a preservation process (comprising fossilization, sampling, and identification), and 3) the underlying diversification process determining the frequency and temporal distribution of speciation (or origination if dealing with higher taxa) and extinction events. In our notation, the age of events is expressed as time before present.

### The ADE Model

We have developed an ADE model in which the probability for a lineage to become extinct is a function of its age, defined as the elapsed time since its

origination. Based on the ADE model described by Hagen et al. (2015), we define the extinction rate of a taxon of age  $\sigma$  as:

$$\mu(\sigma) = \frac{\Phi}{\Psi} \left( \frac{\sigma}{\Psi} \right)^{\Phi-1}, \quad (1)$$

where the parameters  $\Phi$  and  $\Psi$  determine how extinction changes as a function of age (Fig. 1a). Under this model, the resulting waiting time until extinction follows a Weibull distribution (Weibull 1939), where  $\Phi \in R^+$  is the shape parameter and  $\Psi \in R^+$  is the scale parameter (Parker and Arnold 1997; Hagen et al. 2015) (Fig. 1b). The value of the shape parameter ( $\Phi$ ) changes the effect of age dependence qualitatively. More specifically, for a shape  $\Phi < 1$ , the extinction rate is higher in younger taxa, declining with taxon age. In contrast, when  $\Phi > 1$ , extinction probability increases with taxon age. Finally, if  $\Phi = 1$ , the extinction process is age independent with a constant rate, thus reducing to a constant rate birth–death model (Fig. 1a, b).

Silvestro et al. 2014b derived the likelihood function for constant rate birth–death models based on a set of fossil lineages with estimated speciation and extinction times indicated by  $s$  and  $e$ . The likelihood function can incorporate ADE by letting the extinction rate be a function of lineage longevity,  $\sigma = s - e$ . Thus, the likelihood of an extinct lineage under an ADE birth–death model with constant speciation rate  $\lambda$  is:

$$P(s, e | \lambda, \Phi, \Psi) \propto \lambda^N \exp(-\lambda\sigma) \times \mu(\sigma) \times \exp\left(-\int_s^e \mu(\sigma) d\sigma\right). \quad (2)$$

Under the ADE model described above, the longevities of lineages follow a Weibull distribution and the likelihood of an extinct lineage with longevity  $\sigma = s - e$  is given by the Weibull probability density function:

$$P(\sigma | \Phi, \Psi) = \frac{\Phi}{\Psi} \left( \frac{\sigma}{\Psi} \right)^{\Phi-1} e^{-(\sigma/\Psi)^\Phi}. \quad (3)$$

#### ADE Models for Incomplete Fossil Data

The inevitable incompleteness of fossil data makes it impossible to observe the entire temporal range of existence of a lineage, as the past history of a taxon is usually only known through a limited number of fossil occurrences. The incompleteness of the fossil record can bias the data in two main ways. First, the observed longevity of a lineage, based on its first and last appearances, is likely to underestimate its true longevity from the time of its origination to the time of extinction (Liow and Stenseth 2007; Silvestro et al. 2014a). Second, some lineages, and especially short-lived taxa, are likely to leave no traces of their existence in the fossil record. Both biases become more severe as preservation rates decrease. Thus, explicitly modeling the preservation process is crucial to estimating origination and extinction

dynamics from fossil data (Foote 2000; Liow and Nichols 2010; Silvestro et al. 2014b).

The first bias can be explicitly addressed by estimating the true times of origination and extinction of a lineage from the fossil data, based on a preservation process. In PyRate, fossil preservation is modeled as a Poisson process with rate  $q$  expressing the expected number of fossil occurrences per lineage per time unit (typically myr) (Silvestro et al. 2014b). A hierarchical Bayesian algorithm enables the joint estimation of the preservation rate and the origination and extinction times using the Markov chain Monte Carlo technique (MCMC; Silvestro et al. 2015b).

The second bias, linked to the fact that short-living taxa are less likely to leave a fossil record than those that are long-living, has not been explicitly addressed in previous PyRate implementations, although it has been shown through simulations that incomplete taxon sampling has limited effects on the accuracy of the estimation of speciation and extinction rates (Silvestro et al. 2014b). The probability of not sampling any fossils for a lineage with longevity equal to  $\sigma$  under a homogeneous Poisson process (HPP) with preservation rate  $q$  is:

$$P(K=0|q, \sigma) = \exp(-q\sigma), \quad (4)$$

that is, the probability of a waiting time of length  $\sigma$  during which no fossil occurrences ( $K=0$ ) are sampled. Thus, the probability for a lineage to be sampled, that is, to have at least one fossil occurrence, is a function of the preservation rate and its longevity (Silvestro et al. 2014b):

$$P(K>0|q, \sigma) = 1 - \exp(-q\sigma). \quad (5)$$

As a result, for a given preservation rate, short-living lineages are less likely to be sampled in the fossil record than those that are long-living (Fig. 1c). This bias has the effect of strongly altering the observed distribution of taxa longevities, which no longer follow a Weibull distribution (Fig. 1d). Instead, the distribution of observed longevities, given a Poisson preservation process and an ADE model of diversification, is proportional to:

$$g(\sigma) = P(\sigma | \Phi, \Psi) \times P(K>0|q, \sigma). \quad (6)$$

The normalized probability of sampling an extinct lineage is therefore:

$$P(\sigma | \Phi, \Psi, q) = g(\sigma) / \int_0^\infty g(x) dx. \quad (7)$$

Extant taxa are lineages whose longevity is truncated at the present (Time 0). Here, we indicate the observed longevity of an extant lineage as  $\sigma_t$  (i.e. time since its origination), emphasizing that only extant lineages with at least one fossil record are included in the analyzed data. The probability of sampling an extant lineage is given by the probability that it left at least one fossil record since its time of origin and that it has not become

extinct before the present:

$$P(\sigma_i | \Phi, \Psi, q) = 1 - \frac{\int_0^{\sigma_i} g(x) dx}{\int_0^{\infty} g(x) dx}. \quad (8)$$

The expected distribution of sampled longevities (Fig. 1d) may be further altered by changes in preservation rates through time. Many empirical studies have shown that fossil preservation is not a constant process (e.g. Foote 2000, 2003; Liow et al. 2015; Pires et al. 2015; Silvestro et al. 2016). Here, we relaxed the assumption of constant preservation by allowing the preservation rates to vary at predefined times of shift. The preservation process is, therefore, modeled by a non-homogeneous Poisson process in which the rate is constant within time frames delimited by times of shifts (e.g. Bacon et al. 2015). The probability of sampling a fossil lineage originating at time  $s$  and becoming extinct at time  $e$  under a non-homogeneous Poisson process with a time-varying preservation rate is:

$$P(K > 0 | q, s, e) = 1 - \exp\left(-\int_s^e q(t) dt\right), \quad (9)$$

where  $q(t)$  is the preservation rate at time  $t$ . In our current implementation, the number and temporal placement of rate shifts is predefined and fixed.

### Implementation

We implemented the ADE model in a Bayesian framework to jointly estimate 1) the origination and extinction times of all sampled lineages, 2) the preservation rates, and 3) the parameters of the ADE model (i.e., the shape and scale parameters of the Weibull distribution). For a given lineage  $i$  with fossil occurrences  $\mathbf{x} = [x_1, \dots, x_K]$ , we use MCMC to sample the parameters from the following posterior distribution:

$$P(s_i, e_i, q, \Phi, \Psi, | \mathbf{x}) \propto P(\mathbf{x} | s_i, e_i, q) \times P(s_i - e_i | \Phi, \Psi, q) \times P(\Phi) P(\Psi) P(q), \quad (10)$$

where the first term of the product represents the likelihood of fossil occurrences, given the times of origination and extinction ( $s, e$ ), and the preservation rate  $q$ , based on a Poisson process of preservation (Silvestro et al. 2014b). The second term provides the probability of sampling a lineage with longevity  $\sigma = s - e$  as described in Equations (7) and (8). The third term includes prior distributions for the Weibull parameters and preservation rates. We used gamma-distributed priors for both preservation rates and the Weibull scale ( $\Psi$ ). Since the null hypothesis for the shape parameter is  $\Phi = 1$ , that is, age-independent extinction, we used a prior with the highest density at 1 by defining it as  $P(\log(\Phi)) = N(0, 2)$ .

Times of origination and extinction are updated using sliding window proposals, whereas preservation rates and Weibull parameters are updated using multiplier proposals (e.g., Ronquist et al. 2007). When running

under a model with shifts in preservation rates, we consider the rates between shifts as independent and estimate them from the data.

### Simulations under the ADE Model

To test the performance of our method, we simulated fossil data under ADE models and various preservation rates. We generated phylogenies under a birth–death process with constant speciation and ADE rates using the *TreeSimGM* R-package (Hagen et al. 2015; Hagen and Stadler 2017). From these phylogenies, we extracted the true speciation and extinction times for all lineages and used them to simulate fossil occurrences under Poisson models of preservation, with constant or time-varying preservation rates. We set the root age of phylogenies to 20 myr and the speciation rate ( $\lambda$ ) to 0.4 in all simulations. To assess the different properties of our ADE model, we generated fossil data sets under four birth–death and preservation scenarios, explained below and summarized in Table 1.

In the first scenario, we simulated lineages under an ADE model with Weibull distributed longevities. For each simulation, we drew the shape parameter of the Weibull distribution from a uniform distribution  $\Phi \sim U[0.5, 2]$ , thus reproducing cases of high infant extinction ( $\Phi < 1$ ), roughly constant extinction probability ( $\Phi \approx 1$ ), and highly delayed extinction ( $\Phi > 1$ ). To avoid unrealistic longevities of lineages, we drew values for the scale parameters ( $\Psi$ ) based on a predefined target mean longevity  $\bar{\sigma}$ . Given a target mean longevity  $\bar{\sigma}$  and a shape parameter  $\Phi$ , the scale parameter is calculated as  $\Psi = \bar{\sigma}_n / \Gamma(1 + 1/\Phi)$ . We randomly sampled the target  $\bar{\sigma}$  for each simulation from a uniform distribution  $\bar{\sigma} \sim U[3, 8]$ . After obtaining the lineages generated under the ADE model, we sampled fossil occurrences according to a constant Poisson process with preservation rate  $q$  drawn at random from a uniform distribution  $q \sim U[0.5, 1.5]$ . The range of preservation rates (expected number of fossil occurrences per lineage/myr) was defined based on values previously estimated from different empirical data sets (e.g., Pires et al. 2015; Silvestro et al. 2015a).

The simulations under the second scenario replicated the same ADE settings used in Scenario I but differed in the preservation process generating the fossil record. Here, we set the preservation rates to shift at time 10 myr. We drew the preservation rates before and after the shift as independent random numbers from uniform distributions  $q_1 \sim U[0.5, 1.5]$ ,  $q_2 \sim U[0.5, 1.5]$  (Table 1).

In the 3rd and 4th scenarios, we introduce a violation of the ADE model assumptions to test the robustness of the method. Specifically, we simulated cases in which the scale parameter ( $\Psi$ ) of the Weibull distribution changes through time. In Scenario III, we simulated a change in the scale parameter, taking place at 5 myr. The initial scale value was drawn based on the same target mean longevity used in previous simulations ( $\bar{\sigma} \sim U[3, 8]$ ), whereas we drew the second scale (after the shift at 5

TABLE 1. Summary of the simulation settings

	ADE simulation scenarios (assessing method accuracy)			
	I	II	III	IV
Speciation rate	$\lambda=0.4$	$\lambda=\lambda 0.4$	$\lambda=\lambda 0.4$	$\lambda=\lambda 0.4$
ADE shape	$\Phi \sim U[0.5, 2]$	$\Phi \sim U[0.5, 2]$	$\Phi \sim U[0.5, 2]$	$\Phi \sim U[0.5, 2]$
ADE shifts through time	No shift	No shift	One at 5 Ma	One at 5 Ma
ADE scale at first time bin	$\Psi$ corrected for $\bar{\sigma} \sim U[3, 8]$	$\Psi$ corrected for $\bar{\sigma} \sim U[3, 8]$	$\Psi_1$ corrected for $\bar{\sigma} \sim U[3, 8]$	$\Psi_1$ corrected for $\bar{\sigma} \sim U[3, 8]$
ADE scale at second and last time bin	—	—	$\Psi_2 \sim U[0.5\Psi_1, 2\Psi_1]$	$\Psi_2 = 0.2\Psi_1$
Fossil preservation	$q \sim U[0.5, 1.5]$	$q_1 \sim U[0.5, 1.5]$ shift at 10 Ma $q_2 \sim U[0.5, 1.5]$	$q_1 \sim U[0.5, 1.5]$ shift at 5 Ma $q_2 \sim U[0.5, 1.5]$	$q_1 \sim U[0.5, 1.5]$ shift at 5 Ma $q_2 \sim U[0.5, 1.5]$
	False-positive simulations (assessing frequency of Type I errors)		Sensitivity to violations to model assumptions (strong rate variation)	
Speciation and Extinction process	Age-independent and constant though time		Age-independent and variable though time	
Shifts though time	No shift		$n \sim \text{Poisson}(2)$	
Shift position	No shift		Sampled uniformly along total age, i.e. 20 Myr, of the simulated trees	
Speciation rate	$\lambda=0.4$		$\lambda_n \sim U[0.1, 1]$	
Extinction rate	$\mu \sim U[0.333, 0.125]$		$\mu_n \sim U[0.1, 1]$ with $\mu_{n=1} = \mu_{n=1} \times 10$	
Fossil preservation	$q \sim U[0.5, 1.5]$		$q \sim U[0.5, 1.5]$	

myr) from  $\Psi_2 = r\Psi_1$ , where  $r \sim \exp(U[\log(0.5), \log(2)])$ . Thus, we simulated up to 2-fold changes in the scale parameters, corresponding to up to 2-fold changes in extinction rates (Equation (1)), in the case of age-independent extinction ( $\Phi = 1$ ). As in Scenario II, we simulated fossil occurrences under a Poisson process with rates ranging from 0.5 to 1.5 and one shift in preservation at 5 myr.

In Scenario IV, we introduce a stronger violation of the model assumption by simulating a very strong decrease in scale ( $\Psi_2 = 0.2\Psi_1$ ), which corresponds to a 5-fold increase in the extinction rate in the case of age-independent extinction ( $\Phi = 1$ ). Under this scenario, an average of about 98% of the total number of lineages in the clade become extinct before the present (Table 2). The settings for the preservation process followed those of Scenarios II and III.

The number of sampled lineages included in each simulation was variable since birth–death simulations were based on constraining the root age and were not conditioned on the number of tips (Hagen et al. 2015). Moreover, preservation rates reduced the number of sampled lineages stochastically. In this way, we simulated 500 data sets under each of the four scenarios and filtered them to remove exceedingly small or large data sets, that is, with fewer than 20 lineages or more than 600 lineages. The number of resulting data sets, together with additional summary statistics, is reported in Table 2. Finally, from all four scenarios, a total of 1145 simulated data sets were included in the analyses (Table 2).

### Data Analysis

We analyzed each simulated data set by running 500,000 MCMC iterations, sampling every 500. We checked the runs for convergence using Tracer v1.6 (Rambaut et al. 2014) after excluding the first 100,000 iterations as burn-in. We summarized the posterior samples of the parameters of interest (preservation rates, shape, and scale) by calculating their posterior means and their 50% and 95% highest posterior density intervals (HPD).

To assess the method's performance, we quantified the accuracy of the estimated parameter values. The overall accuracy of the estimates, generalized as  $\hat{x}$ , was quantified as the mean absolute percentage error (MAPE), defined as  $\text{MAPE} = \frac{1}{n} \sum_i \frac{|\hat{x}_i - x_i|}{x_i}$ , where  $n$  is the number of simulations,  $\hat{x}$  is the estimated parameter value (posterior mean), and  $x$  is the true value. We also plotted the relative errors of the parameter estimates, defined as  $\text{RE} = (\hat{x} - x)/x$ , against the preservation rate and against the number of taxa. These plots allowed us to verify whether varying levels of incompleteness in fossil records and different data set sizes affect the accuracy of parameter estimates.

We assessed the ability to correctly estimate whether extinction is age dependent by looking at how frequently the Weibull shape parameter ( $\Phi$ ) was correctly estimated to be significantly less or greater than 1, that is, age-independent extinction. We first calculated the frequency of erroneous ADE estimates, here defined as the proportion of simulations in which  $\Phi$  was

TABLE 2. Summary statistics for the simulated data sets under each scenario

		Scenario			
		I	II	III	IV
Number of simulations		274	342	246	283
Total number of species in the simulated tree, including extinct lineages	Mean	234.1	234.1	245.4	146.4
	Min	47	47	39	29
	Max	869	869	683	453
Number of sampled species in the fossil data	Mean	168.14	141.83	175.95	104.64
	Min	37	23	33	21
	Max	588	556	486	299
Percentage of extant species in the fossil data	Mean	0.49	0.44	0.5	0.03
	Min	0.22	0.04	0.11	0
	Max	0.81	0.76	0.83	0.18
Sampling proportion	Mean	0.72	0.72	0.71	0.7
	Min	0.48	0.46	0.51	0.39
	Max	0.93	0.92	0.89	0.9
Number of fossil occurrences	Mean	567.53	487.64	591.74	354.02
	Min	88	55	95	51
	Max	2063	1859	1836	1095
Proportion of singletons	Mean	0.14	0.16	0.13	0.3
	Min	0.02	0.03	0.03	0.12
	Max	0.36	0.36	0.32	0.61

erroneously estimated to be significantly less than 1 when the true value was greater than 1, or  $\Phi$  was erroneously estimated to be significantly greater than 1 when the true value was less than 1. We considered the estimated shape parameter to be significantly different from 1 (less than or greater than) when 1 did not fall inside its 95% HPD. Similarly, we calculated the frequency of true positives as the proportion of simulations in which  $\Phi$  was correctly estimated to be significantly less (or greater) than 1. We also repeated these calculations using the 50% HPD to assess a different level of significance.

#### Type I Error

We measured the frequency of false positives (Type I error) by analyzing 303 data sets simulated under age-independent extinction. We simulated clades under similar settings to those in previous simulations (e.g., Scenario I), but fixed the shape parameter of the Weibull distribution to  $\Phi=1$  (Table 1). We generated trees with 20–600 lineages (including both extinct and extant) under constant speciation (set to 0.4) and mean longevity randomly drawn from  $\bar{\sigma} \sim U[3,8]$ . We sampled fossil occurrences from a uniform Poisson process with the preservation rate randomly drawn from  $U[0.5,1.5]$ , analyzed the clades using the ADE model and calculated the 95% HPD of the shape parameter. The frequency of false positives was computed as the proportion of simulations in which the 95% HPD of the shape parameter did not include 1.

To further explore potential biases, we investigated the effect of anagenesis (resulting in the extinction

of a morphospecies and origination of a new one) on our inference of the ADE model. We simulated the longevities of 200 lineages under ADE (Weibull shape set to 2 or 0.5) and under age-independent rate of extinction (Weibull shape equal to 1). The lineages were subjected to several levels of anagenesis. This involved partitioning them into morphospecies, assuming a constant rate of anagenesis leading to the disappearance of a morphospecies and its replacement by a new one. We generated 100 simulations under each setting and assessed the effects of anagenesis on the estimated Weibull shape parameter (see [Supplementary Fig. S6](http://dx.doi.org/10.5061/dryad.r5f70) available on Dryad at <http://dx.doi.org/10.5061/dryad.r5f70>, for more details).

#### Sensitivity of ADE Models to Strong Model Violations

We performed additional simulations (Table 1) to assess the impact of strong model violations (numerous and intense rate changes through time) on the estimation of ADE parameters. In particular, we explored whether time-variable speciation and extinction rates in the absence of age dependency can induce artificial evidence for ADE. Understanding the sensitivity of ADE models to temporal rate variation is crucial to assess the applicability of the method to empirical data sets, where speciation and extinction are known to often follow non-homogeneous processes (e.g., [Stadler 2011a](#); [Silvestro et al. 2015b](#)).

We simulated 300 fossil data sets in the absence of age dependence but under random time variable birth–death models. We set the root age to 20 and sampled the number of rate shifts in speciation and extinction

rates from Poisson distributions with a mean equal to 2. Speciation and extinction rates between rate shifts were sampled from uniform distributions  $U[0.1, 1]$ —thus simulating up to 10-fold rate changes—and the temporal placement of rate shifts was sampled from uniform distributions  $U[0, 20]$ . The initial extinction rate (between the root and the first time of rate shift) was set to be ten times lower than the initial speciation rate to prevent clades from becoming extinct too soon after their origin. Only simulations with a number of species (extinct or extant) between 150 and 500 were kept (simulations with a greater number of lineages were avoided in the interest of feasible computation time). We simulated fossil occurrences based on a uniform Poisson process with preservation rates randomly sampled from a uniform distribution  $U[0.5, 1.5]$ . We analyzed the data sets under the ADE model using the same MCMC settings described above and calculated the frequency of false positives as the proportion of analyses that estimated the Weibull shape parameter to be significantly less than or greater than 1, based on its 95% HPD.

#### *Comparison with Molecular-based Inferences*

We compared the robustness of our fossil-based model with that of an ADE model recently implemented for phylogenies of extant taxa using the R package *TreePar* (Stadler 2011a; Alexander et al. 2016). Although this model is analogous to that presented here, it assumes the time to extinction to follow a Gamma distribution, rather than a Weibull distribution. The interpretation of its estimated shape parameter is qualitatively the same as that of a Weibull distribution, with values different from one indicating evidence of ADE. The *TreePar* ADE model, as our model, assumes that rates of speciation are constant and that extinction rates are only a function of species age.

We then tested how similar model violations affect the results of the ADE model implemented for phylogenies of extant taxa. We simulated 300 phylogenies of extant taxa under birth–death settings similar to those implemented in the fossil simulations using *TreeSim* (Stadler 2011b). To ensure sufficient power to estimate the model parameters, we fixed the number of tips to 1000 (as in Alexander et al. 2016) and did not set any constraints on the root age. We set the number of rate shifts to 2 (at times 2 Ma and 7 Ma) and sampled random speciation and extinction rates as in the fossil-based simulations (Table 1). We carried out maximum likelihood optimizations under the gamma model (ADE) and under the exponential model (age independent birth–death). Finally, we calculated the frequency of false positives using a likelihood ratio test to assess whether the gamma model was inferred to be significantly better than the exponential model, based on a chi-squared distribution with 1 degree of freedom ( $P > 0.95$ ).

#### *Analysis of Mammalian Carnivores: Fossils and Phylogeny*

We analyzed fossil and phylogenetic data of mammalian carnivores to assess whether there is evidence of ADE and to verify whether the two types of data provide consistent results. The analytical workflow followed for these analyses is summarized in Fig. 2.

We used a fossil data set of Carnivora from the Northern Hemisphere recently compiled and thoroughly cleansed by Pires et al. (2015). The data set included 961 species, 78 of which are extant and 7071 fossil occurrences spanning the past 40 myr (with 38% of singletons). The temporal ranges of the fossil occurrences were comparatively narrow (median range 1.65 myr, 95% confidence interval [CI] 0.006–6.275). To incorporate uncertainties associated with the dating of fossil occurrences (typically expressed as minimum and maximum ages), we resampled the ages of each occurrence to generate 10 randomized data sets, following the procedure described by Silvestro et al. (2014b) and implemented in PyRate. We allowed for temporal variation in the preservation rates by estimating independent rate parameters within each geological epoch from the Eocene to the present, but merged Pleistocene and Holocene into a single time frame, because the latter is considerably shorter than any previous epoch. We then analyzed the 10 resampled data sets (3,000,000 MCMC iterations, sampling every 10,000) under the ADE model and combined all the posterior samples to summarize the estimated shape and scale parameters of the Weibull distribution. We assessed the degree of variability of the estimated shape and scale parameters across resampled data sets by calculating the coefficients of variation among the posterior mean from each data set.

Previous analyses of the Carnivora fossil record have shown that both speciation and extinction rates underwent substantial changes in the early diversification of the clade (Eocene and Oligocene) and towards the present (Quaternary), whereas the diversification process was essentially stable in the Neogene (23.03–2.58 Ma) (Liow and Finarelli 2014; Pires et al. 2015). As we found in our simulations time-varying speciation and extinction rates that resulted in frequent false positives (see Results), we repeated the ADE analyses on a subsample of the Carnivora fossil data set. This data set included exclusively species with all known occurrences found in the Neogene. The subsampled data set encompassed 568 species and 2,948 occurrences and was analyzed under the same settings used for the full data set.

We then analyzed a dated phylogeny of all extant carnivores comprising 286 species, with crown age at about 65 Ma (Nyakatura and Bininda-Emonds 2012) using the *TreePar* package to identify evidence of ADE. We ran maximum likelihood optimizations under the model with gamma-distributed longevities (i.e., ADE) and exponentially distributed longevities (age-independent extinction) and compared the fit of the two



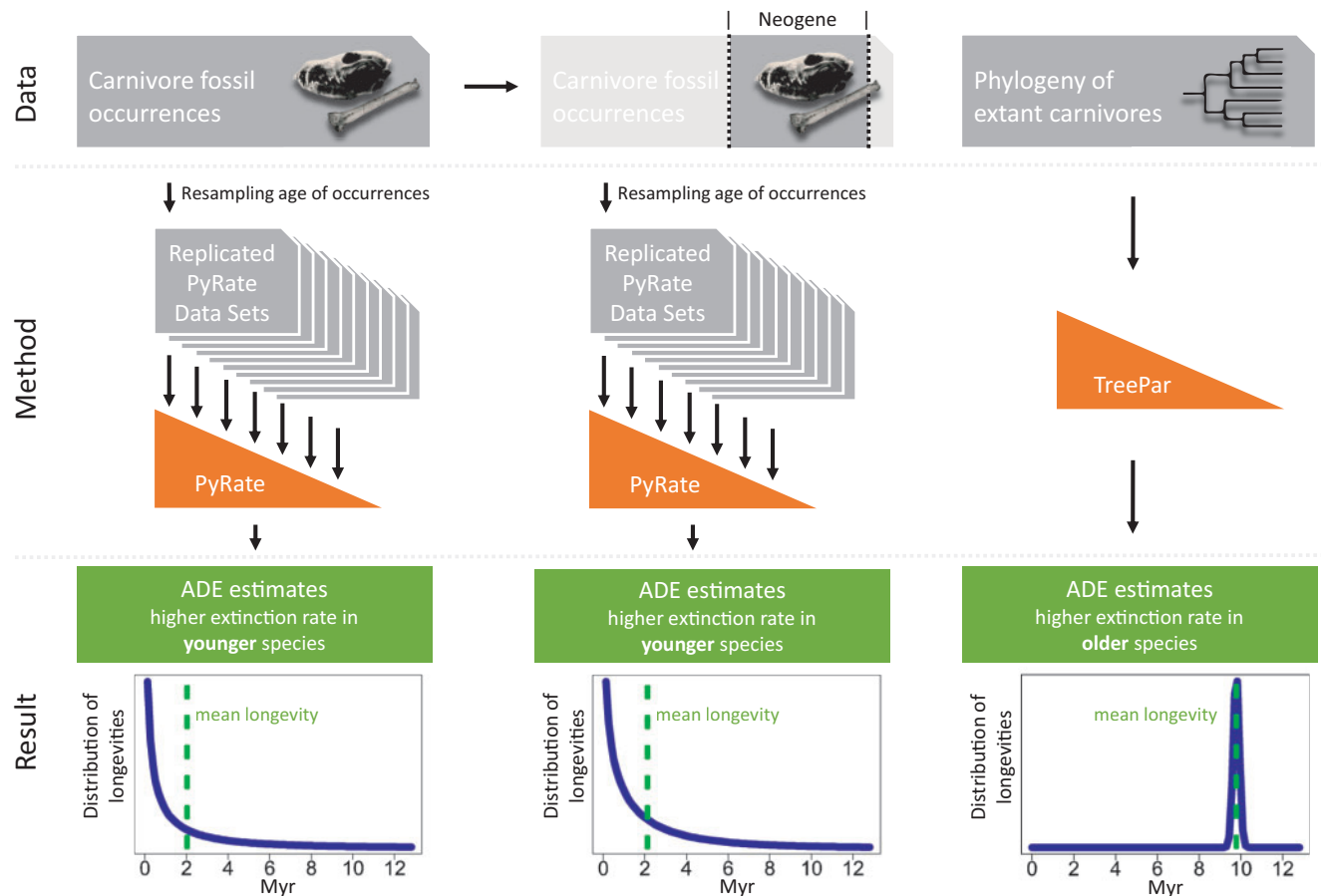


FIGURE 2. Workflow for ADE testing of incomplete fossil and phylogenetic data of terrestrial carnivores (top), the methodology applied in this study (middle), and the main empirical results for carnivores (bottom). To test for possible effects of unstable diversification rates, a subset of the fossil data was taken for the Neogene (23.03–2.58 myr), a period with known diversification stability for carnivores. It was not possible to implement a similar approach for phylogenies. The software used are shown by the triangles. The bottom plots summarize the results of the estimated longevities (straight line) and the mean expected longevity (dashed line). All fossil analyses consistently indicated ADE with higher extinction rates in younger species, whereas the phylogenetic data also suggest ADE but with higher extinction rates in older species.

models using a likelihood ratio test with 1 degree of freedom.

#### Data and Software Availability

The ADE birth–death model was implemented in PyRate (available at: <https://github.com/dsilvestro/PyRate>). Other Python and R scripts, simulated and empirical data used in this study are available from the Dryad.

## RESULTS

### Estimating ADE from Fossil Data

The main parameters of interest in our ADE model are the preservation rates, the shape of the Weibull distribution (which changes the effect of species longevity on extinction rate qualitatively), and the scale of the Weibull distribution (which, for a given shape

value, determines the expected mean species longevity). These parameters were accurately estimated across our simulations (Table 3; Fig. 3). The MAPE around the estimated preservation rates was below 0.1 when the preservation rate was constant (Scenario I; Fig. 3a). In cases of time-varying preservation (Scenarios II–IV), the accuracy of the estimated preservation rate preceding the shift ( $q_1$ ) was slightly lower, but with MAPE, still below 0.2 (Fig. 3b–d). This is possibly a result of the lower number of lineages in the early diversification of clades. Overall, the accuracy of estimated preservation rates increased with higher values of  $q$  (Supplementary Fig. S1 available on Dryad) and an increasing number of lineages (Supplementary Fig. S2 available on Dryad). The proportion of singletons in the data had little effect on the accuracy of the estimated parameters (Supplementary Fig. S3 available on Dryad). The precision of the estimates, here calculated as the size of the 95% CI, was similar across simulations (Supplementary Fig. S4 available on Dryad).

TABLE 3. Accuracy of the estimated ADE parameters

Scenario	Shape	Scale	Preservation ( $q_1$ )	Preservation ( $q_2$ )
I	0.176 (0.175)	0.114 (0.12)	0.067 (0.054)	—
II	0.194 (0.213)	0.142 (0.165)	0.167 (0.152)	0.067 (0.065)
III	0.205 (0.214)	—	0.13 (0.132)	0.065 (0.057)
IV	0.241 (0.273)	—	0.133 (0.109)	0.081 (0.077)

Note: Mean absolute percentage errors (MAPE) calculated across all simulations for each scenario (standard deviations given in parentheses). All parameters were generally estimated with high levels of accuracy. MAPE were not calculated for the scale parameter in Scenarios III and IV, because data were simulated under time-variable scales, whereas a single scale parameter is estimated by our ADE model. Nevertheless, the estimated values can be visually compared to the mean of the simulated scale values (Fig. 2c and d). Parameters  $q_1$  and  $q_2$  in Scenarios II–IV represent the preservation rates before and after a time of shift, respectively.

The shape and scale parameters of the Weibull distribution ( $\Phi, \Psi$ ) were accurately estimated when the data conform to the model assumptions (Scenarios I and II) with MAPE around 0.15 (Table 3; Fig. 3a, b). Simulation with small and moderate violations of the model assumptions (up to 5-fold temporal changes in the scale parameter; Scenarios III and IV) yielded decreased accuracy in the estimated shape (Fig. 3c, d), with MAPE between 0.2 and 0.25. Large values of  $\Phi$  were generally underestimated in Scenario IV. However, the accuracy of the estimated shape and scale parameters increased with an increasing number of lineages under all scenarios (Supplementary Fig. S2 available on Dryad). In contrast, we did not find a strong correlation between preservation rates and accuracy of the ADE parameters, suggesting that the model works equally well under a wide range of preservation rates (Supplementary Fig. S1 available on Dryad).

The proportion of simulations where  $\Phi$  was erroneously estimated to be significantly greater or less than 1 was below 1% under all four scenarios (Table 4), that is, even under moderate violations of the model assumptions (Scenarios III and IV). The power of our approach to significantly support ADE (when true) is higher when the true value of  $\Phi$  is substantially less than or greater than 1 (Fig. 4). For instance, under moderately small or large shape parameters (e.g.,  $\Phi < 0.7$ ,  $\Phi > 1.5$ ), the correct ADE model was identified with statistical confidence in 70–90% of the cases under Scenario I. The power of the method, as expected, decreases when the true value of  $\Phi$  is closer to 1 (Fig. 4). Although small model violations do not appear to strongly affect the frequency of true positives (Fig. 4c), stronger violations have the effect of strongly reducing the power of the method (Fig. 4d). Similarly, the size of the data set influenced the power of the method (Supplementary Figure S5 available on Dryad). Under no or moderate model violation, the correct ADE model was identified with statistical confidence in about 80% of the simulations with data sets of 100 to 150 lineages and the power of the test increased with larger number of lineages (Supplementary Figure S5a, b, c available on Dryad). Under strong model violations (Scenario IV), the power was overall lower (Fig. 4) but still appeared to increase with increasing size of

the data sets (Supplementary Figure S5d available on Dryad).

Our method shows high robustness against false positives. Our additional simulations based on age-independent extinction and constant rates of speciation and extinction found the frequency of Type I errors to be 3.3%, indicating that the method is able to correctly reject ADE in favor of a constant extinction model. Finally, additional simulations (Supplementary Fig. S6 available on Dryad) showed that when the rate of anagenesis (pseudo-extinction and pseudo-speciation) is substantially higher than the rate of extinction, support of ADE tends to weaken in favor of age-independent models. Thus, constant processes of pseudo-extinction and pseudo-speciation do not lead to Type I error (Supplementary Fig. S6 available on Dryad).

#### *ADE in Mammalian Carnivores and Behavior of ADE Models under Strong Model Violations*

The analysis of the fossil record of terrestrial carnivores revealed strong evidence for ADE, where the extinction rate is higher in younger species and decreases with species age (Supplementary Table S1 available on Dryad). The estimated shape parameter of the Weibull distribution was 0.57 (95% HPD 0.50–0.65), that is, significantly less than 1. The estimated scale parameter was 1.25 (95% HPD 0.94–1.57), resulting in an estimated mean species longevity of 2.02 myr (95% HPD 1.76–2.27 myr). The coefficients of variation of the shape and scale parameters estimated across 10 resampled data sets were very small (0.017 for shape and 0.034 for scale), thus indicating that resampling fossil ages from the respective temporal ranges accounted for little variation in the estimates.

A very different evolutionary scenario is inferred through the analysis of a phylogeny of extant carnivore species using *TreePar*. The gamma (ADE) model had a significantly higher likelihood compared with the alternative exponential model (age independent). However, the estimated effect of age on extinction rates contrasted with the fossil-based inference. The shape parameter of the gamma distribution was 4655.11, that is, significantly greater than 1, and therefore indicates that extinction is lower in younger species and increases

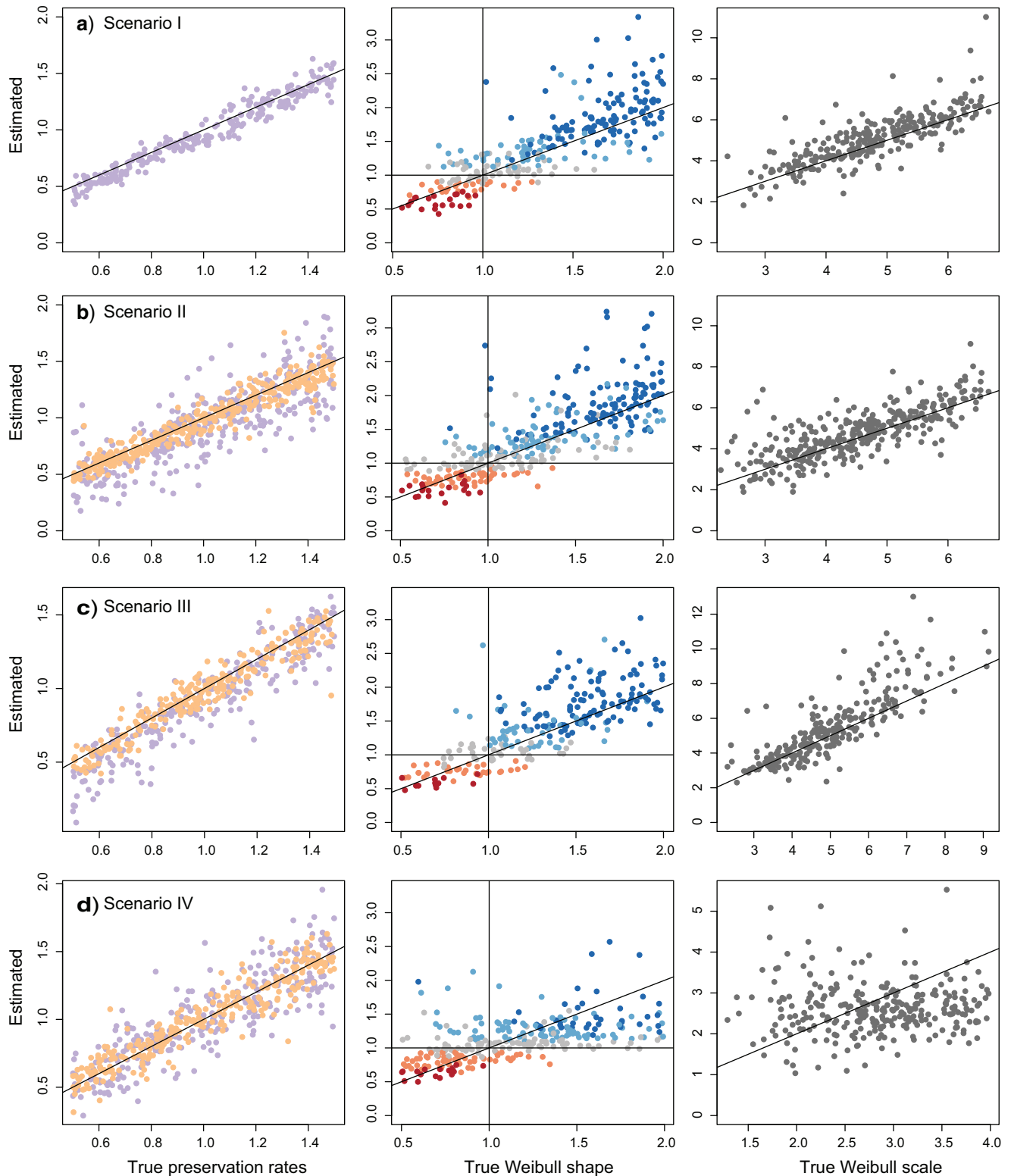


FIGURE 3. Accuracy of parameter estimates under different preservation and ADE settings. Estimated rates (posterior means from a total of 1145 simulations) are plotted against true rates based on different simulation Scenarios I–IV, from left to right: preservation rates ( $\eta$ ), Weibull shape parameter ( $\Phi$ ), and Weibull scale parameter ( $\Psi$ ). Distribution of longevity has curvature dictated by  $\Phi$  while mean longevity is a function of  $\Phi$  and  $\Psi$ . Points below the diagonal lines represent underestimates, points above represent overestimates. For the preservation rate  $\eta$ , estimates of the preservation rate preceding the shift are plotted in purple and preservation rates after the shift in orange (only for Scenarios II–IV). For the shape parameter, values significantly greater or less than 1 based on the 95% HPD interval are plotted in dark blue and dark red, respectively. Estimates greater or less than 1 based on the 50% HPD interval are plotted in light blue and light red. The estimates that could not be distinguished from  $\Phi = 1$  based on the 50% HPD are plotted in light grey. Scenarios III and IV were simulated under a model with one shift in the scale parameters, thus the x-axis shows the mean of the two true scale parameters.

TABLE 4. Frequency of erroneous ADE estimates under the four simulated scenarios

Scenario	Frequency of erroneously estimated shape parameters	
	95% HPD	50% HPD
I	0	0.051
II	0.006	0.07
III	0	0.041
IV	0.003	0.113

Note: Erroneous estimates are identified as simulations in which the shape parameter of the Weibull distribution is wrongly estimated to be greater than 1, when the true value was less than 1, and *vice versa*. We calculated whether the estimated shape was significantly different from 1 based on the 95% and 50% HPD. These results show that the method is able to qualitatively recover the correct model of evolution, even under moderate violations of its model assumptions.

with species age. The estimated scale parameter was 0.0021, resulting in a mean species longevity of 9.78 myr.

The substantial conflict between fossil-based and phylogenetic analyses may be explained by an erroneous estimation of the ADE parameters in either model as a result of strong violations of model assumptions. The analysis of simulated fossil and phylogenetic data sets with strongly varying speciation and extinction rates through time indicates that such model violations frequently result in artificial evidence for ADE (Fig. 5a). In fossil-based simulations, our method erroneously found evidence of ADE in 17.5% of the tests, based on the 95% HPD around the shape parameter of the Weibull distribution. Phylogenetic data resulted in a considerably higher rate of false positives (63%),

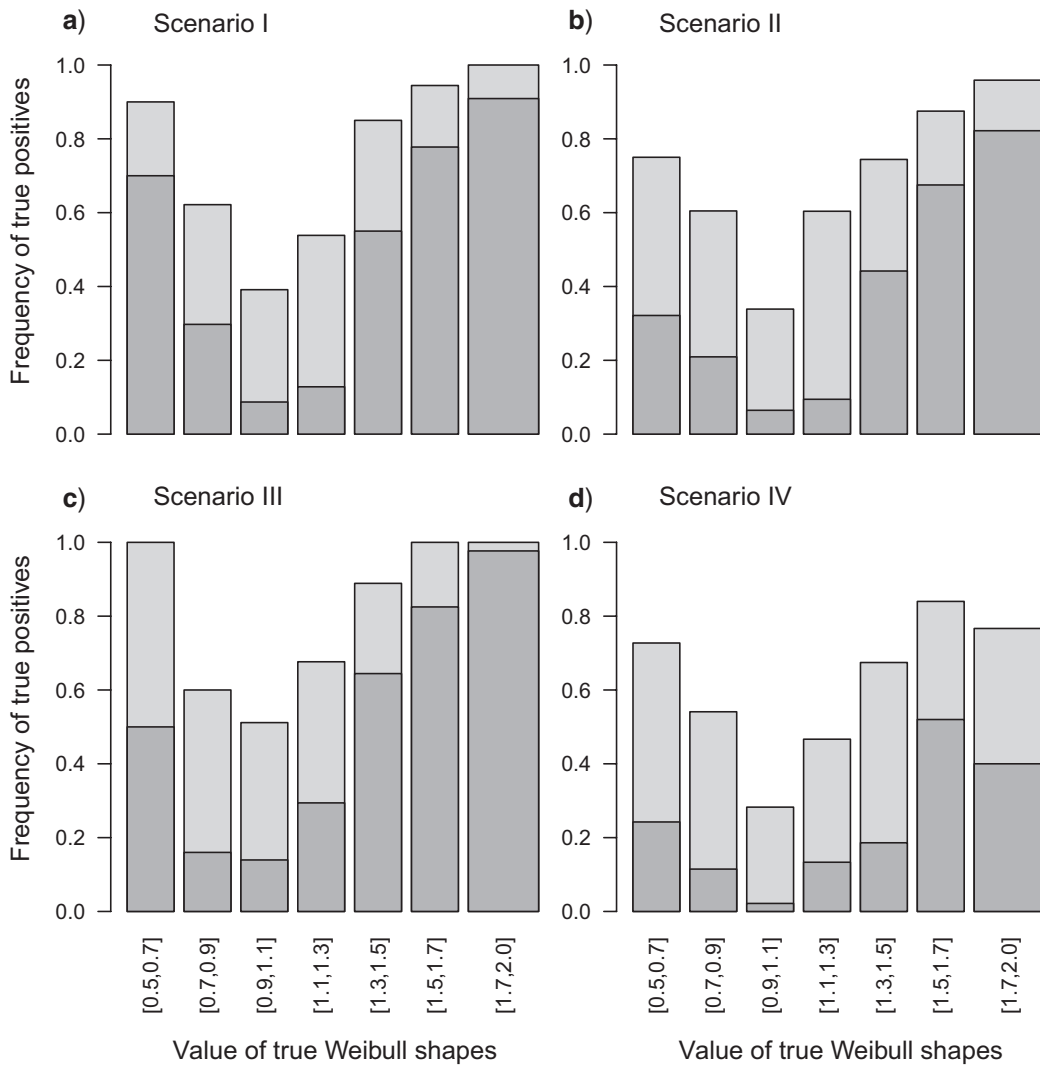


FIGURE 4. Power to correctly detect ADE. We considered as true positives simulations in which the shape parameter was correctly estimated to be greater or less than one based on the 50% HPD (light gray) and on the 95% HPD (dark gray) True shape values were grouped into seven categories and the frequency of true positives is shown within each category.

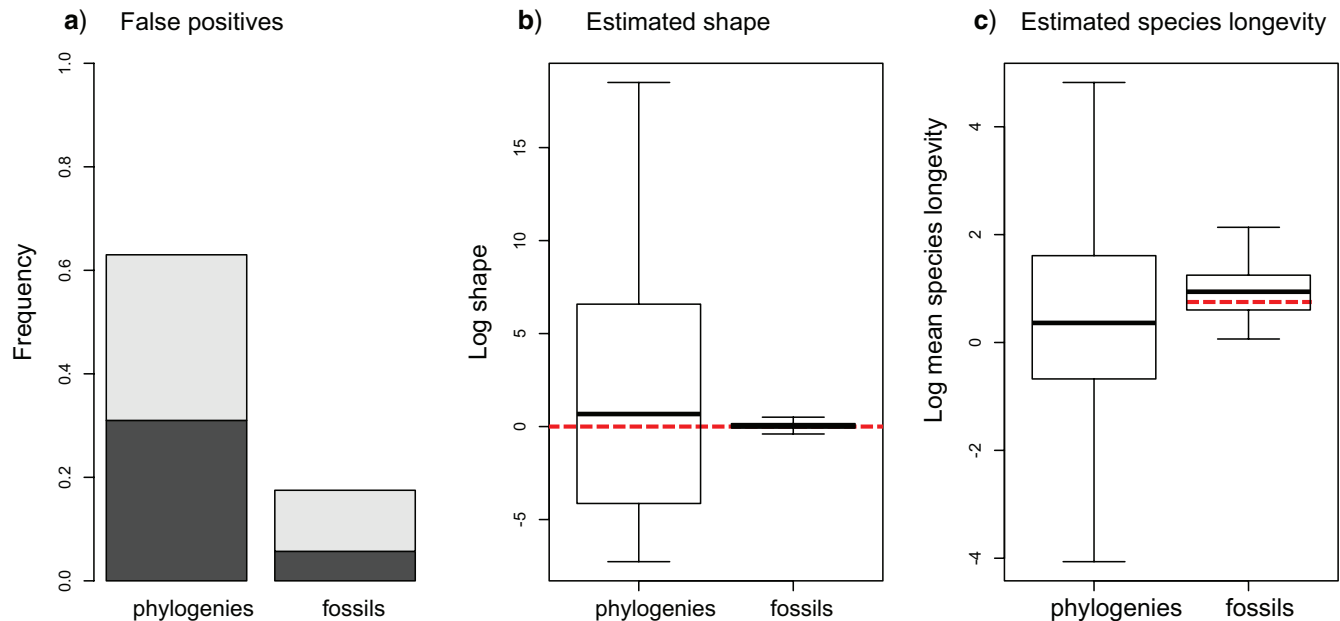


FIGURE 5. Robustness of fossil-based and phylogenetic ADE models to violations of assumptions. We simulated 300 phylogenies with 1000 tips and 300 fossil data sets under time-variable (but age-independent) speciation and extinction rates. Data sets were analyzed using a phylogenetic ADE model (Alexander et al. 2016) and our fossil-based ADE model. a) Proportion of false positives based on a likelihood ratio test for phylogenies and 95% HPD for fossil data. Light gray indicates the frequency of simulations, where the shape parameter was estimated to be significantly less than 1; dark gray indicates instances, where the shape parameter was estimated to be significantly greater than 1. b) Estimated shape parameters in log scale: the dashed line indicates the shape parameters under age-independent extinction models (shape = 1). Note that the phylogenetic method uses a gamma distribution, whereas the fossil-based method uses a similar Weibull distribution. c) Estimated log of the mean species longevity: dashed line represents the empirical mean longevity of extinct species in the fossil data sets. Boxes and whiskers indicate the 50% and 95% ranges, respectively; the median is indicated by the black horizontal bar.

based on likelihood ratio tests. Fossil-based parameter estimates appear to be closer to the true values and less prone to extreme values than phylogeny-based estimates (Fig. 5b, c). In the case of fossil data sets, we could compare the estimated mean species longevity with the empirical mean longevity of extinct simulated lineages, averaged over all simulations. The mean species longevity was estimated with high accuracy from the fossil data, despite the model violations and high frequency of false positives (Fig. 5c).

Since our simulations indicate that the correct estimation of ADE is sensitive to time-varying speciation and extinction rates, the contrasting results obtained for Carnivora using fossils and phylogenies are likely affected by the documented temporal heterogeneity of carnivore diversification (Pires et al. 2015; Silvestro et al. 2015a). However, the additional analyses based on the Neogene fossil record—a period during which speciation and extinction rates had previously been shown to be stable (Liow and Finarelli 2014)—were qualitatively similar to those based on the full data set. The estimated shape parameter of the Weibull distribution was 0.73 (95% HPD 0.61–0.84), that is, significantly less than 1. The estimated scale parameter was 1.73 (95% HPD 1.28–2.20), resulting in an estimated mean species longevity of 2.12 myr (95% HPD 1.78–2.45 myr).

## DISCUSSION

Assessing whether the age of a lineage affects the probability of its becoming extinct is crucial to understand the mechanisms underlying species survival and diversification and is at the center of fundamental evolutionary theories such as Van Valen's Red Queen hypothesis (Van Valen 1973, 1977; Pearson 1995; Liow et al. 2011). Most paleontological research aimed to prove or disprove ADE has studied either species in extremely well-preserved lineages such as planktonic foraminifera (Pearson 1992; Parker and Arnold 1997; Doran et al. 2006; Ezard et al. 2011) or at higher taxonomic units (genera, families) in clades with good fossil records such as Brachiopoda (Van Valen 1973; Raup 1975). Exceptionally preserved fossil record yields good taxon sampling and facilitates the estimation of species' (or higher taxa) longevities, thus allowing the direct fit of survivorship curves to the observed longevity of lineages. However, the application of these methods to the large majority of other clades with poorer species-level sampling is likely to lead to biased results (Foote and Raup 1996).

We have presented a new method to estimate extinction as a function of species age using the fossil record. Our model, implemented in a Bayesian framework, allows for a joint estimation of 1) the times of origination and extinction of sampled lineages, 2) the preservation rates, and 3) the parameters related

to ADE. The model explicitly incorporates different potential sources of uncertainty and biases in the fossil record. First of all, it extends the observed longevity of sampled lineages beyond the first and last appearances, estimating the times of origination and extinction based on the preservation process. This procedure is important to correct for the otherwise underestimated longevity of species (Liow and Stenseth 2007; Silvestro et al. 2014a) and to enable the analysis of singletons, which have an observed longevity equal to zero (Silvestro et al. 2014b). Secondly, our ADE model, unlike most models, is able to explicitly correct for the biases linked to unobserved species by incorporating the signature that missing lineages leave in the distribution of observed longevities as a function of the preservation rate (Fig. 1). This is crucial because preservation rates affect the probability of sampling short-lived lineages more severely than long-lived species. Temporal heterogeneity of the preservation rate is also integrated in the model, thus accounting for time-variable biases, a condition likely to be true in most empirical data sets (Foote and Miller 2007). Finally, our method is able to also use extant species to infer the parameters of the ADE model and, therefore, the resulting expected longevity. We show through simulations that the parameters of the model are accurately inferred under a range of parameter settings. Overall, the method has good power to correctly identify ADE when true and is able to reject it when extinction is age-independent.

#### *Confounding Effects and Potential Artifacts*

Time-variable extinction, that is, clade-wide changes in the extinction rate through time, is known to generate confounding effects, when attempting to infer age dependency from the longevity of lineages (Pearson 1992, 1995). In our simulations, we found that these effects are negligible or limited if the changes in extinction rates are comparatively small (up to 5-fold rate changes). This magnitude of rate variation is quite common in empirical data sets (Pires et al. 2015). However, bias becomes more substantial when speciation and extinction rates change strongly and repeatedly during the diversification history of a clade. Under such circumstances, the reliability of estimates decreases and the ADE model becomes more prone to errors. This problem can be potentially overcome by analyzing taxa originating in different time periods separately (Raup 1978; Pearson 1995) or, as we did in our empirical analysis, subsampling the data to only include lineages within a time window with overall stationary extinction rates. In future implementations, both types of rate variations could be incorporated in a single model, potentially building upon previously described methods (Pearson 1992; Doran et al. 2006; Ezard et al. 2011; Smits 2015).

Another potential source of bias is the presence of “pseudo-extinction” and “pseudo-speciation” events in the data, that is, the splitting of a lineage into different morphospecies linked to anagenesis that did not actually

involve a cladogenetic event or the termination of a lineage (Pearson 1992, 1995; Benton and Pearson 2001). Analyses of planktonic foraminifera based on morphospecies (Doran et al. 2006) and evolutionary species (the latter not affected by pseudo-extinction and pseudo-speciation) (Ezard et al. 2012) found similar evidence of ADE, where extinction rates decrease with lineage age. Ezard et al. (2012) investigated the effect of different species concepts on longevity under a Weibull model and showed they have little effect on the estimated shape parameters. With simulations, we showed that, when morphospecies (rather than evolutionary species *sensu* Ezard et al. 2012) are used as lineages in an ADE analysis, constant processes of pseudo-extinction and pseudo-speciation do not generate spurious evidence of ADE, rather they tend to reduce the estimated effects of age dependency (Supplementary Fig. S6 available on Dryad). Hence, under these settings, our approach is conservative with respect to finding age-dependency. Temporal variation of the rates of anagenesis are likely to make ADE estimates more prone to error, similarly to time-varying speciation and extinction rates. This could happen, for example, during adaptive radiations if changes in morphology and diversification rates are coupled, requiring special attention when analyzing clades that experience severe changes in rates. However, because birth–death models in PyRate do not differentiate between the extinction of a lineage and pseudo-extinction, time-varying rates of pseudo-extinction will be identified as heterogeneity in extinction rates, which in turn should call for caution when interpreting ADE results. Placing the lineages in a phylogenetic framework may help determine the validity of taxa and improve the estimation of their extinction rates (Ezard et al. 2011). In the absence of a reliable and complete phylogenetic hypothesis, the concerted influence of anagenesis, taxonomic practices, and limited amount of available data to determine and describe fossil taxa can impact the assignment of fossil occurrences into species (and other taxonomic units) and therefore affect inferences on the longevity of lineages.

Our model, like most other fossil-based and phylogenetic methods, relies on the definition of the taxonomic units implemented in a data set. Although phylogenetic methods are being developed to address species delimitation issues using molecular data (e.g. Jones et al. 2015; Rannala 2015), the identification of extinct taxa is usually strongly dependent on the expertise of paleontologists and systematists who describe and identify fossil specimens. Although providing new solutions to this issue goes beyond the purpose of this study, this potential artifact highlights the importance of taxonomically well-verified data in macroevolutionary inferences.

#### *Estimating ADE from Fossils and from Phylogenies of Extant Taxa*

Phylogenetic comparative methods and the estimation of speciation and extinction rates from phylogenies

of extant taxa have progressed dramatically in recent years, expanding their scope and revealing unsuspected properties of tree shapes and branching times (see [Stadler 2013b](#); [Morlon 2014](#), for recent reviews on the topic). [Alexander et al. \(2016\)](#) demonstrated that evidence of ADE can be detected from molecular phylogenies, despite the fact that extinction events are not directly observed on trees. Although their method constitutes a major achievement in phylogenetic comparative methods, our simulations show that the phylogenetic approach is very sensitive to model violations (Fig. 5). This is not surprising, given the general difficulties in estimating extinction dynamics from extant-taxon phylogenies ([Ricklefs 2007](#); [Quental and Marshall 2010](#); [Stadler 2013a](#)). ADE can be accurately estimated from phylogenies, provided that they are sufficiently large (hundreds of tips) and that the model assumptions are not violated, that is, rates are only dependent on age and are otherwise constant through time ([Alexander et al. 2016](#)).

However, the larger and older a phylogeny is, the lower the chances that the data comply with the model assumptions. To make ADE models more easily applicable to empirical phylogenies, we agree with [Alexander et al. \(2016\)](#) that further methodological development is desirable to relax these assumptions and provide more general and more realistic models of diversification. Although advances in phylogenetic methods are promising, the fossil record arguably remains the most direct and reliable evidence of extinction dynamics and lineages longevity ([Quental and Marshall 2010](#); [Condamine et al. 2013](#)). Despite the better performance of the fossil-based ADE model, similar limitations may exist for fossil-based analyses. In addition, several sources of bias remain unaccounted for in our model such as variation in extinction rates through time.

#### *Extinction in Mammalian Carnivores*

The fossil record of mammalian carnivores provides strong evidence for ADE, in which the recently originated species are much more likely to become extinct than older species. Under the estimated parameter values, the extinction rate for a lineage 0.1 myr after its origination is 0.91 (Equation 1), whereas for a lineage that has lived 1 myr, the extinction rate is considerably lower (0.49). Extinction rates decrease with increasing lineage age – it is reduced to 0.26 after 10 myr. The mean longevity of sampled taxa in the carnivore fossil data (calculated from the estimated times of origination and extinction) was 2.69 myr. This estimate falls within the range of 2.63–2.95 myr identified by [Prothero \(2014\)](#) who revised the taxonomy of the group and applied simple range-through-time differences between first and last appearances. However, the average species longevity inferred by our ADE model was more than 500,000 years shorter (2.02 or 2.12 myr depending on the analysis). We emphasize

that this estimate accounts for unobserved species, whose existence is not documented in the fossil record, but implied by the estimated preservation rate and the parameters of the Weibull distribution (Fig. 1). This result highlights the importance of incorporating fossilization and sampling biases when attempting to infer species longevity.

Importantly, phylogenetic and fossil analyses led to opposite conclusions regarding the effect of age dependence on extinction rates in carnivores, indicating that extinction probability decreases (fossil) or increases (phylogeny) with species age. Although there are examples of consistent results in macroevolutionary analyses of fossil data and phylogenies of extant taxa (e.g., [Morlon et al. 2011](#); [Cantalapiedra et al. 2015](#)), phylogenetic and fossil data have often produced discordant results ([Liow et al. 2010](#); [Quental and Marshall 2010](#); [Rabosky 2010](#), [Silvestro et al. 2014b](#)). Discordances between phylogenetic and fossil analyses could also be related to erroneous dating (of fossil occurrences or phylogenetic branching events) or incongruences in taxonomic units used in the analysis. In recent years, it has therefore become progressively clear that efforts should focus on understanding the sources of these discrepancies, which may be linked to different biases both in the data and in the methods and model assumptions ([Slater et al. 2012b](#); [Fritz et al. 2013](#); [Slater and Harmon 2013](#)).

In the case of ADE models, we have shown that both phylogenetic and fossil analyses are prone to frequent errors in the presence of violations of the model assumptions, namely time-variable speciation and extinction rates. Consequently, it is reasonable to interpret the observed incongruence between analyses as the result of biased inferences. In the case of fossil-based analyses, however, we were able to reduce the risk of biased results by restricting the analysis to species from the Neogene, where overall speciation and extinction rates for carnivores have been independently shown to be relatively constant over time ([Liow and Finarelli 2014](#); [Pires et al. 2015](#)). We therefore consider the ADE parameters estimated from the fossil record of Carnivora to be more reliable than those obtained from the phylogeny of extant taxa. Further research is needed to fully understand the sources of conflict between phylogenetic and fossil-based analyses, such as the comparability of the species concept used in the two different frameworks. Integrating paleontological and neontological data in a single analytical framework will likely render macroevolutionary inferences more realistic and robust ([Ronquist et al. 2012](#); [Slater et al. 2012a](#); [Heath et al. 2014](#)).

#### CONCLUSIONS

We developed a new method to estimate ADE in a Bayesian framework while explicitly correcting for the effects that preservation has on the observed longevity of lineages and on the probability of their sampling.

Because of its ability to overcome sampling biases, we expect our method to be applicable to a wider range of organisms and at lower taxonomic levels (species) than other available approaches. Hence, it may also allow for more reliable interpretations of what affects species longevity by explicitly incorporating the uncertainties of the fossil record compared to methods that read species' longevities literally (e.g., Smits 2015).

Older fossil-based studies, such as the foundational paper of Van Valen (1973), tended to use genus longevities in their analyses, whereas more recent studies have used the level of species. Of particular interest is the potential effect that taxonomic resolution might have on our inferences regarding the generality of ADE models. Ezard et al. (2016) and Crampton et al. (2016) have shown that the choice of taxonomic level in foraminifera and graptolites, respectively, affects extinction rate inferences, with evidence of age dependence at the species level but not at the genus level. This suggests that Van Valen's law of constant extinction may be taxonomy dependent. In fact, Raup and Stanley (1978) have hypothesized that ADE might not act simultaneously at the species and genus level. The method presented here explicitly corrects for biases linked to poor fossil quality and may therefore open up an avenue for broader empirical analyses at different taxonomic levels.

#### SUPPLEMENTARY MATERIAL

Data available from the Dryad Digital Repository: <http://dx.doi.org/10.5061/dryad.r5f70>.

#### ACKNOWLEDGMENTS

We thank Rachel Warnock, Tanja Stadler, Xavier Meyer, and Nicolas Salamin for Alexandra Gavryushkina, discussions and input. We also thank the editors, Michael Alfaro and Thomas Near, Thomas Ezard, and one anonymous reviewer for valuable comments on this manuscript.

#### FUNDING

This work was supported by the Swedish Research Council (B0569601 to A.A. and 2015-04748 to D.S.); the São Paulo Research Foundation (FAPESP) (grant no. 2012/04072-3 to T.B.Q.); the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013, ERC Grant Agreement 331024 to A.A.); and the Swedish Foundation for Strategic Research and a Wallenberg Academy Fellowship to A.A.

#### REFERENCES

Alexander H.K., Lambert A., Stadler T. 2016. Quantifying age-dependent extinction from species phylogenies. *Syst. Biol.* 65: 35–50.

- Bacon C.D., Silvestro D., Jaramillo C., Smith B.T., Chakrabarty P., Antonelli A. 2015. Biological evidence supports an early and complex emergence of the Isthmus of Panama. *Proc. Natl. Acad. Sci. U.S.A.* 112: 6110–6115.
- Benton M.J. 2016. Origins of biodiversity. *PLoS Biol.* 14: e2000724.
- Benton M.J., Pearson P.N. 2001. Speciation in the fossil record. *Trends Ecol. Evol.* 16: 405–411.
- Cantalapiedra J.L., Hernández Fernández M., Azanza B., Morales J. 2015. Congruent phylogenetic and fossil signatures of mammalian diversification dynamics driven by tertiary abiotic change. *Evolution* 69: 2941–2953.
- Condamine F.L., Rolland J., Morlon H. 2013. Macroevolutionary perspectives to environmental change. *Ecol. Lett.* 16: 72–85.
- Crampton J.S., Cooper R.A., Sadler P.M., Foote M. 2016. Greenhouse–icehouse transition in the Late Ordovician marks a step change in extinction regime in the marine plankton. *PNAS* 113: 1498–1503.
- Doran N.A., Arnold A.J., Parker W.C., Huffer F.W. 2006. Is extinction age dependent? *Palaios* 21: 571–579.
- Ezard T.H., Quental T.B., Benton M.J. 2016. The challenges to inferring the regulators of biodiversity in deep time. *Phil. Trans. R. Soc. B* 371: 20150216.
- Ezard T.H.G., Aze T., Pearson P.N., Purvis A. 2011. Interplay between changing climate and species' ecology drives macroevolutionary dynamics. *Science* 332: 349–351.
- Ezard T.H.G., Pearson P.N., Aze T., Purvis A. 2012. The meaning of birth and death (in macroevolutionary birth–death models). *Biol. Lett.* 8: 139–142.
- Finnegan S., Payne J.L., Wang S.C. 2008. The red queen revisited: reevaluating the age selectivity of phanerozoic marine genus extinctions. *Paleobiology* 34: 318–341.
- Foote M. 2000. Origination and extinction components of taxonomic diversity: general problems. *Paleobiology* 26: 74–102.
- Foote M. 2003. Origination and extinction through the phanerozoic: a new approach. *J. Geol.* 111: 125–148.
- Foote M., Miller A.I. 2007. *Principles of paleontology*. New York (NY): W. H. Freeman.
- Foote M., Raup D.M. 1996. Fossil preservation and the stratigraphic ranges of taxa. *Paleobiology* 22: 121–140.
- Fritz S., Schnitzler J., Eronen J.T., Hof C., Böhning-Gaese K., Graham C. 2013. Diversity in time and space: wanted dead and alive. *Trends Ecol. Evol.* 28: 509–516.
- Hagen O., Hartmann K., Steel M., Stadler T. 2015. Age-dependent speciation can explain the shape of empirical phylogenies. *Syst. Biol.* 64: 432–440.
- Hagen O., Stadler T. 2017. Treessim: simulating phylogenetic trees under general bellman harris models with lineage-specific shifts of speciation and extinction in R. *Methods Ecol. Evol.* doi:10.1111/2041-210X.12917
- Heath T.A., Huelsenbeck J.P., Stadler T. 2014. The fossilized birth–death process for coherent calibration of divergence-time estimates. *PNAS* 111: E2957–E2966.
- Johnson M.T.J., FitzJohn R.G., Smith S.D., Rausher M.D., Otto S.P. 2011. Loss of sexual recombination and segregation is associated with increased diversification in evening primroses. *Evolution* 65: 3230–3240.
- Jones G., Aydin Z., Oxelman B. 2015. Dissect: an assignment-free bayesian discovery method for species delimitation under the multispecies coalescent. *Bioinformatics* 31: 991–998.
- Jones G.R. 2011. Tree models for macroevolution and phylogenetic analysis. *Syst. Biol.* 60: 735–746.
- Lambert A. 2010. The contour of splitting trees is a Lévy process. *Ann. Probab.* 38: 348–395.
- Lambert A., Stadler T. 2013. Birth–death models and coalescent point processes: the shape and probability of reconstructed phylogenies. *Theor. Popul. Biol.* 90: 113–128.
- Liow L.H., Finarelli J.A. 2014. A dynamic global equilibrium in carnivoran diversification over 20 million years. *Proc. R. Soc. Lond. B* 281: 20132312.
- Liow L.H., Nichols J.D. 2010. Estimating rates and probabilities of origination and extinction using taxonomic occurrence data: capture–recapture approaches. In *Short courses in paleontology*:



- quantitative paleobiology (eds Hunt G, Alroy J), pp. 81–94. New Haven, CT: Yale University Printing and Publishing.
- Liow L.H., Quental T.B., Marshall C.R. 2010. When can decreasing diversification rates be detected with molecular phylogenies and the fossil record? *Syst. Biol.* 59: 646–659.
- Liow L.H., Reitan T., Harnik P.G. 2015. Ecological interactions on macroevolutionary time scales: Clams and brachiopods are more than ships that pass in the night. *Ecol. Lett.* 18: 1030–1039.
- Liow L.H., Stenseth N.C. 2007. The rise and fall of species: implications for macroevolutionary and macroecological studies. *Proc. R. Soc. Lond. B Biol. Sci.* 274: 2745–2752.
- Liow L.H., Van Valen L., Stenseth N.C. 2011. Red queen: from populations to taxa and communities. *Trends Ecol. Evol.* 26: 349–358.
- May M.R., Hohna S., Moore B.R. 2016. A Bayesian approach for detecting the impact of mass-extinction events on molecular phylogenies when rates of lineage diversification may vary. *Methods Ecol. Evol.* 7: 947–959.
- Morlon H. 2014. Phylogenetic approaches for studying diversification. *Ecol. Lett.* 17: 508–525.
- Morlon H., Parsons T.L., Plotkin J.B. 2011. Reconciling molecular phylogenies with the fossil record. *PNAS* 108: 16327–16332.
- Muller H.J. 1964. The relation of recombination to mutational advance. *Mutat Res Fundam Mol Mech Mutagen* 1: 2–9.
- Nee S. 2006. Birth-death models in macroevolution. *Annu. Rev. Ecol. Syst.* 37: 1–17.
- Nyakatura K., Bininda-Emonds O.R.P. 2012. Updating the evolutionary history of Carnivora (Mammalia): a new species-level supertree complete with divergence time estimates. *BMC Biol* 10: 12.
- Parker W.C., Arnold A.J. 1997. Species survivorship in the cenozoic planktonic foraminifera: a test of exponential and Weibull models. *Palaios* 12: 3–11.
- Pearson P.N. 1992. Survivorship analysis of fossil taxa when real-time extinction rates vary: the Paleogene planktonic foraminifera. *Paleobiology* 18: 115–131.
- Pearson P.N. 1995. Investigating age-dependency of species extinction rates using dynamic survivorship analysis. *Hist. Biol.* 10: 119–136.
- Pigot A.L., Owens I.P.F., Orme C.D.L. 2012. Speciation and extinction drive the appearance of directional range size evolution in phylogenies and the fossil record. *PLoS Biol.* 10: e1001260.
- Pires M.M., Silvestro D., Quental T.B. 2015. Continental faunal exchange and the asymmetrical radiation of carnivores. *Proc. R. Soc. Biol. Sci. Ser. B* 282: 20151952.
- Prothero D.R. 2014. Species longevity in North American fossil mammals. *Integr. Zool.* 9: 383–393.
- Quental T.B., Marshall C.R. 2010. Diversity dynamics: molecular phylogenies need the fossil record. *Trends Ecol. Evol.* 25: 434–441.
- Rabosky D.L. 2006. Likelihood methods for detecting temporal shifts in diversification rates. *Evolution* 60: 1152–1164.
- Rabosky D.L. 2010. Extinction rates should not be estimated from molecular phylogenies. *Evolution* 64: 1816–1824.
- Rambaut A., Suchard M., Xie D., Drummond A. 2014. Tracer v1.6. Available from: URL <http://beast.bio.ed.ac.uk/Tracer>.
- Rannala B. 2015. The art and science of species delimitation. *Curr. Zool.* 61: 846–853.
- Raup D.M. 1975. Taxonomic survivorship curves and Van Valen's Law. *Paleobiology* 1: 82–96.
- Raup D.M. 1978. Cohort analysis of generic survivorship. *Paleobiology* 4: 1–15.
- Raup D.M., Stanley S.M. 1978. Principles of paleontology. New York (NY): W. H. Freeman.
- Ricklefs R.E. 2007. Estimating diversification rates from phylogenetic information. *Trends Ecol. Evol.* 22: 601–610.
- Ronquist F, van der Mark P, Huelsenbeck J.P. 2009. Bayesian phylogenetic analysis using MrBayes. In: Lemey P, Salemi M, Vandamme A-M, eds. The phylogenetic handbook. New York: Cambridge University Press, 210–266.
- Ronquist F., Klopfstein S., Vilhelmsen L., Schulmeister S., Murray D.L., Rasnitsyn A.P. 2012. A total-evidence approach to dating with fossils, applied to the early radiation of the hymenoptera. *Syst. Biol.* 61: 973–999.
- Silvestro D., Antonelli A., Salamin N., Quental T.B. 2015a. The role of clade competition in the diversification of North American canids. *PNAS* 112: 8684–8689.
- Silvestro D., Cascales-Miñana B., Bacon C.D., Antonelli A. 2015b. Revisiting the origin and diversification of vascular plants through a comprehensive bayesian analysis of the fossil record. *New Phytol.* 207: 425–436.
- Silvestro D., Salamin N., Schnitzler J. 2014a. Pyrate: a new program to estimate speciation and extinction rates from incomplete fossil data. *Methods Ecol. Evol.* 5: 1126–1131.
- Silvestro D., Schnitzler J., Liow L.H., Antonelli A., Salamin N. 2014b. Bayesian estimation of speciation and extinction from incomplete fossil occurrence data. *Syst. Biol.* 63: 349–367.
- Silvestro D., Zizka A., Bacon C.D., Cascales-Miñana B., Salamin N., Antonelli A. 2016. Fossil biogeography: a new model to infer dispersal, extinction and sampling from palaeontological data. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 371: 20150225.
- Slater G.J., Harmon L.J. 2013. Unifying fossils and phylogenies for comparative analyses of diversification and trait evolution. *Methods Ecol. Evol.* 4: 699–702.
- Slater G.J., Harmon L.J., Alfaro M.E. 2012a. Integrating fossils with molecular phylogenies improves inference of trait evolution. *Evolution* 66: 3931–3944.
- Slater G.J., Harmon L.J., Wegmann D., Joyce P., Revell L.J., Alfaro M.E. 2012b. Fitting models of continuous trait evolution to incompletely sampled comparative data using approximate bayesian computation. *Evolution* 66: 752–762.
- Smits P.D. 2015. Expected time-invariant effects of biological traits on mammal species duration. *PNAS* 112: 13015–13020.
- Stadler T. 2011a. Mammalian phylogeny reveals recent diversification rate shifts. *Proc. Natl Acad. Sci. U.S.A.* 108: 6187–6192.
- Stadler T. 2011b. Simulating trees with a fixed number of extant species. *Syst. Biol.* 60: 676–684.
- Stadler T. 2013a. How can we improve accuracy of macroevolutionary rate estimates? *Syst. Biol.* 62: 321–329.
- Stadler T. 2013b. Recovering speciation and extinction dynamics based on phylogenies. *J. Evol. Biol.* 26: 1203–1219.
- Van Valen L. 1973. A new evolutionary law. *Evol. Theory* 1: 1–30.
- Van Valen L. 1977. The Red Queen. *Am. Nat.* 111: 809–810.
- Weibull W. 1939. A statistical theory of the strength of materials. Stockholm (Sweden): Generalstabens Litografiska Anstalts Förlag.
- Wiltshire J., Huffer F.W., Parker W.C. 2014. A general class of test statistics for Van Valen's Red Queen hypothesis. *J. Appl. Stat.* 41: 2028–2043.