Distributions of extinction times from fossil ages and tree topologies: the example of some mid-Permian synapsid extinctions

Gilles Didier[1] and Michel Laurin[2]

[1] IMAG, Univ Montpellier, CNRS, Montpellier, France

[2] CR2P ("Centre de Paléontologie – Paris"; UMR 7207), CNRS/MNHN/SU, Muséum National d'Histoire Naturelle, Paris, France

September 16, 2021

Abstract

Given a phylogenetic tree that includes only extinct, or a mix of extinct and extant taxa, where at least some fossil data are available, we present a method to compute the distribution of the extinction time of a given set of taxa under the Fossilized-Birth-Death model.

Our approach differs from the previous ones in that it takes into account (i) the possibility that the taxa or the clade considered may diversify before going extinct and (ii) the whole phylogenetic tree to estimate extinction times, whilst previous methods do not consider the diversification process and deal with each branch independently. Because of this, our method can estimate extinction times of lineages represented by a single fossil, provided that they belong to a clade that includes other fossil occurrences.

We assess and compare our new approach with a standard previous one using simulated data. Results show that our method provides more accurate confidence intervals.

This new approach is applied to the study of the extinction time of three Permo-Carboniferous synapsid taxa (Ophiacodontidae, Edaphosauridae, and Sphenacodontidae) that are thought to have disappeared toward the end of the Cisuralian (early Permian), or possibly shortly thereafter. The timing of extinctions of these three taxa and of their component lineages supports the idea that the biological crisis in the late Kungurian/early Roadian consisted of a progressive decline in biodiversity throughout the Kungurian.

Introduction

Reconstructing the history of the diversification of life on Earth has long been one of the main goals of evolutionary biology. In this great enterprise, the fossil record plays a central role because it gives direct evidence (even if fragmentary) of the biodiversity at various times (Carroll, 1988). It even documents spectacular changes in the rates of cladogenesis (evolutionary divergence of a lineage that splits into two lineages, a process that we here equate with speciation), anagenesis, and extinction, which occurred more or less regularly in the history of life. These were often caused by environmental changes, some of which may have resulted from intense volcanism (Wignall *et al.*, 2009), impacts of large meteorites (Basu *et al.*, 2003; Tabor *et al.*, 2020), a combination of both (Arens and West, 2008), or simply transgressions, regressions (Hallam, 1989), or glaciations, among others. Such changes are associated with evolutionary radiations (Ronquist *et al.*, 2012; Slater, 2015; Gavryushkina *et al.*, 2016; Brocklehurst, 2017; Ascarrunz *et al.*, 2019) that occur when the diversification of a taxon accelerates significantly, and mass extinction events (Axelrod and Bailey, 1968; Lewin, 1983; Raup and Sepkoski, 1984; Stanley, 1988; MacLeod, 1996; Benton, 2003; Ward *et al.*, 2005; Retallack *et al.*, 2006; Wignall *et al.*, 2009; Bond *et al.*, 2010; Ruta *et al.*, 2011; Sidor *et al.*, 2013; Lucas, 2017; Brocklehurst, 2018), during which the extinction rate of many taxa increases greatly, but typically for a short time.

So far, most studies of these phenomena that emphasized the fossil record have used the taxic approach, which consists of counting the number of taxa of a given rank (most frequently, families or

genera; more rarely, species) in various time bins and estimating fluctuations in extinction and origination rates (Raup and Sepkoski, 1984; Benton, 1985, 1989; Alroy, 1996; Day *et al.*, 2015a; Brocklehurst, 2018). Several limitations are inherent to this approach.

First, some early studies relied on databases that included many paraphyletic or even polyphyletic taxa, and thus confused pseudoextinction with genuine extinction (Patterson and Smith, 1987), even though the extinction of a paraphyletic taxon often coincides with the extinction of several smaller clades therein. Indeed, this was identified by Benton (1989) as one of the main aspects that could be improved in subsequent studies, and subsequent developments proved him right (Uhen, 1996; Fara, 2004; Marjanović and Laurin, 2008). Some recent analyses using the taxic approach even include a phylogenetic correction to these biodiversity counts by accounting for ghost lineages (Ruta *et al.*, 2011; Martin *et al.*, 2014; Jouve *et al.*, 2017). Progress has also been made on how to integrate events from various sections into an optimal sequence (Sadler, 2004), and this has been applied to one of the Permian mass extinction events (Day *et al.*, 2015a).

Second, counts of taxa at a given Linnaean nomenclatural level (except for species, if these are conceptualized as an evolutionary lineage) are poor measures of biodiversity (Bertrand *et al.*, 2006; Laurin, 2010), because taxa of a given level (i.e., Linnaean rank) share no objective properties (Ereshefsky, 2002) and may include one to many lineages. For this reason, better ways to quantify biodiversity were developed (Faith, 1992). We do not imply that lineage-level analyses provide a complete picture of evolution of biodiversity because other aspects are relevant, such as disparity. Some key evolutionary events, such as the Cambrian explosion, may be remarkable because of the increase in disparity rather than in speciation rate (which is poorly constrained in the Ediacarian). To an extent, counts of taxa of higher ranks may capture this, but in a indirect and imprecise way, and the evolution of disparity is better tackled by quantitative measures designed specifically to capture this (see, e.g., Wilson *et al.*, 2013).

A third problem of the classical taxic approach is that the known stratigraphic ranges of taxa typically underestimate their true stratigraphic range (real age of appearance and extinction), a problem that is likely to be especially acute for taxa with a poor fossil record (Strauss and Sadler, 1989). Most recent analyses using the taxic approach attempt to compensate indirectly for the incompleteness of the fossil record (see, e.g., Foote, 2003; Lu *et al.*, 2006), but this does not yield a clear idea about the timing of exinction of individual lineages.

Fourth, counting taxa in time bins can create two types of artefacts. First, if the time bins are relatively long (like geological periods or stages), the resulting counts may give the impression that origination or extinction events are concentrated at the limits between two consecutive bins, whereas in fact, the diversity of a taxon may have changed more or less gradually throughout the time bin (Day *et al.*, 2015a; Lucas, 2017). Some methods have been devised to minimize this problem (Foote, 2003), but when more detailed stratigraphic data are available, other methods may be more appropriate to better assess whether the changes are abrupt or gradual. However, this raises another problem: for taxa that have a sparse fossil record, simple sampling effects may give the false impression that extinctions have been gradual (which is the second type of artefact evoked above). This is called the "Signor-Lipps effect" because of the landmark study by Signor *et al.* (1982), even though Shaw (1964) described it earlier, according to MacLeod (1996). Again, some taxic studies have tackled this problem to an extent (Lu *et al.*, 2006), but these methods yield data on taxonomic global turnover rates in a given time interval, rather than a fine-scale view of the timing of extinction of individual taxa.

To establish a better understanding of the dynamics of fluctuations in biodiversity over time, it is thus useful to assess as accurately as possible the stratigraphic ranges of taxa. Early developments in this field tackled both ends (origination and extinction) of the stratigraphic ranges of taxa (Strauss and Sadler, 1989; Marshall, 1990, 1997; Wagner, 2000). Most recent methodological developments have addressed the problem of taxon origination by inferring how much of the earliest phase of each taxon's history remains hidden from the known fossil record, which may be useful to date the nodes of the Tree of Life (Tavaré *et al.*, 2002; Marshall, 2008; Laurin, 2012; Warnock *et al.*, 2012; Sterli *et al.*, 2013; Warnock *et al.*, 2015; Didier and Laurin, 2020). However, determining when taxa became extinct is also interesting, especially to better understand past biological crises. Mass extinction events have been increasingly studied in the last decades, especially for the end-Permian event (e.g., Benton, 2003; Ward *et al.*, 2005; Retallack *et al.*, 2006; Wignall *et al.*, 2009; Bond *et al.*, 2010; Ruta *et al.*, 2011; Sidor *et al.*, 2013; Lucas, 2017; Brocklehurst, 2018), a trend that is partly fueled by the rising concern about the current anthropogenic biodiversity crisis (Wake and Vredenburg, 2008; Barnosky *et al.*, 2011; Ceballos *et al.*, 2015, 2017). Thus, time is ripe to return to the question of timing of extinction of taxa.

Most of the approaches that addressed this question were derived from the seminal work of Strauss and Sadler (1989), which can provide confidence intervals for the origination and extinction time of a taxon when its fossilization potential is constant in time. Later, Marshall (1994, 1997) and Marjanović and Laurin (2008) extended this work to the case where the fossilization potential varies through time. In the same way, Silvestro *et al.* (2014a,b) considered a model where the fossilization recovery rate follows a PERT distribution (a generalized form of the beta distribution) between the origin and the end of a lineage, which is used in a Bayesian framework with priors defined from a birth-death process, to estimate the speciation and the extinction times. Among the approaches which are not derived from that of Strauss and Sadler (1989), let us mention that of Bradshaw *et al.* (2012), which is based on the method of McInerny *et al.* (2006) and that of Alroy (2014). We refer to Laurin (2012), Wang and Marshall (2016) and Marshall (2019) for recent reviews on this topic.

Studies of mass extinction events focus on patterns affecting taxa of various sizes, some of which include many lineages. For such studies, the methodology of Strauss and Sadler (1989) and of previous appraoches might not be appropriate because they consider each taxon as if it were composed of a single lineage that does not diversify. Thus, the derivative approaches do not take into account the possibility that the considered taxon (whether it is composed of one or more lineages) may have given birth to one or more lineages that left no fossil record before going extinct, possibly well after all lineages documented in the fossil record (Fig. 1). Neglecting this possibility could be justified in the case where the diversification rates are low with regard to the fossil recovery rate. Unfortunately, our previous studies suggest the opposite situation in the datasets considered in Didier *et al.* (2017) and Didier and Laurin (2020). It follows that one could expect the extinction times (and the stratigraphic range extensions) provided by Strauss and Sadler (1989) to be inaccurate in some cases. This problem may be minor when estimating the stratigraphic range of a single nominal species, but it is probably more severe when estimating the extinction time of a clade known to have included several species, as in the case of the Permo-Carboniferous taxa (Ophiacodontidae, Edaphosauridae and Sphenacodontidae) studied below. Note that taking into account the diversification process to assess the time of extinction is much more important for a clade that became extinct long ago (i.e., tens of thousands of years ago or more) than for lineages that became extinct in historical times and for which sighting records are available (Rivadeneira *et al.*, 2009). The timescales involved in the latter case ensure that no speciation (cladogenesis) event may occur between the last observation and the extinction of the considered lineage.

To better estimate extinction time by considering additional lineages that may have left no fossil record, the fossilized birth-death model (FBD model) could be used. The FBD model assumes that fossil finds follow a Poisson process, which is also assumed by Strauss and Sadler (1989), but it also models the diversification of taxa as a birth-death process. Given that the parameters that characterize the FBD model include an extinction rate, it should be possible to use this process to estimate the probability distribution of extinction times. So far, the FBD model has been used to date cladogenetic events (Stadler and Yang, 2013; Heath *et al.*, 2014; Didier and Laurin, 2020) but usually not extinction, with the exception of Brocklehurst (2020), who used tip dating with the FBD to assess Olson's extinction and to reject the idea of Olson's gap. Evaluating extinction times through the FBD would be very useful to determine the extent to which the Signor-Lipps effect has biased our perspective on mass extinction events. It could also be useful to reassess the reliability and stratigraphic significance of some taxa as index fossils, at least those with a relatively sparse fossil record for which reliable phylogenies exist; such cases are presumably fairly rare in the marine realm, but may be more common in continental biochronology (Steyer, 2000; Day *et al.*, 2013, 2015b). Indeed, stratigraphic correlations of continental strata, at least when relying on vertebrate fossils, often use higher-ranking taxa (nominal genera or families), especially when strata located on different continents are assessed (Rubidge, 2005; Lucas, 2018; Lucas and Shen, 2018).

Another concern with the approach of Strauss and Sadler (1989) is that it requires several fossils of a taxon to provide a confidence interval that bounds the corresponding extinction time, which makes it unsuitable for lineages or small clades with a low fossilization rate. Moreover, because it is computed independently on each extinct taxon (without consideration of its close relatives and the tree structure of this set of lineages), the level of precision provided by the method developed by Strauss and Sadler (1989) depends on the number of fossils present on each terminal branch. This point can be a major issue for datasets where the fossil recovery is low (see Simulation Study below). This limitation does not apply to the method we propose, in which the extinction-time distribution of taxa with a single fossil on their terminal branch can be determined, if this branch belongs to a clade with a sufficient number

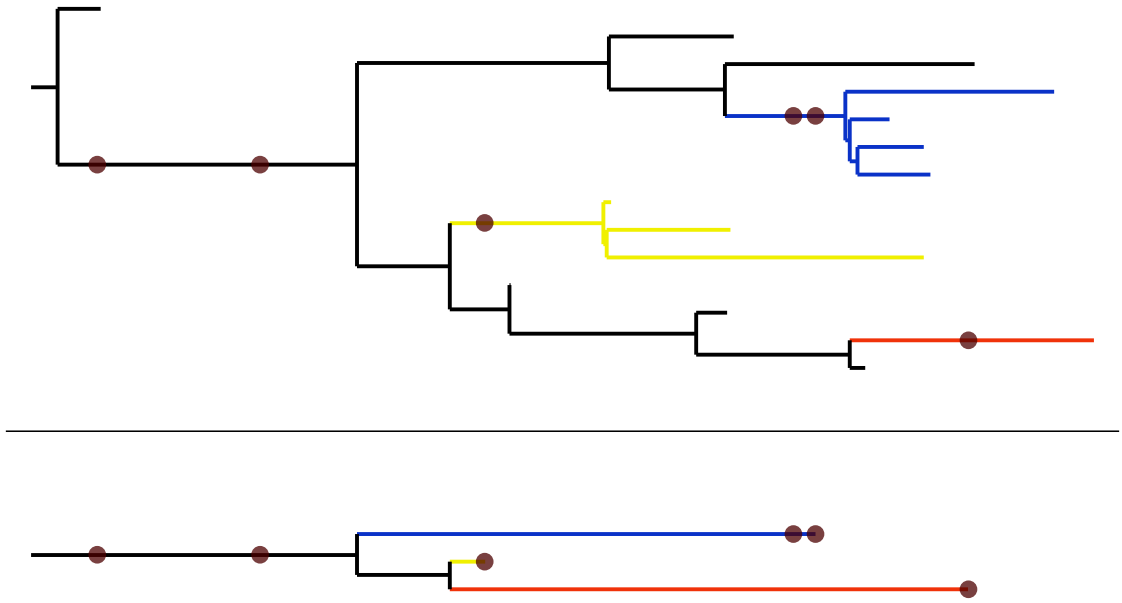

Figure 1: A simulated extinct clade with sampled fossils represented by brown dots. Top: The clade's complete evolutionary history. Bottom: The portion of the clade's history observable from the known fossil record. Note that the 'blue' and 'green' taxa diversify before going extinct, but that these diversification events are not recorded in the known fossil record.

of fossils. This results largely from the fact that all the data in a given dataset (which must represent a clade that may be truncated at any given time in the past) are used to assess the FBD parameters and hence, are considered in the computations of extinction time densities of all its branches.

Below, we extend the FBD model to estimate extinction times of taxa that may consist of one to many lineages. Specifically, given a dataset that consists of a phylogenetic tree of just extinct, or extinct and extant taxa, where at least some fossil data are available but without divergence times, we compute the probability that a given set of taxa (known to be extinct at the present time) goes extinct before a time t. The computation of this probability density is a direct extension of the method provided in Didier and Laurin (2020). We also provide an explicit formula for the probability that a given subset of extinct taxa (typically a clade) goes extinct before another one under the FBD model (Appendix B).

Our approach differs conceptually from that of Strauss and Sadler (1989) and subsequent works in the sense that, in our method, the extinction date is the time of a particular event whose distribution can be computed from the parameters of the model. By contrast, in the approach of Strauss and Sadler (1989), the extinction time is a parameter of the model, which has to be estimated or bounded in a confidence interval, and requires a Bayesian framework to obtain posterior distributions.

We adapted the Monte Carlo Markov Chain (MCMC) importance sampling procedure devised by Didier and Laurin (2020) in order to deal with the common case where fossil ages are provided as stratigraphic time intervals and to integrate distributions over the parameters of the model. Our approach allows incorporating phylogenetic uncertainty by estimating the extinction times over a set of trees.

The method presented below, like our previous works in this field, assumes a homogeneous fossilization rate through time. This assumption about the quality of the fossil record is made by most methods based on the FBD model. In practice, this requirement is never met, but mild violations of this assumption, like random fluctuations through time, should affect the reliability of the method less than pervasive trends over the considered time interval. Note that this hypothesis would not be reasonable if we had adopted what Lucas (2017) called the "best sections" analysis, which focuses on a region where a given taxon has a rich fossil record for a given period. Given that the FBD models diversification of clades, and that the clades that we analyze had a cosmopolitan distribution, it would be inappropriate to restrict our analysis to a single fossiliferous sedimentary basin. Note also that this assumption is analogous with the molecular

clock, which initially was assumed, for computation purposes, to be strict and universal (Zuckerkandl and Pauling, 1965), before local or relaxed clock methods (Cooper and Penny, 1997; Sanderson, 2002; Drummond *et al.*, 2006) were developed to account for rate variations that had been suspected to occur from the very beginning. Likewise, the FBD is probably amenable to such developments, but these are beyond the scope of this study

Our approach is first assessed and compared to that of Strauss and Sadler (1989) on simulated datasets. It is then applied to study the extinction of three Permo-Carboniferous taxa: Ophiacodontidae, Edaphosauridae and Sphenacodontidae. Two additional studies, which are not directly related to the extinction time question, are presented in appendices:

- We exploited the simulated data to assess the relevance of the posterior distributions of speciation, extinction and fossilization rates from phylogenetic trees with fossils, which can be obtained from our approach (Appendix D).
- The dataset used here incorporates a significantly larger amount of fossil data from the Paleobiology Database than that of Didier and Laurin (2020) (though it includes only a subset of taxa of this previous dataset). We applied the approach of Didier and Laurin (2020) to compute the divergence time densities and the posterior densities of the speciation, extinction and fossilization rates of this new dataset (Appendix E).

The computation of the extinction time distribution and of its confidence upper bound at a given threshold was implemented as a computer program and as a R package, both available at `https://github.com/gilles-didier/DateFBD`.

Empirical example: Permian extinction of Ophiacodontidae, Edaphosauridae and Sphenacodontidae

We illustrate our method with an empirical example from the rich fossil record of Permian synapsids. Synapsida originated in the Carboniferous and experienced a few evolutionary radiations, the first one of which, in the Late Carboniferous and early Permian, gave rise to taxa that have long been known as "pelycosaurs" (Romer and Price, 1940; Reisz, 1986; Benson, 2012), but which will be called here "Permo-Carboniferous synapsids". Among these taxa arose the stem lineage of therapsids, probably in the Late Carboniferous (Sidor, 2001; Amson and Laurin, 2011; Spindler, 2014; Angielczyk and Kammerer, 2018). Therapsids become increasingly common in the Roadian (early middle Permian) fossil record (Reisz and Laurin, 2002; Abdala *et al.*, 2008), and dominated several ecological niches from the Wordian (mid-Guadalupian) to the end of the Permian (Smith *et al.*, 2012). All other synapsid clades appear to have became extinct before the end of the Guadalupian (Modesto *et al.*, 2011). Therapsida experienced several evolutionary radiations, including one that gave rise to mammals, in the Triassic or in the Jurassic (King and Beck, 2020).

Up to four mass extinction events have been recognized in the Permian fossil record of synapsids (Lucas, 2017), and a brief review of these is relevant to understand the context of the present study and to justify the taxonomic sample. The first may have occurred at the Artinskian/Kungurian boundary (about 282 Ma), or possibly at the Sakmarian/Artinskian boundary (about 290.1 Ma), or it may be a long decline that occurred throughout the Sakmarian and Artinskian (Benton, 1985, 1989; Brocklehurst *et al.*, 2013). This extinction was suggested by early studies (Olson and Vaughn, 1970; Olson, 1982), and Brocklehurst *et al.* (2013) mentioned it too, as a late Sakmarian extinction. Benton (1989) included it in his list of major mass extinctions (along with the P/Tr and K/Pg events), and stated that Ophiadocontidae, Edaphosauridae, and Sphenacodontidae were among the taxa that became extinct then. However, Lucas (2017) argued that it was a "non-event" that represents a normal level of faunal turnover and he later asserted that it simply represents the end of his very long-lasting (about 15-20 Ma) Coyotean chronofauna (Lucas, 2018, p. 429), which extends approximately from the late Gzhelian to the late Sakamrian. In any case, it is now clear that Ophiadocontidae, Edaphosauridae, and Sphenacodontidae persisted at least until the Kungurian, which is well after Lucas' Coyotean chronofauna.

The second possible mass extinction event in Permian tetrapods, which we study here, may have occurred near the Kungurian/Roadian stage boundary (Sahney and Benton, 2008; Brocklehurst *et al.*, 2013; Lucas, 2017; Brocklehurst, 2018), which is also the Cisuralian (early Permian)/Guadalupian (middle Permian) series boundary (272.3 Ma). It seems plausible that some of the extinctions (among others,

those of Ophiadocontidae, Edaphosauridae, and Sphenacodontidae) that had at some point been postulated to have taken place in the Sakmarian and/or Artinskian actually took place toward the end of the Kungurian, or slightly later. The observed stratigraphic range of Ophiadocontidae and Edaphosauridae ends shortly before the top of the Kungurian, whereas Sphenacodontidae may well extend into the early Roadian, given the controversial and poorly constrained age of the San Angelo Formation (see below). Sahney and Benton (2008) called this event "Olson's extinction" and considered that it was the first of three Permian crises (they seem to have ignored the possibility of an earlier crisis in the Artinskian, which their data were not designed to reveal). They did not date it very precisely, mentioning only that it had taken place in the Roadian (272.3-268.8 Ma) and/or Wordian (268.8-265.1 Ma), but estimated that it "reveals an extended period of low diversity when worldwide two-thirds of terrestrial vertebrate life was lost" (Sahney and Benton, 2008, p. 760). Olroyd and Sidor (2017) suggested, among other hypotheses, that extinction of most Permo-Carboniferous synapsids at the end of the Kungurian could have allowed their non-competitive replacement by therapsids, which is supported by the results of Sahney and Benton (2008) and Brocklehurst *et al.* (2013). This hypothesis is tested indirectly here (see below). Lucas (2018, p. 430) suggested that rather than a single large crisis, a few events (which he called "Redtankian events", after his Redtankian chronofauna) took place in the Kungurian. Lucas (2018, p. 430) suggested that ophiacodontids became extinct before edaphosaurids, but that both clades were extinct in the early Redtankian. Sphenacodontids became extinct later because they occur in the Littlecrotonian (Lucas, 2017, p. 43). However, this is based on a literal interpretation of the fossil record; no attempts have been made at assessing confidence intervals for the extinction times of relevant taxa, as far as we know, although Brocklehurst (2018, 2020) studied "Olson's extinction" through other methods. Our study aims at filling this gap.

The third mass extinction event took place near the end of the Capitanian (259.8 Ma), the last stage of the Guadalupian (middle Permian), around the time of the Emeishan volcanism in southern China (Day *et al.*, 2015a; Lucas, 2017). It wiped out the dinocephalians (a fairly large clade of Guadalupian therapsids), although this apparently occurred gradually in the *Tapinocephalus* Assemblage Zone (Day *et al.*, 2015b). Other therapsid taxa also appear to have been affected by this crisis, which apparently also influenced the marine realm (Day *et al.*, 2015a). Varanopid synapsids also appear to have become extinct then (Modesto *et al.*, 2011); this was the last of the Permo-Carboniferous synapsid clades to become extinct because caseids are not currently known after the Roadian. Maddin *et al.* (2008) assigned a mid-Capitanian age to the caseid *Ennatosaurus tecton*, but more recently, Golubev (2015) assigned a Roadian age to the locality (Moroznitsa) of the holotype. Lucas (2017) argued that parareptiles were also affected by this crisis, but other studies suggests that parareptiles were only affected by background extinctions throughout the Guadalupian (Ruta *et al.*, 2011; Cisneros *et al.*, 2020). Day *et al.* (2015a) estimated that there was a 74-80% loss of generic richness in amniotes in this crisis, and that it was not as severe as the end-Permian crisis. On the contrary, Lucas (2017) considered that this extinction event was more severe for amniotes than the much better-known end-Permian event and that both may have lasted longer than previously thought. In the seas, goniatitids and most fusulinids disappeared around that time, but ammonoids do not appear to have been strongly affected (Lucas and Shen, 2018, p. 34). Furthermore, Day *et al.* (2015a, p. 6) attributed this extinction to the short-term cooling and darkness induced by the Emieshan volcanism, as well as related acidification.

The fourth and best-known of the Permian mass extinction event took place at the Permian/Triassic boundary (251.9 Ma). Like most such events, it has been studied mostly in the marine realm (Raup and Sepkoski, 1982; Payne, 2005; Brayard *et al.*, 2009; Romano *et al.*, 2013), where it is estimated that between 80 and 96% of species were eliminated (Sahney and Benton, 2008), but recent studies have shown a significant crisis in continental vertebrates as well (Benton, 2003; Ward *et al.*, 2005). Thus, Smith *et al.* (2012, p. 47) reported that out of 41 therapsid genera present in the middle of the *Dicynodon* Assemblage Zone, only three survived the end-Permian extinction (not taking into consideration ghost lineages that imply a slightly greater proportion of survivors). The extinctions on land and in the seas seem to have been approximately synchronous (Smith and Botha-Brink, 2014). Lucas (2017) claims that this extinction event lasted longer than previously claimed and occurred in a stepwise manner. In fact, it was once thought that this crisis lasted over the last 10 Ma of the Permian (Erwin, 1990), but that was before the end-Guadalupian crisis was identified. More recent studies point to a much shorter crisis, especially on the continents. Among recent works on this topic, Viglietti *et al.* (2021) differs by using a fine stratigraphic resolution of 13 time bins lasting about 300 000 years each, thus covering in detail evolution of tetrapod biodiversity in the Karoo around the P/Tr boundary. This approach was timely

because as they stated, using a fine stratigraphic scale is important to better understand mass extinction events. It allowed Viglietti *et al.* (2021) to determine that the crisis among continental tetrapods of the Karoo Basin lasted about 1 Ma, which is longer than the marine crisis, which is believed to have lasted only about 61 000 years (Burgess *et al.*, 2014; Liu *et al.*, 2020). These studies highlight the importance of using high-resolution stratigraphic data to study mass extinction events.

As can be seen from this very brief review, a recurring question is how long each crisis lasted. In other words, did most taxa become extinct at about the same time (in a few tens of thousand years, possibly a few hundred thousand years), or were the extinctions spread over a few million years? Settling this question ideally requires abundant, well-dated, and geographically-widespread data, as well as appropriate analytical methods to discriminate between genuine gradual extinctions and the Signor-Lipps effect on taxa with a relatively scarce fossil record, as is typically the case for continental vertebrates.

Our dataset is relevant to assess extinction times that span the first and second of the possible tetrapod Permian mass extinction events. Our method is not designed to assess fluctuations in extinction rates; rather, our objective is to obtain a better understanding of the timing of extinction of various taxa to either corroborate or refute previous statements about such events. More specifically, we test the following hypotheses:

1. Many ophiacodontids, edaphosaurids, and sphenacodontids had become extinct well before the end of the Kungurian (which is consistent with a prolonged crisis, or a series of crises, rather than with a single, catastrophic, sudden event at the end of the Kungurian; this can also test the existence of the first of the four crises listed above, near the Artinskian/Kungurian boundary).

2. Ophiadocontidae, Edaphosauridae, and Sphenacodontidae became extinct (gradually or not) by the end of the Kungurian, at the latest (Benton, 1989; Brocklehurst *et al.*, 2013; Lucas, 2017);

3. These three clades became extinct in the following order: Ophiacodontidae, Edaphosauridae, and Sphenacodontidae (Lucas, 2018, p. 430).

We test the first hypothesis by verifying the proportion of terminal branches (observed nominal species) of these three clades that became extinct before the end of the Kungurian (i.e., more than 95% of their extinction probability is before the end of the Kungurian, which is 272.3 Ma). A substantial proportion of lineages becoming extinct before the end of the Kungurian would be compatible with a gradual extinction of these clades, even though additional tests will be required to prove this hypothesis. If a high proportion of the early extinctions were concentrated in time (especially around the Artinskian/Kungurian boundary), this would be compatible with (but would not prove) the possibility of a crisis around that time.

We test the second hypothesis by verifying if the extinction-density probability of the three clades is compatible with an extinction of these clades by the end of the Kungurian. Contrary to previous methods, ours considers the extinction times of lineages that have not been preserved in the fossil record but that are very likely to have existed because of the speciation, extinction and fossilization rates. The fact that Didier *et al.* (2017, p. 981) estimated that only about 14% of the eupelycosaur lineages (defined as an internode on the tree) had left a fossil record suggests that taking into consideration unobserved lineages can have a major impact on our estimates of the extinction times of these three large eupelycosaur clades. Hence, these times should be somewhat later than the extinction of the most recent known lineage of each of these clades, and potentially, substantially more recent than the last observed fossil of each of these clades.

The third hypothesis is tested by looking at the peak probability density of extinction time and the end of the 95% confidence interval of the extinction time of all three clades and by computing the probability of a given clade becoming extinct before another clade under the FBD model.

Note that testing hypotheses 1 and 2 above amounts to testing indirectly the suggestion that the replacement of Permo-Carboniferous synapsids by therapsids was non-competitive (Olroyd and Sidor, 2017, p. 593) because a literal reading of the fossil record suggests that therapsid diversification accelerated sharply in the Roadian and that this is after the extinction of Ophiacodontidae, Edaphosauridae, and Sphenacodontidae. If this is correct, this replacement was non-competitive. On the contrary, if Ophiacodontidae, Edaphosauridae, and Sphenacodontidae became extinct only in the Wordian or later (time at which theraspids were already abundant in most terrestrial assemblages), this would suggest a competitive replacement.

These three hypotheses are not mutually exclusive. For instance, these clades could have become extinct gradually by the end of the Kungurian in the order suggested by Lucas (2018, p. 430), in which case all three hypotheses would be correct. Alternatively, these three clades might also have become extinct suddenly after the Kungurian and in a different sequence than stipulated by Lucas (2018, p. 430), even though this is not what a literal interpretation of their fossil record suggests; in this case, all three hypotheses would be false. But any combination is possible.

Corrections for multiple tests are not required in this study because the goal of the paper is not to test n times that n species became extinct before the Kungurian/Roadian boundary (which would indeed amount to making a high number of tests). Rather, we want to get a picture of the pattern of extinction (gradual vs simultaneous) and assess where it fits compared to the Kungurian/Roadian boundary. The actual number of tests (for which we report probabilities) performed in this study is fairly low.

Methods

Methods to estimates extinction times which are presented below all return a *confidence upper bound of the extinction date at a given order* (95% is the usual choice), i.e., the time t which is such that the probability for the extinction date to be anterior to t is equal to the order required.

Branch-by-branch approaches

We present here several previous approaches which deal with each taxon independently of the rest of the tree, in the sense that in order to estimate the extinction time of a given (extinct) taxon, they only require the number n of its fossils and their ages $f_1, f_2, \ldots, f_n$. We assume that the fossil ages are given in chronological order, so f_1 and f_n are the most ancient and most recent fossil ages respectively.

The approach of Strauss and Sadler (1989) makes the assumption that the fossil ages of a given extinct taxon are uniformly distributed into a time interval of unknown bounds, which are its speciation and extinction ages. By using properties of uniformly distributed samples, they provided, amongst others, an explicit expression of confidence upper bound of the extinction age of a given taxon at order o, namely $f_n + (1-o)^{-\frac{1}{n-1}}(f_n - f_1)$.

McInerny *et al.* (2006) proposed a method to infer extinction times from sighting records under a uniformity assumption close to that Strauss and Sadler (1989). Let us briefly expose their approach by transposing it into the geological case. The method assumes that the probability of fossilization per unit of time is constant and estimates it as $\frac{n-1}{f_n - f_1}$ (in the case where the first fossil age is used as start time, the number of fossiliferous horizons considered reduce to $n-1$ and let us assume here that at most a single fossiliferous horizon is observed per unit of time). The probability for the extinction time to be posterior to any time $T > f_n$ is that of not observing any fossil between f_n and time T given that the considered lineage does not go extinct before T, which follows a geometric distribution, $\left(1 - \frac{n-1}{f_n - f_1}\right)^{T - f_n}$, and from which we can get a confidence upper bound at a given order o: $f_n + \frac{\log(1-o)}{\log\left(1 - \frac{n-1}{f_n - f_1}\right)}$. A continuous version of this approach is basically obtained by replacing the geometric distribution by an exponential distribution in the formula just above. The probability for the the extinction time to be posterior to a time T becomes $\exp\left(-\frac{n-1}{f_n - f_1}(T - f_n)\right)$ and the confidence upper bound at order o is then $f_n - \frac{\log(1-o)}{\left(\frac{n-1}{f_n - f_1}\right)}$. Note that $\frac{n-1}{f_n - f_1}$ is now interpreted as the rate of a Poisson process modelling the fossilization (and we no longer have to assume that at most a single fossiliferous horizon is observed per unit of time).

Bradshaw *et al.* (2012) proposed a method based on that of McInerny *et al.* (2006) in order to estimate extinction times from the fossil record. They first consider a weighted combination of the results provided by the method of McInerny *et al.* (2006) on the k most recent fossils for $k = 1$ to $n-1$ with weights inversely proportional to $f_n - f_{n-k}$, in order to give more weight to the most recent fossils. They also take into account the uncertainty in dating fossils by resampling the fossil ages from Gaussian distributions and by considering 95% confidence limits.

Alroy (2014) proposed a Bayesian iterative approach to infer extinction date – see Solow (2016) and Alroy (2016) for a discussion of this method. Briefly, the time after the most recent fossil age is divided into equal subintervals indexed on $1, 2, \ldots$. Under the notations of Alroy (2014), let p_s be the probability that a particular interval contains a fossil age given that the taxon is not extinct at this time interval, E be

a prior probability of extinction and ϵ_t be the posterior probability of extinction before the tth interval (Alroy, 2014, proposed different ways to estimate the two first quantities). The method iteratively computes the posterior probability of extinction before all intervals t by setting $\epsilon_1 = \frac{E}{E+(1-E)(1-p_s)}$ and, for all $t > 1$, $\epsilon_t = \frac{\epsilon_{t-1}+(1-\epsilon_{t-1})E}{\epsilon_{t-1}+(1-\epsilon_{t-1})E+(1-\epsilon_{t-1})(1-E)(1-p_s)}$. The confidence upper bound of the extinction date at order o is the upper bound of the last interval t such that $\epsilon_t \leq o$.

A basic global method

The approach of McInerny *et al.* (2006) and its continuous version can be adapted in order to take into account the whole tree by considering all the fossiliferous horizons for the estimation of the probability or the rate of fossilization. This global estimation is performed by considering all the differences between the ages of the successive fossils of the phylogenetic tree. In the continuous case, these time differences are assumed to be exponentially distributed with the rate of the Poisson process modelling the fossilization, and we consider its maximum likelihood estimate $\hat{\varphi}$. The 'global' fossilization rate estimate $\hat{\varphi}$ is next used to get the confidence upper bound at order o of the extinction time of any extinct taxa of the tree as above, i.e., $f_n - \frac{\log(1-o)}{\hat{\varphi}}$ in the continuous case.

The FBD model

The FBD model was introduced in Stadler (2010) and has been referred to as the "Fossilized-Birth-Death" model since Heath *et al.* (2014). This model assumes that the diversification process starts with a single lineage at the origin time, which is one of its parameters. Next, each lineage alive evolves independently until its extinction and may be subject during its lifetime to events of speciation (here equated with cladogenesis, which leads to the birth of a new lineage), extinction (which terminates the lineage) or fossilization (which leaves a fossil of the lineage dated at the time of the event) which occur with respective rates λ, μ and ψ, which are the main parameters of the model. Last, the extant lineages (if any) are sampled at the present time with probability ρ, the last parameter of the model.

Let us recall the probabilities of the following basic events, derived in (Stadler, 2010; Didier *et al.*, 2012, 2017) under the FBD model, which will be used to compute various probability distributions on the observable part of realizations of the FBD process (Fig. 1 here or Didier and Laurin, 2020, fig. 1). The probability $\mathbf{P}(n,t)$ that a single lineage starting at time 0 has n descendants sampled with probability ρ at time $t > 0$ without leaving any fossil (i.e., neither from itself nor from any of its descendants) dated between 0 and t is given by

$$\mathbf{P}(0,t) = \frac{\alpha(\beta-(1-\rho)) - \beta(\alpha-(1-\rho))\mathrm{e}^{\omega t}}{\beta-(1-\rho)-(\alpha-(1-\rho))\mathrm{e}^{\omega t}} \text{ and}$$
$$\mathbf{P}(n,t) = \frac{\rho^n(\beta-\alpha)^2\mathrm{e}^{\omega t}\left(1-\mathrm{e}^{\omega t}\right)^{n-1}}{\left(\beta-(1-\rho)-(\alpha-(1-\rho))\mathrm{e}^{\omega t}\right)^{n+1}} \text{ for all } n > 0,$$

where $\alpha < \beta$ are the roots of $-\lambda x^2 + (\lambda+\mu+\psi)x - \mu = 0$, which are always real (if λ is positive) and are equal to

$$\frac{\lambda+\mu+\psi \pm \sqrt{(\lambda+\mu+\psi)^2 - 4\lambda\mu}}{2\lambda},$$

and where $\omega = -\lambda(\beta-\alpha)$. Note that $\mathbf{P}(0,t)$ is the probability that a lineage alive at time 0 goes extinct before t without leaving any fossil dated between 0 and t.

The probability density $\mathbf{D}(t)$ for a lineage alive at time 0 to go extinct exactly at time t (without leaving any fossil) is basically obtained by deriving $\mathbf{P}(0,t)$ with regard to t. We get that

$$\mathbf{D}(t) = \frac{\lambda(\alpha-(1-\rho))(\beta-(1-\rho))(\beta-\alpha)^2\mathrm{e}^{\omega t}}{(\beta-(1-\rho)-(\alpha-(1-\rho))\mathrm{e}^{\omega t})^2}.$$

In the particular case where $\rho = 1$, expressions above simplify to

$$\mathbf{P}(0,t) = \frac{\alpha\beta(1-\mathrm{e}^{\omega t})}{\beta-\alpha\mathrm{e}^{\omega t}} \;,\; \mathbf{P}(n,t) = \frac{(\beta-\alpha)^2\mathrm{e}^{\omega t}\left(1-\mathrm{e}^{\omega t}\right)^{n-1}}{\left(\beta-\alpha\mathrm{e}^{\omega t}\right)^{n+1}} \text{ for all } n > 0 \text{ and } \mathbf{D}(t) = \frac{\mu(\beta-\alpha)^2\mathrm{e}^{\omega t}}{(\beta-\alpha\mathrm{e}^{\omega t})^2}. \tag{1}$$

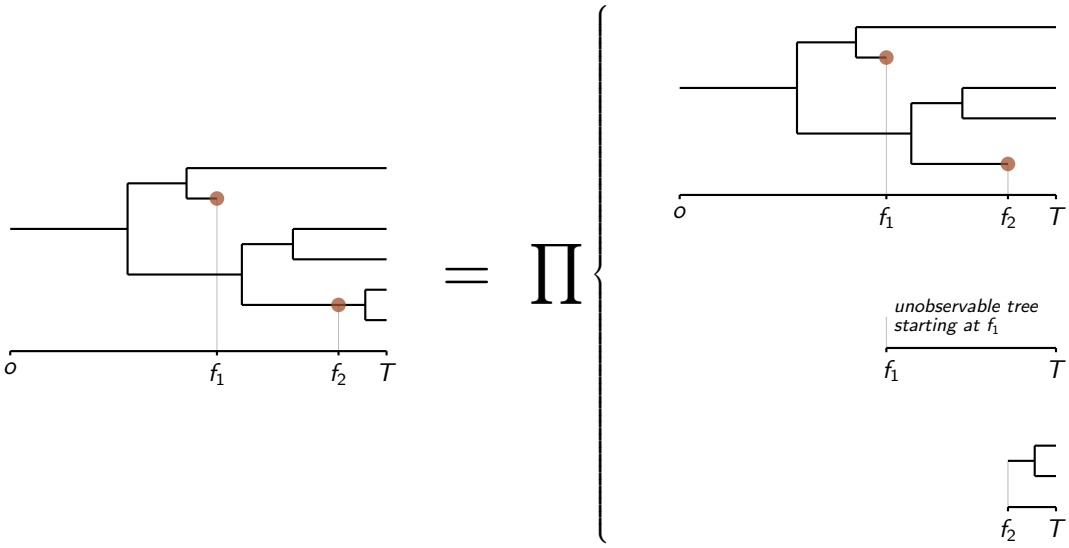

Figure 2: Decomposition of the probability density of a phylogenetic tree of extinct and extant taxa with fossils (figured by brown dots) as the product of the probability densities of "basic trees" by cutting it at each fossil find (adapted from Didier and Laurin, 2020).

Extinction time distributions

We consider here a dataset consisting of a phylogenetic tree $\mathcal{T}$ of extant and extinct taxa with fossils, which is interpreted as the observable part of a realization of the FBD process, and of the fossil age vector $\mathbf{f}$. We aim to compute the joint probability density of $(\mathcal{T}, \mathbf{f})$ and that a particular subset S of taxa of $\mathcal{T}$ goes extinct before a time t under the FBD model. This question makes sense only if the taxa of S all go extinct before the present time. In the empirical example provided below, these subsets consist of three clades, but the method can handle any set (monophyletic or not) of taxa.

This joint probability density can be computed in the exact same way as that of the dataset $(\mathcal{T}, \mathbf{f})$ presented in Didier and Laurin (2020). To show this, we adapted figure 3 from Didier and Laurin (2020), which illustrates the fact that, from the Markov property, the probability density of $(\mathcal{T}, \mathbf{f})$ can be written as the product of that of the "basic trees" obtained by splitting $\mathcal{T}$ at each fossil find (Fig. 2). The tree of Figure 2 contains only one extinct taxon that goes extinct between f_1, the age of its most recent fossil, and the present time T. The contribution factor of the basic tree starting from time f_1, which is unobservable from f_1 to the present time, is the probability that a lineage alive at f_1 goes extinct before the present time T without leaving any fossil dated between f_1 and T, i.e, $\mathbf{P}(0, T - f_1)$. Computing the joint probability density of the dataset of Figure 2 and that its extinct lineage goes extinct before a given time t (resp. exactly at a given time t) is performed by setting the contribution factor of the basic tree starting from time f_1 to $\mathbf{P}(0, t - f_1)$ (resp. to $\mathbf{D}(t - f_1)$).

Let us now consider a subset S of extinct taxa of $\mathcal{T}$, for instance, an extinct clade of $\mathcal{T}$. The joint probability density of the dataset $(\mathcal{T}, \mathbf{f})$ and that S goes extinct before a given time t, which is basically the joint probability density of $(\mathcal{T}, \mathbf{f})$ and that all the taxa of S goes extinct before t, is obtained by setting, for all taxa $n \in S$, the contribution factor of the unobservable basic tree pending from the leaf n, which starts from time $\ell_{\mathbf{f},n}$, i.e., the age of the most recent fossil of n, to $\mathbf{P}(0, t - \ell_{\mathbf{f},n})$, that is the probability that a lineage alive at $\ell_{\mathbf{f},n}$ goes extinct before t without leaving any fossil dated between $\ell_{\mathbf{f},n}$ and t. The computation of this joint probability density is thus computed in the very same way and with the same algorithmic complexity as that of the probability density of $(\mathcal{T}, \mathbf{f})$ presented in Didier and Laurin (2020).

The computation presented above allows us to determine the probability distribution of the extinction time of a subset of taxa in the case where the tree topology, the fossil ages and the parameters of the FDB model are exactly known, a situation which is unfortunately never met in practice. Namely, the rates of the FBD model are unknown, morphological data of the fossils may be consistent with several

topologies, and fossil ages are provided as time intervals of plausible ages.

Given a priori probability distributions over the parameters and the data, a standard way to deal with the uncertainty on the model parameters and the data consists in integrating the extinction time distribution over all the possible values of the FBD parameters, of the fossil ages (constrained within confidence intervals) and over all tree topologies included in a population of trees. We implemented the numerical computation of this integration by Monte Carlo Markov Chain (MCMC, see Appendix A).

Note that determining the whole distribution of the extinction time is generally not required. Getting a confidence upper bound of the extinction date at a given order (usually 95%) provides sufficient information for most purposes. In our framework, this bound is basically the quantile at the order required of the extinction time distribution.

If one is interested in comparing the extinction dates of two given extinct clades (or more generally two subsets of extinct taxa) A and B, it is possible to compute the probability that A goes extinct before B under the FBD model. In Appendix B, we provide an explicit formula for this probability.

Empirical dataset

Our dataset includes a subset of the taxa represented in the dataset used by Didier and Laurin (2020), from which we simply extracted the smallest clade that includes Ophiacodontidae, Edaphosauridae and Sphenacodontidae. However, we incorporated all the information that we could find about the relevant taxa in the Paleobiology Database (Alroy *et al.*, 2012). Thus, we now have more fossil occurrences than in our previous studies that used this dataset (Didier *et al.*, 2017; Didier and Laurin, 2020). Most of the fossil occurrences in our database are from south-western North America, but a few (like *Archaeothyris florensis*) come from the eastern part of that continent, as well as from Europe (such as *Haptodus baylei*). We also updated the geological ages using the recent literature (see below). Finally, we updated our supermatrix (which is used to produce the population of source trees) to include more recent studies that update scores for various taxa and provide scores for taxa for which the earlier version of our matrix had no data (the position of these taxa was specified using a skeletal constraint). Specifically, we replaced the matrix of Benson (2012) by that of Mann and Paterson (2020), which is the most recent update of that matrix. To better resolve the phylogeny of edaphosaurids and sphenacodontids, we incorporated also the matrix from Brocklehurst and Brink (2017). This allowed us to reduce the number of taxa in the skeletal constraint, which we updated, notably to reflect current ideas about the affinities of *Milosaurus* (Brocklehurst and Fröbisch, 2018). The 100 trees used here were sampled randomly from three searches that each yielded 200 000 equiparsimonious trees (the maximal number of trees that we allowed in the search) that apparently include at least two tree islands, and we verified that this sample included trees of both islands. Additional searches with a lower number of maximal trees and a different seed number for the random addition sequence were carried out to verify that these were the shortest trees. We also carried out a single search with a maximal number of 15 000 000 trees. All these searches confirmed that we appear to have recovered all tree islands and that these include some of the shortest trees. All these new searches included only the taxa studied here, plus *Varanops brevirostris* (the varanopid with the highest number of characters scored into our supermatrix), which was used as the outgroup.

Our method, contrary to those derived from Strauss and Sadler (1989), Marshall (1994), and Marshall (1997), ideally requires stratigraphic data expressed in absolute age (Ma) rather than height in a section because we model events in time. It could use section height, provided that sedimentation rates were more or less homogeneous in time, but this would be less interesting (evolution happens in geological time, not in strata) and would be meaningless if applied to sections of different basins in which sedimentation rates were not necessarily comparable. In any case, these would still have to be correlated to each other. We thus relied on the literature to convert stratigraphic position of the fossils into approximate ages (represented by age ranges, which our method samples using a flat distribution). We tried to adopt the established consensus, as reported, for part of the fossiliferous basins in SW USA, by Nelson *et al.* (2013) and the age assignments usually attributed to these formations as interpreted (for much of the relevant stratigraphic range) by Lucas (2018, fig. 4), using the most recent geological timescale, as summarized by Lucas and Shen (2018, fig. 9) and Schneider *et al.* (2020). For instance, according to Lucas (2018, fig. 4), the Moran Formation extends from the latest Asselian to the late Sakmarian, and the overlying Putnam Formation extends from that time to the end of the Sakmarian (approximately). Given that Lucas and Shen (2018, fig. 9) report an age interval of 298.9 to 295 Ma for the Asselian and 295 to 190.1 Ma for the Sakmarian, we assigned age intervals of 295.5 to 291.5 Ma to the Moran Formation, and from

291.5 to 290 Ma to the Putnam Formation, and to the fossils found therein. Thus, the oldest of the two occurrences of *Lupeosaurus kayi* documented in the Paleobiology Database is from the Moran Formation and was thus assigned an age of 295.5 to 291.5 Ma, as can be seen in the supplements. This dating scheme admittedly involves much interpolation between strata that have been dated by radiometric methods, but it can be updated as the geological timescale is refined. As in our previous studies, we entered all documented stratigraphic horizons for all included taxa. Even if a few horizons were too close in stratigraphic height to distinguish their ages, they were entered separately, but with the same age range. For instance, *Sphenacodon ferocior* occurs in four localities of the Abo Formation (297 to 290.1 Ma), according to the Paleobiology Database; it has thus been scored as occurring four times in strata of this age bracket because we assume that each locality represents a distinct level. However, our analysis is carried out at the fossiliferous horizon level; even if many specimens of a given taxon were found in a given horizon, this is still scored as a single occurrence. In the absence of more detailed information, we also assume that a single level in each locality yielded each of the relevant taxa.

Contrary to Lucas (2018), we consider that most of the Pease River and El Reno Groups are Guadalupian, as previously suggested by one of us and as supported by other studies based on various types of evidence (Clapham, 1970; Reisz and Laurin, 2001, 2002; Nelson and Hook, 2005; Foster *et al.*, 2014; Soreghan *et al.*, 2015). In this study, a single taxon occurs in these strata, namely *Dimetrodon angelensis*, in the San Angelo. We consider that the range of possible ages of this formation, which is at the base of the Pease River Group, straddles the Kungurian/Roadian boundary, from 275 Ma to 271 Ma. The stratigraphic database used here, along with the trees, is available in the supplements.

Simulation study

In order to assess the accuracy of the upper bound of the extinction time corresponding to a given confidence level (here the usual 95%) obtained from the approaches presented in Section 'Methods', we simulated the diversification and the fossilization of a clade under the FBD model with five fossilization rates. The simulated data consist of the observable parts of the simulated diversification processes but we also stored the extinction time of all the extinct taxa (including those that are not observable) in order to assess the accuracy of different approaches.

In the simulation case, we have access to the (exact) tree topology, the fossil ages and the model parameters. In this section, we shall assume that both the tree topology and the fossil ages are exactly known but consider both the confidence upper bounds obtained from the model parameters used to simulate the data, which is the reference method since the extinction ages follow the distributions from which these upper bounds are computed in this case, and the confidence upper bounds obtained without the knowledge of the model parameters, i.e. by dealing with their uncertainty by integrating over all the possible model parameters (under the assumption that they follow improper uniform distributions; see the section above and Appendix A).

For each simulation, the "real" extinction times of all the extinct taxa were stored and compared to the upper bounds obtained from each method. Namely, the accuracy of each method is assessed with regard to two features:

1. their percentage of errors, that is the proportion of situations in which the real extinction date is posterior to the upper bound provided and,

2. the average width of the confidence interval, that is the average duration separating the age of the most recent simulated fossil of each taxon from the 95% confidence upper bound of its estimated extinction time.

The simulations were obtained by running a FDB process with speciation and extinction rates $\lambda = 0.2$ and $\mu = 0.19$ per lineage and per million year during 200 millions years. These settings yield an expected lineage duration of about 5 millions years. Note that we simulate diversification and fossilization through evolutionary time, not through strata; thus, all simulations aim at assessing the performance of methods at inferring confidence intervals of extinction times, rather than levels in a section. We considered five fossilization rates: $\psi = 0.005$, 0.01, 0.1, 1 and 5. The simulations were filtered for technical reasons. We kept only those leading to an observable tree of size from 50 to 100 nodes and containing more than 20 fossils.

The six following ways to obtain a 95% confidence upper bound of the extinction date were assessed:

S&S: the method of Strauss and Sadler (1989),

Alr: the method of Alroy (2014); since several ways of computing the quantities used to computed the posterior probabilities required by this method are proposed in Alroy (2014), we tested some of them and kept those leading to the best performance (in particular, we set $p_s = \frac{n_p - 1}{n_p - 1 + n_a}$, under the notation of Alroy, 2014, which allows us to apply this approach to terminal branchs with at least two fossiliferous horizons),

McI: the continuous version of the approach of McInerny *et al.* (2006),

Glo: the global version of **McI** (Section 'A basic global method'),

Int: the 95% quantile of the extinction time distribution obtained by integrating uniformly over the speciation, extinction and fossilization rates (numerically estimated from the MCMC procedure presented in Appendix A, which was applied with the following settings. For each fossil recovery rate and each of the 1000 corresponding simulated trees, we discarded the first 10000 iterations of the Markov chain; then, we kept 1 iteration out of 10 until getting a sample of 5000 sets of rates from which the 95% quantiles of the extinction time distributions associated to all the extinct taxa are computed),

Ref: the 95% quantile of the extinction time distribution computed from the parameters used to simulate the data, which is the reference method (unfortunately, this method cannot be applied in practice because we do not know the actual rates of speciation, extinction and fossilization in real situations).

We did not display the results obtained by the method of Bradshaw *et al.* (2012) because they were systematically worse than those obtained with method **McI**. This was expected since the method of Bradshaw *et al.* (2012) is a modified version of that of McInerny *et al.* (2006) in a way which cannot improve performances in our simulation context, i.e., uniformity of the fossil rate and exact fossil ages.

All the methods assessed do not apply to all cases. Methods **S&S**, **Alr** and **McI** requires at least two fossils in the branch of the considered extinct taxa while method **Glo** can be applied only on trees containing at least a branch bearing more than two fossils (it implies that if methods **S&S**, **Alr** and **McI** can be applied then so can be method **Glo**). Only methods **Int** and **Ref** can be applied to all extinct taxa.

Tables 1 to 3 display the results obtained over 1000 simulated trees for each fossilization rate.

Fos. rate	% Accepted trees	% Branch feasible	% **Glo** feasible	Total extinct taxa
0.005	0.27	7.42	85.36	20 787
0.01	0.66	10.19	95.62	25 084
0.1	1.44	33.83	100.00	35 928
1	2.59	75.02	100.00	36 088
5	2.96	93.07	100.00	35 820

Table 1: Simulated extinct taxa statistics for all fossilization rates. Column 2 displays the percentage of simulated trees which satisfy the size and number of fossils required. Columns 3 (% Branch feasible) and 4 (% **Glo** feasible) display the percentage of extinct taxa on which methods **S&S**, **Alr** and **McI** can be applied and on which method **Glo** can be applied, respectively. The last column shows the total number of extinct taxa observed over all the simulations. The average number of extinct taxa per tree is obtained by dividing the last column by the number of (accepted) simulated trees, here 1000, for each fossil recovery rate, and it ranges from 20 (top row) to 36 taxa (bottom row), on average.

Column 2 of Table 1 displays the acceptation percentages of trees, that is the percentages of simulated trees which satisfy our constraints on the size of the tree and on the number of fossils (so we have to simulate 1000 trees with fossils divided by this percentage, in our protocol, to get 1000 simulated trees satisfying these constraints for all fossil recovery rates). We observe that these percentages are low, ranging from 0.27 to 2.96. One thus expects the sample of 1000 simulated trees for each fossilization rate to be biased with regard to the corresponding FBD rates, and may worry about the effect of this bias on our analyses. Fortunately, this effect seems limited. Appendix D displays the posterior distributions

of rates obtained from the MCMC and shows that they are well centered around the rate values used to simulated the corresponding trees.

As expected, Table 1 shows that the proportion of extinct taxa on which the methods **S&S**, **McI** and **Alr** can be applied, i.e. those corresponding to terminal branches bearing at least two fossils, may be very low when the fossil recovery rate is small (Table 1, Column 3). Within the covered simulation space, this proportion ranges from about 7% when the fossil recovery rate is four times lower than the speciation rate, to about 93% when the fossil recovery rate is 25 times the speciation rate (note that in 3 of the 5 sets of simulations, the fossil recovery rate is much higher than the net diversification rate, which is speciation minus extinction rates). This proportion increases with the fossil recovery rate. In our simulations, the "branch-by-branch" approaches require a fossil recovery rate several times higher than the speciation and extinction rates in order to be applicable on the majority of extinct taxa. These requirements may be severely limiting because our current dataset yields fossilization rates only about twice as high as the speciation and extinction rates, and the previous version yielded much lower fossilization rates (Didier and Laurin, 2020). Even a neogene camelid dataset, which could be expected to have a much denser fossil record, only yielded fossilization rates a little more than twice as high as those of speciation and extinction (Geraads *et al.*, 2020). Method **Glo** can be applied to all extinct branch belonging to trees which contain at least a branch bearing at least two fossils. There is a relatively small proportion of cases where this method cannot be applied, and these occur only for the two lowest fossil recovery rates (Table 1, Column 4).

The total number of extinct taxa, displayed in the last column of Table 1, starts by increasing with the fossilization rate, then fluctuates around 36000 for all the simulations when the fossilization rate is close to, or greater than the speciation and extinction rates. This saturation phenomena was expected because phylogenetic trees simulated with the speciation and extinction rates of our protocol have a certain average number of extinct branches among which only those with fossils appear as extinct taxa in the final simulated tree (Fig. 1) and the probability that an extinct branch bears at least a fossil tends to 1 as the fossil recovery rate increases.

Fos. rate	**S&S**	**McI**	**Alr**	**Glo**		**Int**			**Ref**		
0.005	2.40	13.35	11.34	6.48	*8.02*	4.93	*5.30*	**5.38**	4.67	*5.07*	**5.12**
0.01	3.80	16.87	13.97	6.18	*6.93*	5.32	*5.90*	**5.80**	5.05	*5.20*	**5.18**
0.1	3.92	18.30	11.85	4.15	*4.33*	5.16	*5.54*	**5.46**	4.86	*4.95*	**4.95**
1	4.93	14.62	6.32	4.64	*4.69*	4.78	*4.86*	**4.86**	4.74	*4.78*	**4.78**
5	5.27	9.40	2.09	5.17	*5.18*	5.20	*5.20*	**5.20**	5.11	*5.10*	**5.10**

Table 2: Error percentages obtained from the 95% confidence upper bounds provided by the six methods on the simulated extinct taxa where the **S&S**, **McI** and **Alr** computations were feasible (plain text), on those belonging to a simulated tree where a global fossil recovery rate can be estimated (italics; only for methods **Glo**, **Ref** and **Int**) and on all the extinct taxa (bold; only for methods **Ref** and **Int**).

The *error percentage*, that is the percentage of cases where the "real" extinction date is posterior to the confidence upper bound provided, fluctuates around 5%, which is the level of risk required here, both for methods **Ref** and **Int** (Table 2). This was expected for method **Ref** because its confidence upper bound is computed according to the actual distribution of the extinction date under the simulation model in this case. The closeness of the error percentage of method **Int** to the level of risk required suggests that integrating the extinction date distribution over all the possible FBD rates leads to accurate results (see also Appendix D). The error percentage obtained from methods **S&S**, **McI**, **Alr** and **Glo** depend heavily on the fossil recovery rate and is generally far from the level of risk required (it is actually close to the level of risk required only for methods **S&S** and **Glo** with the two highest fossil recovery rates).

As expected, the average difference between the confidence upper bound provided by each method and the age of the most recent fossil, which we call the *mean confidence interval width*, tends to be reduced when the fossil discovery rate increases for all the methods assessed (Table 3). The mean confidence interval widths obtained with the method **Int** are close to those obtained from the reference method **Ref**, again supporting the relevance of our way of dealing with the uncertainty on the speciation, extinction and fossilization rates. The mean confidence interval widths provided by methods **S&S**, **McI**, **Alr** and **Glo** are systematically greater or equal to those of methods **Int** and **Ref**. In particular, the mean confidence interval widths of method **S&S** are several times greater than those obtained with the

Fos. rate	**S&S**	**McI**	**Alr**	**Glo**			**Int**			**Ref**	
0.005	244.0	40.5	36.0	40.4	*41.2*	26.3	*25.8*	**25.9**	26.1	*26.0*	**26.0**
0.01	181.1	30.5	28.9	30.8	*30.6*	21.6	*21.2*	**21.2**	21.7	*21.8*	**21.8**
0.1	50.5	10.3	14.0	10.3	*10.3*	9.0	*9.0*	**9.0**	9.2	*9.2*	**9.2**
1	6.3	2.2	4.7	2.2	*2.2*	2.2	*2.2*	**2.2**	2.2	*2.2*	**2.2**
5	0.9	0.6	1.6	0.6	*0.6*	0.6	*0.6*	**0.6**	0.6	*0.6*	**0.6**

Table 3: Mean confidence interval width in million years, i.e. the mean difference between the 95% confidence upper bounds provided by the three methods and the most recent fossil, on the simulated extinct taxa where the **S&S**, **McI** and **Alr** computations were feasible (plain text), on those belonging to a simulated tree where a global fossil recovery rate can be estimated (italics; only for methods **Glo**, **Int** and **Ref**) and on all the extinct taxa (bold; only for methods **Int** and **Ref**).

reference method, notably when the fossil recovery rate is low. Namely, it is about ten times greater when the fossil discovery rate is 0.05, and is still almost twice as large for a fossilization rate of 5 (Table 3). Note that, for the fossilization rates 1 and 5, the error percentages obtained with the method of Strauss and Sadler (1989) are not significantly lower than those obtained with methods **Ref** and **Int** (Table 2), despite the fact that its mean confidence interval widths are larger than those of these two methods. The same behavior is observed for methods **McI**, **Alr** and **Glo** for the fossilization rates 0.005 and 0.01. Though both methods **McI** and **Glo** have mean confidence interval widths very close to that of the reference for the two highest fossilization rates, only method **Glo** performs as well as methods **Ref** and **Int** for these fossilization rates, since the corresponding error percentages of method **McI** are higher than required.

Tables 2 and 3 show that method **Glo** performs systematically better than method **McI**. Let us recall that the only difference between these two approaches is the way in which the fossilization rate is estimated: branch by branch for **McI**, versus on the whole tree for **Glo**. This suggests that, in the case where the fossil process can be assumed to be homogeneous through the tree, a global estimation does improve the accuracy of extinction time estimation.

Method **Int** performs almost as well as the reference method **Ref** in the sense that both its error percentages and its mean confidence interval widths are very close to those of **Ref**, which are the best achievable under the assumptions used to simulate the data. The performance of method **Glo** is almost as good as that of methods **Int** and **Ref** only for the fossilization rates 1 and 5 (and quite correct with the fossilization rate 0.1). Roughly speaking, method **Glo** differs from method **Int** in the sense that it does not take into account the diversification process (both methods assume that the fossilization follows a Poisson process). The respective performances of **Int** and **Glo** suggest that taking into account the diversification in the extinction time estimation does matter in the case where the fossilization rate is lower than the diversification rates but may be not essential when the fossil recovery rate is high.

Appendix C displays the results of a simulation study assessing the 50% and 75% confidence upper bounds provided by all the methods considered here. We observe the same general behavior as with the 95% confidence upper bound.

Results for the empirical example: extinction times for ophiacodontids, edaphosaurids and sphenacodontids

The tree showing the probability density distributions of extinction times of the tips shows that many lineages of all three clades were probably extinct well before the Kungurian/Roadian boundary (by considering only the tree displayed in Fig. 3). The extinctions seem to be spread out throughout the Kungurian, rather than being concentrated at any given time. These individual extinction times were also computed by taking into account the hundred trees of our dataset and are shown in greater detail in Figures 4, 5, 6 and 7. This shows, unsurprisingly, that for nearly all taxa, the peak extinction time probability is well before the Kungurian/Roadian boundary, and that the 95% confidence interval does not reach that boundary, in most cases. One obvious exception is *Dimetrodon angelensis*, the most recent taxon included in this study (from the San Angelo Formation of the Pease River Group). Its peak density is near the Kungurian/Roadian boundary, and the tail of the distribution suggests that it could have

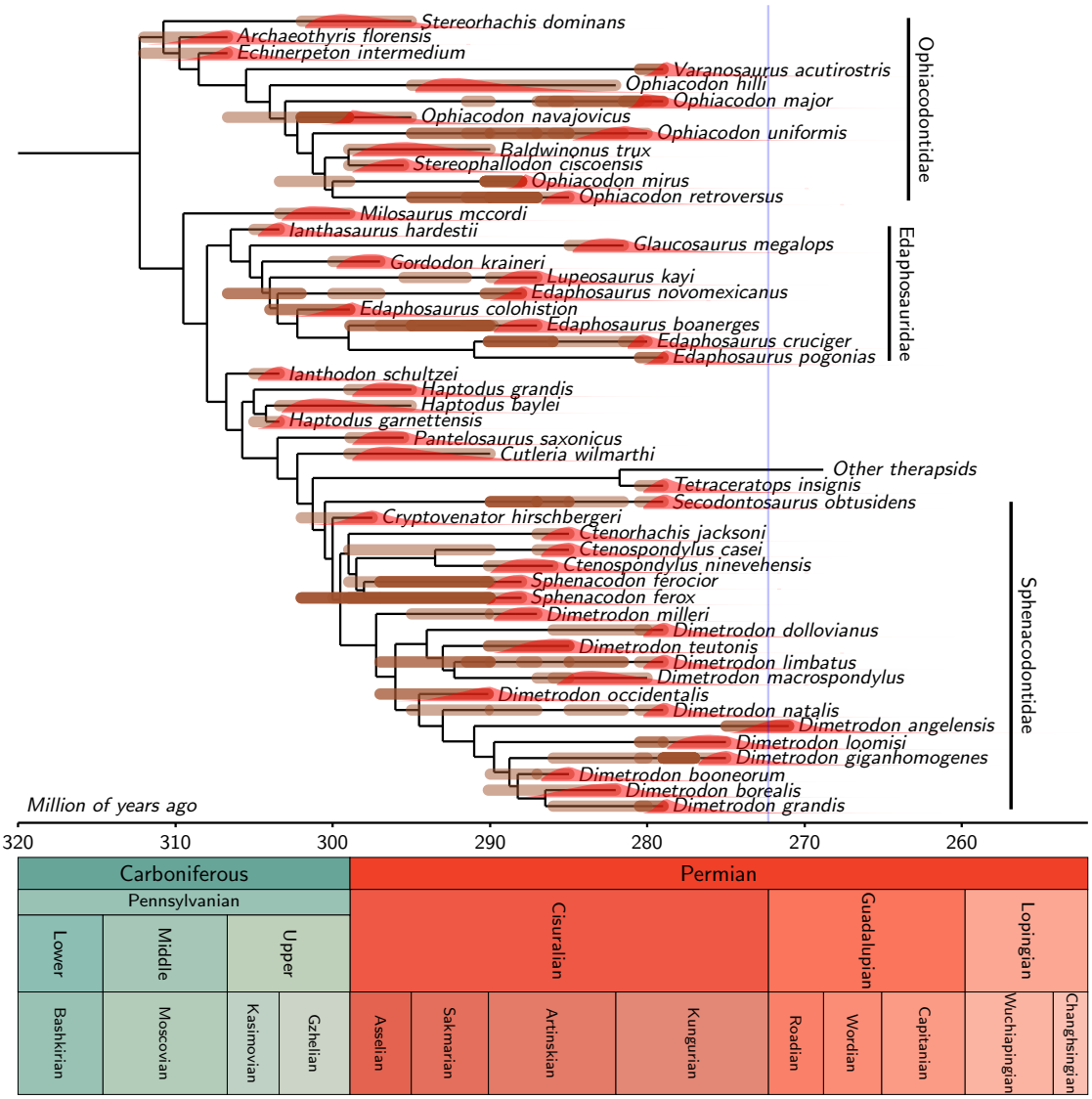

Figure 3: One of the 100 equally parsimonious trees of our dataset and the extinction time probability density distributions of its extinct taxa. Intervals of possible ages for each fossil are represented as thick brown segments with a certain level of transparency (darker brown segments correspond to overlapping intervals). The extinction time probability distributions are represented in red. These sometimes overlap the intervals of possible ages of the last fossils of a given branch, because fossil ages are sampled in these intervals to compute the distributions. The thin blue line represents the Kungurian/Roadian (Cisuralian/Guadalupian) boundary. Only the branch tips are time-calibrated; the position of nodes is set to the median of the distribution of the corresponding speciation time, as shown in Figure 10 (Appendix E).

become extinct well into the Guadalupian, although the posterior probability of the most recent fossil age of *D. angelensis*, biased towards the lower bounds of its interval, could be consistent with an older extinction time (Figure 12). Despite this exception, our first hypothesis is confirmed; most lineages of the three clades appear to have become extinct before the end of the Kungurian. These results provide strong support for the first of our three hypotheses.

Our results show that the taxa Ophiacodontidae and Edaphosauridae were probably extinct by the end of the Kungurian, with probabilities that these two clades persisted into the Roadian of only 9.1% and 8.0% for these two clades, respectively; this is also shown graphically by the fact that the 95% confidence interval on their extinction time only extends to the mid-Roadian, approximately (Fig. 8). Sphenacodontidae may well have become extinct later because the 95% confidence interval on its extinction time extends into the earliest Capitanian (Fig. 8). The probability that Sphenacodontidae extended into the Roadian is high, at 86.8%. In addition, the peak probability of its extinction time is near the mid-Roadian, significantly later than for Ophiacodontidae and Edaphosauridae, whose peak density of extinction is in the early to mid-Kungurian. Thus, our results provide weak support for our second hypothesis; although the probability that Ophiacodontidae and Edaphosauridae were extinct by the end of the Kungurian is high, the probability that Sphenacodontidae persisted into the Roadian appears fairly high. Note that these results consider the probable extinction times of unobserved lineages of these three clades, in addition to all those observed in the fossil record.

Contrary to previous suggestions (Lucas, 2018, p. 430), Ophiacodontidae does not appear to have become extinct before Edaphosauridae; in fact our computations suggest the reverse, but the difference in timing is not statistically significant (the probability that Ophiacodon became extinct after Edaphosauridae is 0.533), so these extinctions should provisionally be considered more or less simultaneous. However, the difference between extinction time of Sphenadocondidae and that of the other two clades is marginally significant; the probability that Sphenacodontidae became extinct after Ophiacodontidae is 0.953, and for Edaphosaurida, this is 0.958. Thus, we cannot confirm the third hypothesis; in fact, our results suggest that Ophiacodontidae may have become extinct slightly later than Edaphosauridae. However, one element of this hypothesis is confirmed: Sphenacodontidae does appear to have become extinct last.

The mean fossilization, speciation, and extinction rates are 0.2306, 0.1348 and 0.1352, respectively. These rates are slightly higher for speciation and extinction, but much higher for fossilization, than those reported for a more inclusive dataset of amniotes (Didier *et al.*, 2017). That study reported rates of fossilization, speciation, and extinction of 3.214 E-3, 9.594 E-2, and 9.488 E-2, respectively.

For all these results, topology appears to have only subtle effects on extinction ages, although the effect on nodal ages is slightly greater, which is to be expected given that topology and clade content are intimately linked.

Discussion

The simulation study shows that taking into account all lineages in a clade, as well as the diversification process rather than only the fossilization events of a single lineage leads to confidence upper bounds of extinction dates which are both tighter and more accurate than those obtained from previous approaches, notably that of Strauss and Sadler (1989), which was designed to assess the true local stratigraphic range of a taxon in a section. Moreover, our approach can determine confidence upper bounds of the extinction time of taxa that have left very few or even a single fossil, a situation that is problematic for the methods that deal with each branch independently.

Comparisons between extinction times of individual ophiacodontid, edaphosaurid and sphenacodontid species (Fig. 4, 5 and 6) suggest that lineages of these three clades became extinct at various times, which is more consistent with a long, low-intensity crisis than a brief, catastrophic event. These lineage-specific extinction times do not suggest a discrete extinction event associated with the Artinskian/Kungurian boundary, but they are compatible with (without proving) a mild, diffuse crisis spanning much of the Artinskian and Kungurian, as suggested by Benton (1985, 1989). These findings rather support the first of the three hypotheses that we test here.

The finding that one of three clades studied here (Sphenacodontidae) has a fairly high probability of having become extinct in the Guadalupian may seem surprising at first sight, given the few records (two occurrences of *Dimetrodon angelensis*) that may occur after the Kungurian (these may be either late Kungurian or early Roadian). Even for Edaphosauridae and Ophiacodontidae, an extinction in the early

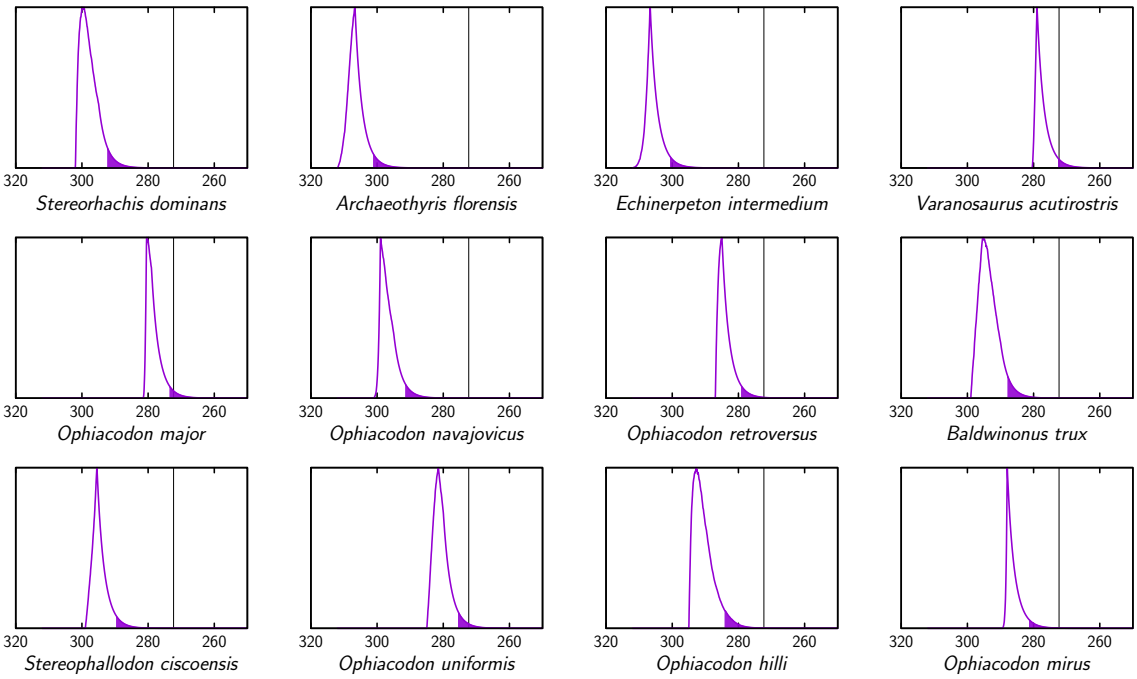

Figure 4: Extinction time probability density distributions of ophiacodontids (X axis, in Ma). The colored part under each distribution starts at its 95% confidence upper bound. The thin vertical lines represent the Kungurian/Roadian (Cisuralian/Guadalupian) boundary.

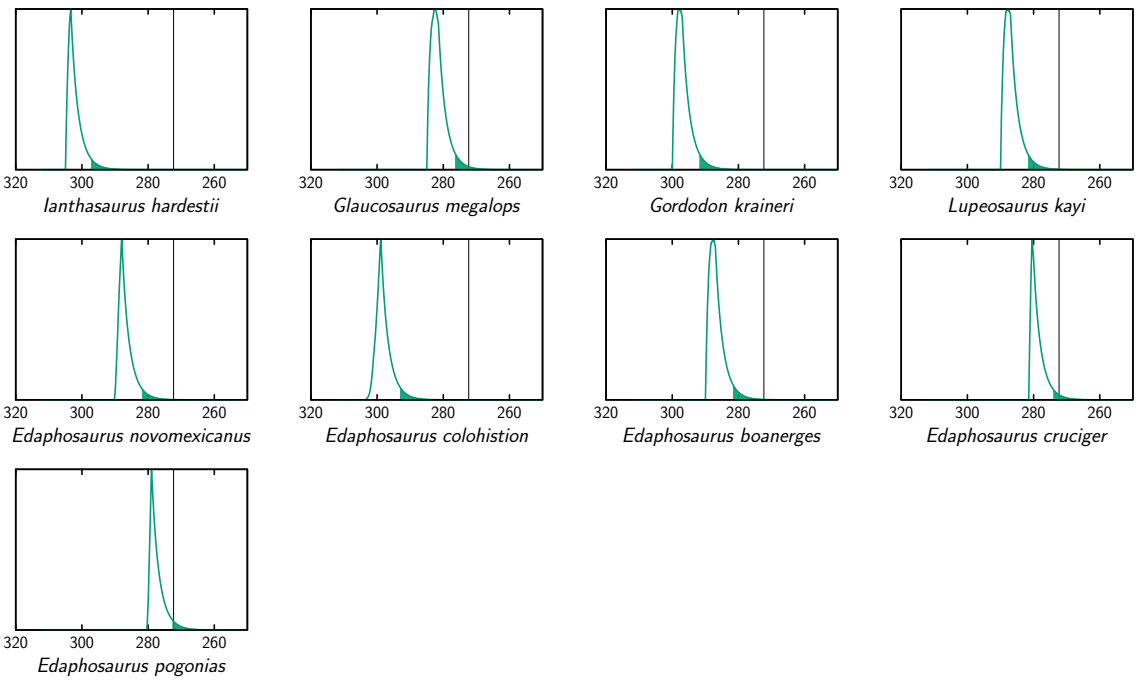

Figure 5: Extinction time probability density distributions of edaphosaurids (X axis, in Ma). The colored part under each distribution starts at its 95% confidence upper bound. The thin vertical lines represents the Kungurian/Roadian (Cisuralian/Guadalupian) boundary.

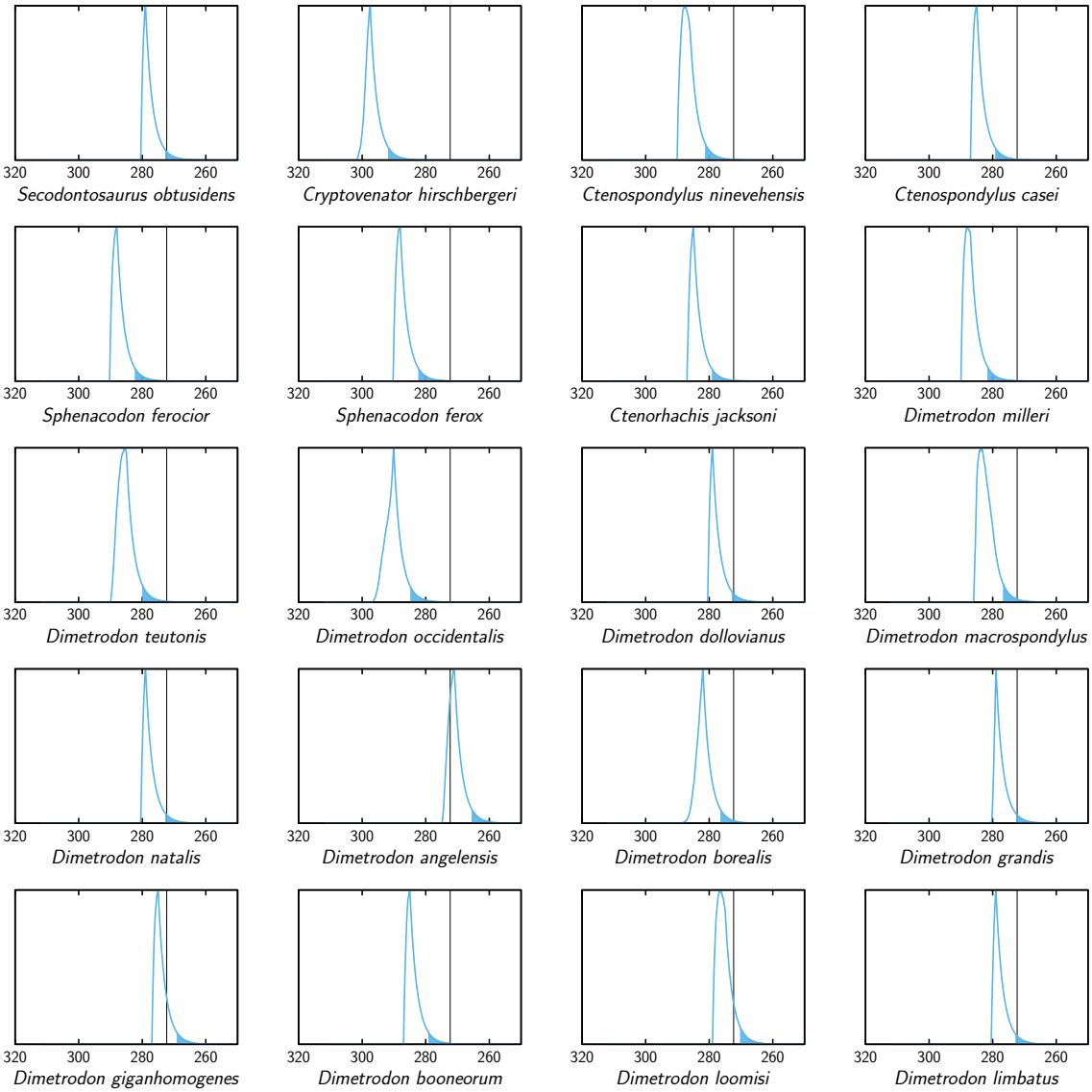

Figure 6: Extinction time probability density distributions of sphenacodontids (X axis, in Ma). The colored part under each distribution starts at its 95% confidence upper bound. The thin vertical lines represent the Kungurian/Roadian (Cisuralian/Guadalupian) boundary.

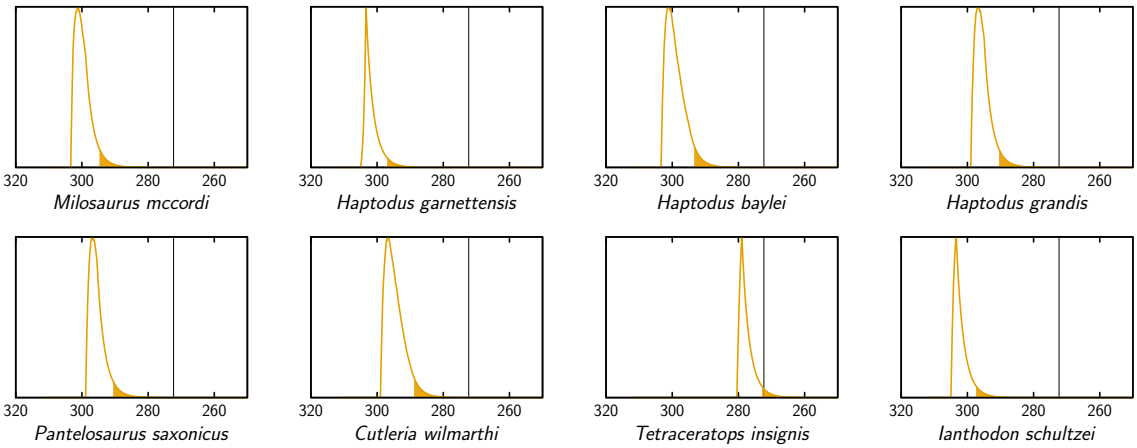

Figure 7: Extinction time probability density distributions of the taxa of our dataset that are not ophiacodontids, edaphosaurids or sphenacodontids (X axis, in Ma). The colored part under each distribution starts at its 95% confidence upper bound. The thin vertical lines represents the Kungurian/Roadian (Cisuralian/Guadalupian) boundary.

Roadian cannot be ruled out because the 95% confidence interval of their stratigraphic range extends into the mid-Roadian.

The extinction time of the most recent lineages of each clade and the extinction of the three large clades to which they belong (Fig. 8) reveals only small differences, except for Sphenacodontidae. This suggests that in these cases, the unobserved lineages (without a fossil record) of these clades did not add much to their stratigraphic extension toward the present. This probably reflects our increased sampling of fossiliferous horizons, which results in a much greater estimated fossilization rate. Didier *et al.* (2017, p. 981) estimated that "only about 14% of the Permo-Carboniferous eupelycosaur lineages (defined as an internode on the tree) have left a fossil record." Using the new rate estimates, this proportion increases to 46%, which implies far fewer missing lineages. It is no surprise that for Sphenacodontidae, unobserved lineages seem to have a greater impact on the extinction time probability density, given that this is the most speciose of the three clades. The peak probability of extinction of Sphenacodontidae in the late Roadian is congruent with the conclusion of Brocklehurst *et al.* (2013). We obtained these results despite being fairly conservative in our assessment of the age of the San Angelo Formation; the range of possible ages that we assigned to it (from 275 to 271 Ma) extends from late Kungurian to early Roadian, but about two-thirds of this interval is in the Kungurian, and only a third in the Roadian.

Our interpretation of these results depends on the quality of the fossil record, which has to be of sufficient quality, on a worldwide basis (records in one location can compensate to an extent gaps in another) to meet the assumptions of our method, and incomplete enough to be compatible with the persistence of unobserved taxa for several million years. In other words, if the fossil record were extremely patchy, no method, no matter how sophisticated, would be able to extract well-constrained, reliable extinction dates from it. Conversely, if the fossil record were nearly complete, there would be no need for analyses because the history of taxa could be directly read in rocks. For studies such as ours (and those cited above) to make sense, an incomplete but not hopeless fossil record is required.

Is the fossil record sufficient to apply our method? Two kinds of gaps (temporal and geographic) can be problematic. In south-western North America, where synapsids have an excellent fossil record in the Cisuralian (as shown by our data), the continental fossil record of synapsids is poor in the Roadian, where it is restricted to the Chickasha and Flowerpot formations of Oklahoma, and probably, by the San Angelo Formation in Texas. It has even been claimed that there is no Roadian fossil record of synapsids in North America because these formations have been argued to be Kungurian (Lucas, 2018), but our literature review shows otherwise (Clapham, 1970; Reisz and Laurin, 2001, 2002; Nelson and Hook, 2005; Foster *et al.*, 2014; Soreghan *et al.*, 2015). In any case, the North American strata that yielded Permian synapsids were deposited close to the paleo-equator. Recent studies show that there is a Roadian synapsid fossil record in Russia, given that the Kazanian is equivalent to the Roadian

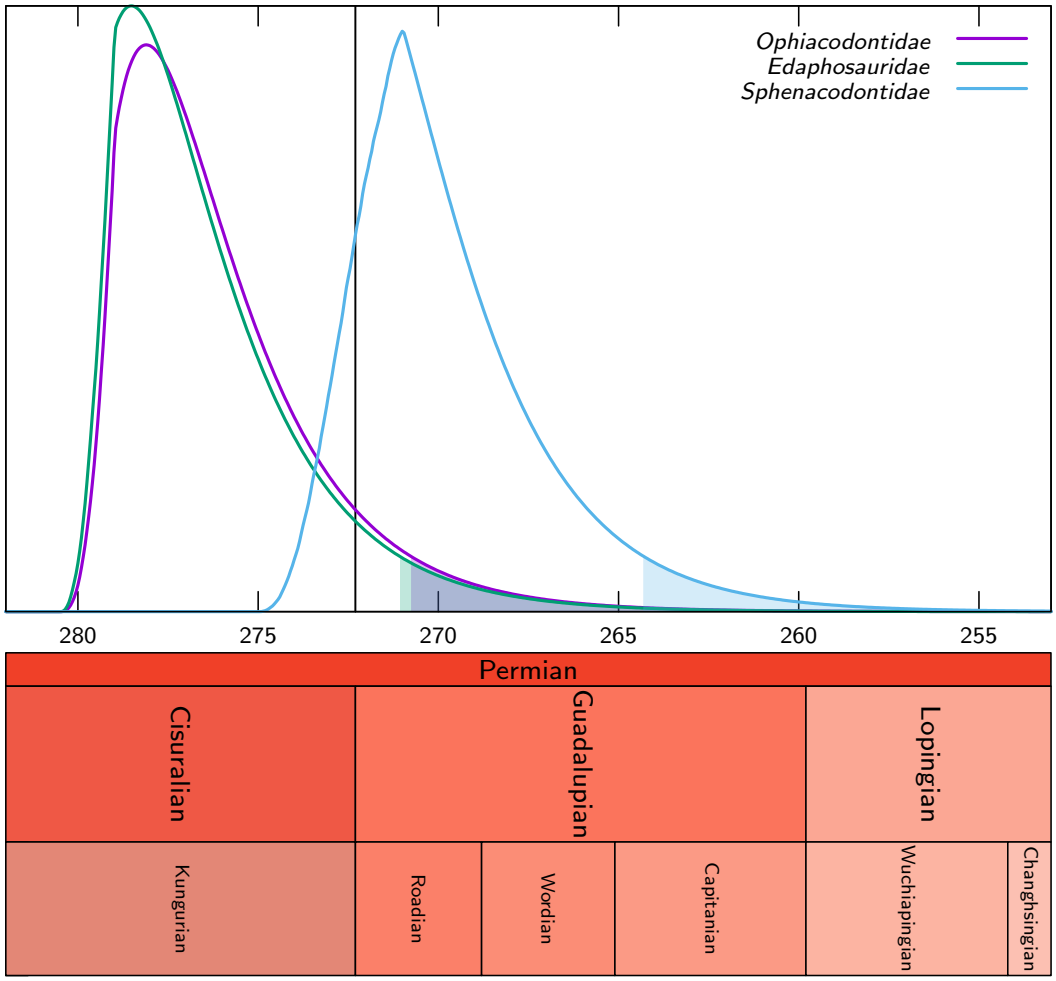

Figure 8: Probability density distributions of the extinction times (X axis, in Ma) of the clades Ophiacodontidae, Edaphosauridae and and Sphenacodontidae based on our dataset. The colored part under each distribution starts at its 95% confidence upper bound. The vertical line represents the Kungurian/Roadian (Cisuralian/Guadalupian) boundary.

(Davydov *et al.*, 2018). From the Wordian on, the synapsid fossil record is more widespread and is especially dense in the Karoo basin in South Africa (Rubidge and Day, 2020), in addition to Russia. These more recent strata were deposited in higher paleo-latitudes of about 25-30 degrees North for the Russian European Platform, and nearly 60 degrees South for the Karoo Basin (Schneider *et al.*, 2020). This spatio-temporally disjunct synapsid fossil record has previously been commented by Brocklehurst *et al.* (2013).

To what extent can these geographic gaps in the synapsid fossil record invalidate our analyses? If permo-carboniferous synapsids were restricted to North America, where their fossil record stops in the Roadian (and probably, early in the Roadian), this would indeed be problematic. However, caseids extend into the Roadian (Maddin *et al.*, 2008; Golubev, 2015) and varanopids survived until the Capitanian (Modesto *et al.*, 2011), and both are documented by a sparse fossil record after the Cisuralian, and mostly outside North America. Even in the Cisuralian, caseids and varanopids have a poorer fossil record than ophiacodontids, edaphosaurids and sphenacodontids. This was noticed long ago by Olson (1965, 1968, 1975), who postulated that these taxa lived in a somewhat more upland environment than other Permo-Carboniferous synapsids. This hypothesized difference in habitat still seems plausible (Angielczyk and Kammerer, 2018, p. 130), even though Lambertz *et al.* (2016) postulated an aquatic lifestyle for at least the geologically most recent caseids, which have highly cancellous bone. The fossil record is inherently biased in favor of aquatic taxa (Shipman, 1981) because most fossiliferous deposits were deposited under water (Brett, 1995). Thus, the presence of a fossil record of caseids and varanopids in the Guadalupian suggests that the absence of ophiacodontids, edaphosaurids and sphenacodontids from that fossil record, despite their inferred habitat (apparently closer to the water than the areas inhabited by caseids and varanopids), is real, rather than a taphonomic artefact. This conclusion is also supported by the fact that ophiacodontids, edaphosaurids and sphenacodontids are also known from Europe (Berman *et al.*, 2001), where amniotes have a Guadalupian and Lopingian fossil record (Schneider *et al.*, 2020). This conclusion is further supported by the association of various combinations of these taxa (caseids and/or varanopids from Olson's caseid chronofauna, with ophiacodontids, edaphosaurids and/or sphenacodontids, from another chronofauna that apparently inhabited more low-land environments) in some early Permian localities, such as El Cobre Canyon (Berman *et al.*, 2015), the Archer City Bonebed (Sander, 1989; Reisz *et al.*, 2010), Fort Sill (MacDougall *et al.*, 2017) and Bromacker (Berman *et al.*, 2014). Thus, the absence of ophiacodontids, edaphosaurids and sphenacodontids from Guadalupian strata of Russia and South Africa suggests that they became extinct no later than in the Guadalupian (early in the Guadalupian, in the case of ophiacodontids and edaphosaurids).

The temporal heterogeneities in the quality of the fossil record appear to be less problematic, for Paleozoic synapsids, than the geographic bias, to the extent that there does not appear to be a trend of decreasing record. The Roadian appears to be a time of relatively poor fossil record of amniotes, but this record improves strongly in the Wordian (Olroyd and Sidor, 2017), which suggests that there is no decreasing trend in the quality of the fossil record of Permian synapsids. The assumption of a homogeneous fossilization rate through time, which is assumed by most methods (including ours) to study extinction events, does not imply that an approximately constant number of synapsid taxa (or specimens) should be known in each geological stage because the clade diversified through time and was affected by extinction events. We share the reservations expressed by Benton *et al.* (2011) about using the number of amniote-bearing localities to assess fluctuations in the quality of the synapsid fossil record through time (e.g., Brocklehurst *et al.*, 2013) because synapsids form half of amniotes from a cladistic perspective, so amniote diversification should be strongly correlated to synapsid diversification over time. This is problematic because Marjanović and Laurin (2008) showed that a simple exponential diversification model of lissamphibians best explained the variations in their fossil record, whereas the area of rock exposures of various ages played a negligible role. It is thus not surprising that Brocklehurst *et al.* (2013, p. 486) found that this method (which they called "residual diversity") conflicted with their two other metrics (species counts and phylogeny-corrected counts) in suggesting that synapsid diversidy had decreased between the Wordian and the Capitanian, when other methods indicate a climax in synapsid diversity. To these problems in assessing biases in our assessment of paleobiodiversity, we must add the anthropogenic bias, which consists in the uneven effort that has been spent looking for fossils in strata of various ages, as pointed out by Brocklehurst *et al.* (2013, p. 481). Thus, assess objectively the quality of the fossil record remains a challenging problem, especially in the terms relevant to our method, namely the probability of discovery of fossiliferous horizons per lineage per million years of those lineages that existed (most of which are probably not known).

Is the Roadian synapsid fossil record incomplete enough to be compatible with the persistence of unobserved taxa (Ophiacodontidae, Edaphosauridae and Sphenacodontidae) for a few million years? This seems likely. The recent discovery of a late Permian diadectomorph from China (Liu and Bever, 2015) serves as a reminder that the fossil record can still yield surprising discoveries that refute a long-established consensus about the stratigraphic range of taxa; this taxon was previously thought to have become extinct toward the end of the early Permian (Laurin, 2015). The recent discovery of two or three dinocephalian skulls that extend the stratigraphic distribution of this taxon (previously considered typical of the *Tapinocephalus* AZ) into the lowermost Poortjie Member of the Teekloof Formation also shows that even in a densely-studied, highly fossiliferous basin such as the Karoo, the time of extinction of a higher taxon (Dinocephalia is typically ranked as a suborder and considered to include three families) is not immune to revision (Day *et al.*, 2015b). Our method could be used to assess the probability and magnitude of potential stratigraphic range extensions of taxa of various sizes, from single lineages to large clades, in addition to providing an additional statistical tool to better assess the evolution of biodiversity over time.

Our findings improve our understanding of the replacement of Permo-Carboniferous synapsids by therapsids in the Guadalupian. Therapsids must have originated in the late Carboniferous, as implied by their sister-group relationships with sphenacodontids (Reisz, 1986; Gauthier *et al.*, 1988; Sidor, 2001; Benson, 2012; Didier and Laurin, 2020) and by our new dating of the eupelycosaur evolutionary radiation (Fig. 10), and more specifically of the sphenacodontid/therapsid divergence, which probably took place in the Gzhelian (Fig. 11). Angielczyk and Kammerer (2018) suggested that this event took place even earlier, in the Kasimovian, but this depends on the poorly constrained age of the Sangre de Cristo Formation in Colorado. The formation by the same name in New Mexico is from another basin, but a recent study concluded that it is of early Permian age (Lucas *et al.*, 2015). Thus, there is currently no strong evidence that the sphenacodontid/therapsid divergence is older than Gzhelian. Yet, therapsids are unknown so far in the Carboniferous, with the possible exception of the very fragmentary remains (a string of a few vertebrae) from the Moscovian of Nova Scotia that Spindler (2014) interpreted as a therapsid, an interpretation that seems tenuous at best; the original interpretation that these belong to a sphenacodontid seems plausible (Reisz, 1972). Therapsids may be represented by a single specimen of *Tetraceratops* in the Kungurian, even though its affinities are still debated (Amson and Laurin, 2011; Spindler, 2020), and beyond the scope of this study. The first undisputed therapsids (dinocephalians from the Goluysherma Assemblage of the Russian Platform) date from the Roadian (Ivakhnenko, 2003; Golubev, 2015), but they become abundant only toward the end of the Roadian (Brocklehurst *et al.*, 2013, p. 487) or in the early Wordian. Lozovsky (2005, p. 182) reported that therapsids barely make up about 5% of the Mezen (Roadian) fauna and that the obvious domination of therapsids only occurred in the Tatarian (which started in the Wordian), as preserved in the Ocher and Isheevo localities. According to Golubev (2015, fig. 2), the Ocher subassemblage is early Wordian and Isheevo extends from the late Wordian through the mid-Capitanian. Also, some fragmentary Roadian caseid remains had been misinterpreted as therapsids (Brocklehurst and Fröbisch, 2017). Thus, in the early Roadian, the replacement of Permo-Carboniferous clades of synapsids by therapsids had apparently barely started (Brocklehurst *et al.*, 2013). The dynamics of biodiversity changes in synapsids similarly shows that therapsids surpassed other synapsid clades in biodiversity sometime in the Roadian (Brocklehurst *et al.*, 2013, fig. 1). All this suggests a very slow initial therapsid diversification (but this could be tested more rigorously with the method presented here), in contrast with the explosive diversification model postulated by Kemp (2009). But was the replacement of Permo-Carboniferous synapsids by therapsids competitive or not? Our findings shed new light on this question.

The extinction of Ophiacodontidae and Edaphosauridae near the Kungurian/Roadian boundary supports the hypothesis, recently proposed by Olroyd and Sidor (2017), that the replacement of Permo-Carboniferous synapsids by therapsids was non-competitive because the absence of definite therapsids in the Kungurian fossil record suggests that they remained rare at that time, and at least locally. It is thus difficult to envision that they played a major role in the extinction of Ophiacodontidae and Edaphosauridae. The slightly later extinction date of Sphenacodontidae gives a more ambiguous signal on this point. They may also have died out before therapsids had become major competitors, as a literal reading of the Roadian fossil record suggests, but this is less certain. It is conceivable that the replacement of Permo-Carboniferous synapsids by therapsids was partly competitive (for sphenacodontids, and more likely, for caseids and varanopids, which are known to extend at least into the Roadian and Capitanian, respectively), and partly opportunistic, by filling the ecological vacuum left by the extinction of ophiacodontids

and edaphosaurids. A dating analysis of the therapsid evolutionary radiation through the FBD would be helpful to assess this. Nevertheless, our findings suggest that the exinction peak in Permo-Carboniferous clades (Olson's extinction) was near the Kungurian/Roadian boundary, rather than in the Roadian and Wordian, as suggested by Sahney and Benton (2008, p. 760).

Acknowledgments

We thank Robert W. Hook (U. of Texas at Austin) for a wealth of information and discussions about biostratigraphy and lithostratigraphy, especially for the El Reno and Pease River groups, and for putting one of us (ML) in touch with other specialists of Permian stratigraphy. We also thank Gregory P. Wahlman (Wahlman Geological Services) and John R. Groves (Technology Center, Pittsburgh, PA) for information about Permian fusulinids. This draft was improved by comments from Neil Brocklehurst, Charles R. Marshall and Ken Angielczyk. Manon Thomazo helped with the compilation of character and state names in the Mesquite Nexus file (see supplements) and corrected some typographic and stylistic errors in parts of the draft.

References

Abdala, F., Rubidge, B. S., and Van Den Heever, J. (2008). The oldest therocephalians (Therapsida, Eutheriodontia) and the early diversification of Therapsida. *Palaeontology*, **51**(4), 1011–1024.

Alroy, J. (1996). Constant extinction, constrained diversification, and uncoordinated stasis in North American mammals. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **127**(1-4), 285–311.

Alroy, J. (2014). A simple bayesian method of inferring extinction. *Paleobiology*, **40**(4), 584–607.

Alroy, J. (2016). A simple bayesian method of inferring extinction: reply. *Ecology*, **97**(3), 798–800.

Alroy, J., Marshall, C., and Miller, A. (2012). Paleobiology database. `https://paleobiodb.org/#/`.

Amson, E. and Laurin, M. (2011). On the affinities of *Tetraceratops insignis*, an Early Permian synapsid. *Acta Palaeontologica Polonica*, **56**(2), 301–312.

Angielczyk, K. D. and Kammerer, C. F. (2018). 5. Non-mammalian synapsids: the deep roots of the mammalian family tree. In *Mammalian evolution, diversity and systematics*, pages 117–198. De Gruyter.

Arens, N. C. and West, I. D. (2008). Press-pulse: a general theory of mass extinction? *Paleobiology*, **34**(4), 456–471.

Ascarrunz, E., Sánchez-Villagra, M. R., Betancur-R, R., and Laurin, M. (2019). On trends and patterns in macroevolution: Williston's law and the branchiostegal series of extant and extinct osteichthyans. *BMC Evolutionary Biology*, **19**(1), 117.

Axelrod, D. and Bailey, H. (1968). Cretaceous dinosaur extinction. *Evolution*, **22**(3), 595–611.

Barnosky, A. D., Matzke, N., Tomiya, S., Wogan, G. O., Swartz, B., Quental, T. B., Marshall, C., McGuire, J. L., Lindsey, E. L., Maguire, K. C., *et al.* (2011). Has the Earth's sixth mass extinction already arrived? *Nature*, **471**(7336), 51–57.

Basu, A. R., Petaev, M. I., Poreda, R. J., Jacobsen, S. B., and Becker, L. (2003). Chondritic meteorite fragments associated with the Permian-Triassic boundary in Antarctica. *Science*, **302**(5649), 1388–1392.

Benson, R. B. (2012). Interrelationships of basal synapsids: cranial and postcranial morphological partitions suggest different topologies. *Journal of Systematic Palaeontology*, **10**(4), 601–624.

Benton, M. J. (1985). Mass extinction among non-marine tetrapods. *Nature*, **316**(6031), 811–814.

Benton, M. J. (1989). Mass extinctions among tetrapods and the quality of the fossil record. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, **325**(1228), 369–386.

Benton, M. J. (2003). *When life nearly died: the greatest mass extinction of all time*. Thames & Hudson London.

Benton, M. J., Dunhill, A. M., Lloyd, G. T., and Marx, F. G. (2011). Assessing the quality of the fossil record: insights from vertebrates. *Geological Society, London, Special Publications*, **358**(1), 63–94.

Berman, D. S., Reisz, R. R., Martens, T., and Henrici, A. C. (2001). A new species of *Dimetrodon* (Synapsida: Sphenacodontidae) from the Lower Permian of Germany records first occurrence of genus outside of North America. *Canadian Journal of Earth Sciences*, **38**(5), 803–812.

Berman, D. S., Henrici, A. C., Sumida, S. S., Martens, T., and Pelletier, V. (2014). First european record of a varanodontine (Synapsida: Varanopidae): Member of a unique early Permian upland paleoecosystem, Tambach Basin, central Germany. In *Early evolutionary history of the Synapsida*, pages 69–86. Springer.

Berman, D. S., Henrici, A., and Lucas, S. G. (2015). Pennsylvanian-permian red bed vertebrate localities of new mexico and their assemblages. *New Mexico Museum of Natural History and Science Bulletin*, **68**, 65–76.

Bertrand, Y., Pleijel, F., and Rouse, G. W. (2006). Taxonomic surrogacy in biodiversity assessments, and the meaning of Linnaean ranks. *Systematics and Biodiversity*, **4**(2), 149–159.

Bond, D., Wignall, P., Wang, W., Izon, G., Jiang, H.-S., Lai, X.-L., Sun, Y.-D., Newton, R., Shao, L.-Y., Védrine, S., *et al.* (2010). The mid-Capitanian (Middle Permian) mass extinction and carbon isotope record of South China. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **292**(1), 282–294.

Bradshaw, C., Cooper, A., Turney, C., and Brook, B. (2012). Robust estimates of extinction time in the geological record. *Quaternary Science Reviews*, **33**, 14–19.

Brayard, A., Escarguel, G., Bucher, H., Monnet, C., Brühwiler, T., Goudemand, N., Galfetti, T., and Guex, J. (2009). Good genes and good luck: ammonoid diversity and the end-Permian mass extinction. *Science*, **325**(5944), 1118–1121.

Brett, C. E. (1995). Sequence stratigraphy, biostratigraphy, and taphonomy in shallow marine environments. *Palaios*, **10**(6), 597–616.

Brocklehurst, N. (2017). Rates of morphological evolution in Captorhinidae: an adaptive radiation of Permian herbivores. *PeerJ*, **5**, e3200.

Brocklehurst, N. (2018). An examination of the impact of Olson's extinction on tetrapods from Texas. *PeerJ*, **6**, e4767.

Brocklehurst, N. (2020). Olson's Gap or Olson's Extinction? a Bayesian tip-dating approach to resolving stratigraphic uncertainty. *Proceedings of the Royal Society B*, **287**(1928), 20200154.

Brocklehurst, N. and Brink, K. S. (2017). Selection towards larger body size in both herbivorous and carnivorous synapsids during the Carboniferous. *Facets*, **2**(1), 68–84.

Brocklehurst, N. and Fröbisch, J. (2017). A re-examination of the enigmatic Russian tetrapod *Phreatophasma aenigmaticum* and its evolutionary implications. *Fossil Record*, **20**(1), 87–93.

Brocklehurst, N. and Fröbisch, J. (2018). A reexamination of *Milosaurus mccordi*, and the evolution of large body size in Carboniferous synapsids. *Journal of Vertebrate Paleontology*, **38**(5), e1508026.

Brocklehurst, N., Kammerer, C. F., and Fröbisch, J. (2013). The early evolution of synapsids, and the influence of sampling on their fossil record. *Paleobiology*, **39**(3), 470–490.

Burgess, S. D., Bowring, S., and Shen, S.-z. (2014). High-precision timeline for Earth's most severe extinction. *Proceedings of the National Academy of Sciences*, **111**(9), 3316–3321.

Carroll, R. L. (1988). *Vertebrate Paleontology and Evolution*. W. H. Freeman, New York.

Ceballos, G., Ehrlich, P. R., Barnosky, A. D., García, A., Pringle, R. M., and Palmer, T. M. (2015). Accelerated modern human-induced species losses: Entering the sixth mass extinction. *Science advances*, **1**(5), e1400253.

Ceballos, G., Ehrlich, P. R., and Dirzo, R. (2017). Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *Proceedings of the national academy of sciences*, **114**(30), E6089–E6096.

Cisneros, J. C., Kammerer, C. F., Angielczyk, K. D., Fröbisch, J., Marsicano, C., Smith, R. M., and Richter, M. (2020). A new reptile from the lower Permian of Brazil (*Karutia fortunata* gen. et sp. nov.) and the interrelationships of Parareptilia. *Journal of Systematic Palaeontology*, pages 1–23.

Clapham, W. (1970). Permian miospores from the Flowerpot Formation of western Oklahoma. *Micropaleontology*, **16**(1), 15–36.

Cooper, A. and Penny, D. (1997). Mass survival of birds across the Cretaceous-Tertiary boundary: molecular evidence. *Science*, **275**(5303), 1109–1113.

Davydov, V., Biakov, A., Schmitz, M., and Silantiev, V. (2018). Radioisotopic calibration of the Guadalupian (middle Permian) series: Review and updates. *Earth-Science Reviews*, **176**, 222–240.

Day, M., Rubidge, B., Jirah, S., and Almond, J. (2013). Biostratigraphic correlation in the Karoo: the case of the Middle Permian parareptile *Eunotosaurus*. *South African Journal of Science*, **109**(3), 1–4.

Day, M. O., Ramezani, J., Bowring, S. A., Sadler, P. M., Erwin, D. H., Abdala, F., and Rubidge, B. S. (2015a). When and how did the terrestrial mid-Permian mass extinction occur? Evidence from the tetrapod record of the Karoo Basin, South Africa. *Proceedings of the Royal Society B: Biological Sciences*, **282**(1811), 20150834.

Day, M. O., Güven, S., Abdala, F., Jirah, S., Rubidge, B., and Almond, J. (2015b). Youngest dinocephalian fossils extend the *Tapinocephalus* Zone, Karoo Basin, South Africa. *South African Journal of Science*, **111**(3-4), 1–5.

Didier, G. and Laurin, M. (2020). Exact Distribution of Divergence Times from Fossil Ages and Tree Topologies. *Systematic Biology*, **69**(6), 1068–1087.

Didier, G., Royer-Carenzi, M., and Laurin, M. (2012). The reconstructed evolutionary process with the fossil record. *Journal of Theoretical Biology*, **315**(0), 26–37.

Didier, G., Fau, M., and Laurin, M. (2017). Likelihood of Tree Topologies with Fossils and Diversification Rate Estimation. *Systematic Biology*, **66**(6), 964–987.

Drummond, A. J., Ho, S. Y. W., Phillips, M. J., and Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS biology*, **4**(5), e88.

Ereshefsky, M. (2002). Linnaean Ranks: Vestiges of a Bygone Era. *Proceedings of the Philosophy of Science Association*, **2002**(3), 305–315.

Erwin, D. (1990). The end-Permian mass extinction. *Annual Review of Ecology and Systematics*, **21**(1), 69–91.

Faith, D. P. (1992). Conservation evaluation and phylogenetic diversity. *Biological Conservation*, **61**(1), 1–10.

Fara, E. (2004). Estimating minimum global species diversity for groups with a poor fossil record: a case study of Late Jurassic–Eocene lissamphibians. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **207**(1-2), 59–82.

Foote, M. (2003). Origination and extinction through the phanerozoic: a new approach. *The Journal of Geology*, **111**(2), 125–148.

Foster, T. M., Soreghan, G. S., Soreghan, M. J., Benison, K. C., and Elmore, R. D. (2014). Climatic and paleogeographic significance of eolian sediment in the Middle Permian Dog Creek Shale (Midcontinent US). *Palaeogeography, Palaeoclimatology, Palaeoecology*, **402**, 12–29.

Gauthier, J., Kluge, A. G., and Rowe, T. (1988). Amniote phylogeny and the importance of fossils. *Cladistics*, **4**(2), 105–209.

Gavryushkina, A., Heath, T. A., Ksepka, D. T., Stadler, T., Welch, D., and Drummond, A. J. (2016). Bayesian Total-Evidence Dating Reveals the Recent Crown Radiation of Penguins. *Systematic Biology*, **66**(1), 57–73.

Geraads, D., Didier, G., Barr, W. A., Reed, D., and Laurin, M. (2020). The fossil record of camelids demonstrates a late divergence between Bactrian camel and dromedary. *Acta Palaeontologica Polonica*, **65**(2), 251–260.

Golubev, V. (2015). Dinocephalian stage in the history of the Permian tetrapod fauna of Eastern Europe. *Paleontological Journal*, **49**(12), 1346–1352.

Hallam, A. (1989). The case for sea-level change as a dominant causal factor in mass extinction of marine invertebrates. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, **325**(1228), 437–455.

Heath, T. A., Huelsenbeck, J. P., and 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.

Ivakhnenko, M. F. (2003). Eotherapsids from the east European placket (Late Permian). *Paleontological Journal*, **37**(SUPPL. 4).

Jouve, S., Mennecart, B., Douteau, J., and Jalil, N.-E. (2017). Biases in the study of relationships between biodiversity dynamics and fluctuation of environmental conditions. *Palaeontologia Electronica*, **20**(1.18), 1–21.

Kemp, T. S. (2009). Phylogenetic interrelationships and pattern of evolution of the therapsids: testing for polytomy. *Palaeontologica Africana*, **44**, 1–12.

King, B. and Beck, R. M. (2020). Tip dating supports novel resolutions of controversial relationships among early mammals. *Proceedings of the Royal Society B*, **287**(1928), 20200943.

Lambertz, M., Shelton, C. D., Spindler, F., and Perry, S. F. (2016). A caseian point for the evolution of a diaphragm homologue among the earliest synapsids. *Annals of the New York Academy of Sciences*, **1385**(1), 3–20.

Laurin, M. (2010). The subjective nature of Linnaean categories and its impact in evolutionary biology and biodiversity studies. *Contributions to Zoology*, **79**(4), 131–146.

Laurin, M. (2012). Recent progress in paleontological methods for dating the Tree of Life. *Frontiers in genetics*, **3**(130).

Laurin, M. (2015). Stratigraphic range, identification of fragmentary fossils, and the Permo/Triassic crisis. *Comptes Rendus Palevol*, **5**(14), 343–345.

Lewin, R. (1983). Extinctions and the History of Life: Now that, for many at least, asteroid impact has been accepted as a causative agent in mass extinction, attention turns to the wider view. *Science (New York, NY)*, **221**(4614), 935.

Liu, F., Peng, H., Bomfleur, B., Kerp, H., Zhu, H., and Shen, S. (2020). Palynology and vegetation dynamics across the Permian–Triassic boundary in southern Tibet. *Earth-Science Reviews*, page 103278.

Liu, J. and Bever, G. (2015). The last diadectomorph sheds light on Late Palaeozoic tetrapod biogeography. *Biology letters*, **11**(5), 20150100.

Lozovsky, V. R. (2005). Olson's gap or Olson's bridge, that is the question. *New Mexico Museum of Natural History and Science Bulletin*, **30**, 179–184.

Lu, P. J., Yogo, M., and Marshall, C. R. (2006). Phanerozoic marine biodiversity dynamics in light of the incompleteness of the fossil record. *Proceedings of the National Academy of Sciences*, **103**(8), 2736–2739.

Lucas, S. G. (2017). Permian tetrapod extinction events. *Earth-Science Reviews*, **170**, 31–60.

Lucas, S. G. (2018). Permian tetrapod biochronology, correlation and evolutionary events. *Geological Society, London, Special Publications*, **450**(1), 405–444.

Lucas, S. G. and Shen, S.-Z. (2018). The Permian chronostratigraphic scale: history, status and prospectus. *Geological Society, London, Special Publications*, **450**(1), 21–50.

Lucas, S. G., Krainer, K., Dimichele, W. A., Voigt, S., Berman, D. S., Henrici, A. C., Tanner, L. H., Chaney, D. S., Elrick, S. D., Nelson, W. J., *et al.* (2015). Lithostratigraphy, biostratigraphy and sedimentology of the Upper Paleozoic Sangre de Cristo Formation, southwestern San Miguel county, New Mexico. In *New Mexico Geological Society, 66th Annual Field Conference, 2015*. New Mexico Geological Society.

MacDougall, M. J., Tabor, N. J., Woodhead, J., Daoust, A. R., and Reisz, R. R. (2017). The unique preservational environment of the Early Permian (Cisuralian) fossiliferous cave deposits of the Richards Spur locality, Oklahoma. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **475**, 1–11.

MacLeod, N. (1996). Testing patterns of Cretaceous-Tertiary planktonic foraminiferal extinction at El Kef (Tunisia). *Geological Society of America Special Papers*, **307**, 287–302.

Maddin, H. C., Sidor, C. A., and Reisz, R. R. (2008). Cranial anatomy of *Ennatosaurus tecton* (Synapsida: Caseidae) from the Middle Permian of Russia and the evolutionary relationships of Caseidae. *Journal of Vertebrate Paleontology*, **28**(1), 160–180.

Mann, A. and Paterson, R. S. (2020). Cranial osteology and systematics of the enigmatic early 'sail-backed'synapsid *Echinerpeton intermedium* Reisz, 1972, and a review of the earliest 'pelycosaurs'. *Journal of Systematic Palaeontology*, **18**(6), 529–539.

Marjanović, D. and Laurin, M. (2008). Assessing confidence intervals for stratigraphic ranges of higher taxa: the case of Lissamphibia. *Acta Palaeontologica Polonica*, **53**(3), 413–432.

Marshall, C. R. (1990). Confidence Intervals on Stratigraphic Ranges. *Paleobiology*, **16**(1), 1–10.

Marshall, C. R. (1994). Confidence Intervals on Stratigraphic Ranges: Partial Relaxation of the Assumption of Randomly Distributed Fossil Horizons. *Paleobiology*, **20**(4), 459–469.

Marshall, C. R. (1997). Confidence Intervals on Stratigraphic Ranges with Nonrandom Distributions of Fossil Horizons. *Paleobiology*, **23**(2), 165–173.

Marshall, C. R. (2008). A simple method for bracketing absolute divergence times on molecular phylogenies using multiple fossil calibration points. *The American Naturalist*, **171**(6), 726–742.

Marshall, C. R. (2019). Using the Fossil Record to Evaluate Timetree Timescales. *Frontiers in Genetics*, **10**, 1049.

Martin, J. E., Amiot, R., Lécuyer, C., and Benton, M. J. (2014). Sea surface temperature contributes to marine crocodylomorph evolution. *Nature communications*, **5**(1), 1–7.

McInerny, G. J., Roberts, D. L., Davy, A. J., and Cribb, P. J. (2006). Significance of sighting rate in inferring extinction and threat. *Conservation Biology*, **20**(2), 562–567.

Modesto, S. P., Smith, R. M., Campione, N. E., and Reisz, R. R. (2011). The last "pelycosaur" a varanopid synapsid from the *Pristerognathus* Assemblage Zone, Middle Permian of South Africa. *Naturwissenschaften*, **98**(12), 1027–1034.

Nelson, W. J. and Hook, R. W. (2005). Pease River Group (Leonardian-Guadalupian) of Texas: An overview. In S. G. Lucas and K. E. Zeigler, editors, *The Nonmarine Permian*, Bulletin 30, pages 243–250. New Mexico Museum of Natural History and Science.

Nelson, W. J., Hook, R. W., and Chaney, D. S. (2013). Lithostratigraphy of the lower Permian (Leonardian) Clear Fork Formation of north-central Texas. In S. G. Lucas, W. A. DiMichele, J. E. Barrick, J. W. Schneider, and J. A. Spielmann, editors, *The Carboniferous-Permian Transition*, Bulletin 60, pages 286–311. New Mexico Museum of Natural History and Science.

Olroyd, S. L. and Sidor, C. A. (2017). A review of the Guadalupian (middle Permian) global tetrapod fossil record. *Earth-Science Reviews*, **171**, 583–597.

Olson, E. and Vaughn, P. (1970). The changes of terrestrial vertebrates and climates during the Permian of North America. *Forma et functio*, **3**, 113–138.

Olson, E. C. (1965). New permian vertebrates from the Chickasha Formation in Oklahoma. *Oklahoma Geological Survey, Circular*, **70**, 1–70.

Olson, E. C. (1968). The family caseidae. *Fieldiana Geology*, **17**, 225–349.

Olson, E. C. (1975). Permo—Carboniferous paleoecology and morphotypic series. *American Zoologist*, **15**(2), 371–389.

Olson, E. C. (1982). Extinctions of Permian and Triassic nonmarine vertebrates. *Geological Society of America Special Papers*, **190**, 501–512.

Patterson, C. and Smith, A. B. (1987). Is the periodicity of extinctions a taxonomic artefact? *Nature*, **330**(6145), 248–251.

Payne, J. L. (2005). Evolutionary dynamics of gastropod size across the end-Permian extinction and through the Triassic recovery interval. *Paleobiology*, **31**(2), 269–290.

Plummer, M., Best, N., Cowles, K., and Vines, K. (2006). CODA: Convergence Diagnosis and Output Analysis for MCMC. *R News*, **6**(1), 7–11.

Raup, D. and Sepkoski, J. (1984). Periodicity of extinctions in the geologic past. *Proceedings of the National Academy of Sciences*, **81**(3), 801.

Raup, D. M. and Sepkoski, J. J. (1982). Mass extinctions in the marine fossil record. *Science*, **215**(4539), 1501–1503.

Reisz, R. (1972). Pelycosaurian reptiles from the middle Pennsylvanian of North America. *Bulletin of the Museum of Comparative Zoology*, **144**(2), 27–62.

Reisz, R. (1986). *Pelycosauria*, volume 17. Lubrecht & Cramer Ltd.

Reisz, R. R. and Laurin, M. (2001). The reptile *Macroleter*: First vertebrate evidence for correlation of Upper Permian continental strata of North America and Russia. *Geological Society of America Bulletin*, **113**(9), 1229–1233.

Reisz, R. R. and Laurin, M. (2002). Discussion and reply: The reptile *Macroleter*: First vertebrate evidence for correlation of Upper Permian continental strata of North America and Russia Discussion. *Geological Society of America Bulletin*, **114**(9), 1176–1177.

Reisz, R. R., Laurin, M., and Marjanović, D. (2010). *Apsisaurus witteri* from the Lower Permian of Texas: yet another small varanopid synapsid, not a diapsid. *Journal of Vertebrate Paleontology*, **30**(5), 1628–1631.

Retallack, G. J., Metzger, C. A., Greaver, T., Jahren, A. H., Smith, R. M., and Sheldon, N. D. (2006). Middle-Late Permian mass extinction on land. *Geological Society of America Bulletin*, **118**(11-12), 1398–1411.

Rivadeneira, M. M., Hunt, G., and Roy, K. (2009). The use of sighting records to infer species extinctions: An evaluation of different methods. *Ecology*, **90**(5), 1291–1300.

Robert, C. and Casella, G. (2013). *Monte Carlo statistical methods*. Springer Science & Business Media.

Romano, C., Goudemand, N., Vennemann, T. W., Ware, D., Schneebeli-Hermann, E., Hochuli, P. A., Brühwiler, T., Brinkmann, W., and Bucher, H. (2013). Climatic and biotic upheavals following the end-Permian mass extinction. *Nature Geoscience*, **6**(1), 57–60.

Romer, A. S. and Price, L. W. (1940). Review of the Pelycosauria. *Geological Society of America Special Papers*, **28**, 1–534.

Ronquist, F., Klopfstein, S., Vilhelmsen, L., Schulmeister, S., Murray, D. L., and Rasnitsyn, A. P. (2012). A total-evidence approach to dating with fossils, applied to the early radiation of the Hymenoptera. *Systematic Biology*, **61**(6), 973–999.

Rubidge, B. and Day, M. (2020). Biostratigraphy of the *Eodicynodon* Assemblage Zone (Beaufort Group, Karoo Supergroup), South Africa. *South African Journal of Geology 2020*, **123**(2), 141–148.

Rubidge, B. S. (2005). 27th du Toit memorial lecture: re-uniting lost continents-fossil reptiles from the ancient Karoo and their wanderlust. *South African Journal of Geology*, **108**(1), 135–172.

Ruta, M., Cisneros, J. C., Liebrecht, T., Tsuji, L. A., and Mueller, J. (2011). Amniotes through major biological crises: faunal turnover among Parareptiles and the end-Permian mass extinction. *Palaeontology*, **54**(5), 1117–1137.

Sadler, P. M. (2004). Quantitative biostratigraphy-achieving finer resolution in global correlation. *Annual Review of Earth and Planetary Sciences*, **32**, 187–213.

Sahney, S. and Benton, M. J. (2008). Recovery from the most profound mass extinction of all time. *Proceedings of the Royal Society B: Biological Sciences*, **275**(1636), 759–765.

Sander, P. M. (1989). Early Permian depositional environments and pond bonebeds in central Archer County, Texas. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **69**, 1–21.

Sanderson, M. J. (2002). Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Molecular biology and evolution*, **19**(1), 101–109.

Schneider, J. W., Lucas, S. G., Scholze, F., Voigt, S., Marchetti, L., Klein, H., Opluštil, S., Werneburg, R., Golubev, V. K., Barrick, J. E., *et al.* (2020). Late Paleozoic–early Mesozoic continental biostratigraphy—links to the standard global chronostratigraphic scale. *Palaeoworld*, **29**(2), 186–238.

Shaw, A. B. (1964). *Time in stratigraphy*. McGraw-Hill, New York.

Shipman, P. (1981). *Life history of a fossil: an introduction to taphonomy and paleoecology*. Harvard University Press, Cambridge, Massachusetts.

Sidor, C. A. (2001). Simplification as a trend in synapsid cranial evolution. *Evolution*, **55**(7), 1419–1442.

Sidor, C. A., Vilhena, D. A., Angielczyk, K. D., Huttenlocker, A. K., Nesbitt, S. J., Peecook, B. R., Steyer, J. S., Smith, R. M. H., and Tsuji, L. A. (2013). Provincialization of terrestrial faunas following the end-Permian mass extinction. *Proceedings of the National Academy of Sciences*, **110**(20), 8129–8133.

Signor, P. W., Lipps, J. H., Silver, L., Schultz, P., *et al.* (1982). Sampling bias, gradual extinction patterns, and catastrophes in the fossil record. *Geological implications of impacts of large asteroids and comets on the Earth*, **190**, 291–296.

Silvestro, D., Schnitzler, J., Liow, L. H., Antonelli, A., and Salamin, N. (2014a). Bayesian Estimation of Speciation and Extinction from Incomplete Fossil Occurrence Data. *Systematic Biology*, **63**(3), 349–367.

Silvestro, D., Salamin, N., and Schnitzler, J. (2014b). PyRate: a new program to estimate speciation and extinction rates from incomplete fossil data. *Methods in Ecology and Evolution*, **5**(10), 1126–1131.

Slater, G. J. (2015). Iterative adaptive radiations of fossil canids show no evidence for diversity-dependent trait evolution. *Proceedings of the National Academy of Sciences*, **112**(16), 4897–4902.

Smith, R., Rubidge, B., and Van der Walt, M. (2012). Therapsid biodiversity patterns and paleoenvironments of the Karoo Basin, South Africa. *Forerunners of Mammals: Radiation, Histology, Biology. Indiana University Press, Bloomington, Indiana*, pages 30–62.

Smith, R. M. and Botha-Brink, J. (2014). Anatomy of a mass extinction: sedimentological and taphonomic evidence for drought-induced die-offs at the Permo-Triassic boundary in the main Karoo Basin, South Africa. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **396**, 99–118.

Solow, A. R. (2016). A simple Bayesian method of inferring extinction: comment. *Ecology*, **97**(3), 796–798.

Soreghan, G. S., Benison, K. C., Foster, T. M., Zambito, J., and Soreghan, M. J. (2015). The paleoclimatic and geochronologic utility of coring red beds and evaporites: a case study from the RKB core (Permian, Kansas, USA). *International Journal of Earth Sciences*, **104**(6), 1589–1603.

Spindler, F. (2014). Reviewing the question of the oldest therapsid. *Freiberger Forschungshefte C*, **548**(22), 1–7.

Spindler, F. (2020). The skull of *Tetraceratops insignis* (Synapsida, Sphenacodontia). *Palaeovertebrata*, **43**(1), 1–11.

Stadler, T. (2010). Sampling-through-time in birth-death trees. *Journal of Theoretical Biology*, **267**(3), 396–404.

Stadler, T. and Yang, Z. (2013). Dating Phylogenies with Sequentially Sampled Tips. *Systematic Biology*, **62**(5), 674–688.

Stanley, S. M. (1988). Paleozoic mass extinctions: shared patterns suggest global cooling as a common cause. *American Journal of Science*, **288**(4), 334–352.

Sterli, J., Pol, D., and Laurin, M. (2013). Incorporating phylogenetic uncertainty on phylogeny-based palaeontological dating and the timing of turtle diversification. *Cladistics*, **29**(3), 233–246.

Steyer, J. S. (2000). Are European Paleozoic amphibians good stratigraphic markers? *Bulletin de la Société géologique de France*, **171**(1), 127–135.

Strauss, D. and Sadler, P. M. (1989). Classical confidence intervals and Bayesian probability estimates for ends of local taxon ranges. *Mathematical Geology*, **21**(4), 411–427.

Tabor, C. R., Bardeen, C. G., Otto-Bliesner, B. L., Garcia, R. R., and Toon, O. B. (2020). Causes and climatic consequences of the impact winter at the cretaceous-paleogene boundary. *Geophysical Research Letters*, **47**(3), e60121.

Tavaré, S., Marshall, C., Will, O., Soligo, C., and Martin, R. (2002). Using the fossil record to estimate the age of the last common ancestor of extant primates. *Nature*, **416**(6882), 726–729.

Uhen, M. D. (1996). An Evaluation of Clade-Shape Statistics Using Simulations and Extinct Families of Mammals. *Paleobiology*, **22**(1), 8–22.

Viglietti, P. A., Benson, R. B., Smith, R. M., Botha, J., Kammerer, C. F., Skosan, Z., Butler, E., Crean, A., Eloff, B., Kaal, S., *et al.* (2021). Evidence from South Africa for a protracted end-permian extinction on land. *Proceedings of the National Academy of Sciences*, **118**(17).

Wagner, P. J. (2000). The Quality of the Fossil Record and the Accuracy of Phylogenetic Inferences about Sampling and Diversity. *Systematic Biology*, **49**(1), 65–86.

Wake, D. B. and Vredenburg, V. T. (2008). Are we in the midst of the sixth mass extinction? a view from the world of amphibians. *Proceedings of the National Academy of Sciences*, **105**(Supplement 1), 11466–11473.

Wang, S. C. and Marshall, C. R. (2016). Estimating times of extinction in the fossil record. *Biology letters*, **12**(4), 20150989.

Ward, P., Botha, J., Buick, R., De Kock, M., Erwin, D., Garrison, G., Kirschvink, J., and Smith, R. (2005). Abrupt and gradual extinction among Late Permian land vertebrates in the Karoo Basin, South Africa. *Science*, **307**(5710), 709.

Warnock, R. C. M., Yang, Z., and Donoghue, P. C. J. (2012). Exploring uncertainty in the calibration of the molecular clock. *Biology Letters*, **8**(1), 156–159.

Warnock, R. C. M., Parham, J. F., Joyce, W. G., Lyson, T. R., and Donoghue, P. C. J. (2015). Calibration uncertainty in molecular dating analyses: there is no substitute for the prior evaluation of time priors. *Proceedings of the Royal Society of London B: Biological Sciences*, **282**(1798), 1–10.

Wignall, P. B., Sun, Y., Bond, D. P., Izon, G., Newton, R. J., Védrine, S., Widdowson, M., Ali, J. R., Lai, X., Jiang, H., *et al.* (2009). Volcanism, mass extinction, and carbon isotope fluctuations in the Middle Permian of China. *Science*, **324**(5931), 1179–1182.

Wilson, L. A., Colombo, M., Hanel, R., Salzburger, W., and Sánchez-Villagra, M. R. (2013). Ecomorphological disparity in an adaptive radiation: opercular bone shape and stable isotopes in Antarctic icefishes. *Ecology and evolution*, **3**(9), 3166–3182.

Zuckerkandl, E. and Pauling, L. (1965). Evolutionary divergence and convergence in proteins. In V. Bryson and H. J. Vogel, editors, *Evolving genes and proteins*, pages 97–166. Elsevier.

A A Metropolis-Hastings importance sampling procedure for the distribution of the extinction times of an arbitrary subset of taxa

We aim to compute the joint probability of a tree $\mathcal{T}$ with fossils (without its divergence times), a fossil age vector $\mathbf{f}$ and its origin time, and that a particular subset S of (extinct at the present time) taxa of $\mathcal{T}$ goes extinct before a time t. Let us put η_S for the extinction time of the taxa in S and for all taxon n of S, $\ell_{\mathbf{f},n}$ for the age of its most recent fossil with regard to the fossil age vector $\mathbf{f}$. The probability density of interest, $\mathbf{P}(\mathcal{T}, \mathbf{f}, \eta_n \leq t)$, reads

$$\mathbf{P}(\mathcal{T}, \mathbf{f}, \eta_S \leq t) = \mathbf{P}(\mathcal{T}, \mathbf{f}) \prod_{n \in S} \frac{\mathbf{P}(0, t - \ell_{\mathbf{f},n})}{\mathbf{P}(0, T - \ell_{\mathbf{f},n})}.$$

The ratio $\frac{\mathbf{P}(0,t-\ell_{\mathbf{f},n})}{\mathbf{P}(0,T-\ell_{\mathbf{f},n})}$ means that we replace the factor contribution of the unobservable tree ending the evolution of taxon n in $\mathbf{P}(\mathcal{T}, \mathbf{f})$, which is the probability that it goes extinct before the present time, i.e., $\mathbf{P}(0, T - \ell_{\mathbf{f},n})$, by $\mathbf{P}(0, t - \ell_{\mathbf{f},n})$, which is the probability that it goes extinct before t.

The computation of the probability density above holds in the case where both the topology and the fossil ages are exactly known. In practice, there is a lot of uncertainty over these two features. In particular, our dataset contains a set Ω of equiparsimonious topologies and the only information about the age of each fossil f is provided as a geological age interval I_f. Given that with a tree $\mathcal{T} \in \Omega$ and a vector of fossil ages $\mathbf{f} \in \mathbf{I}$, we are able to compute the probability densities $\mathbf{P}(\mathcal{T}, \mathbf{f}, \eta_S \leq t)$ for all times t, determining distributions from our dataset may be performed by summing over all the topologies and integrating over all the time intervals.

We assume prior distributions in which (i) all the topologies in Ω are equiprobable and (ii) the fossil ages are uniformly distributed over their time intervals. We put $\mathbf{I}$ for the product-set of the time intervals $\mathbf{I} = \bigotimes_f \mathrm{I}_f$. The joint probability density above conditioned on the whole dataset is

$$\frac{\sum_{\mathcal{T} \in \Omega} \int_{\mathbf{I}} \mathbf{P}(\mathcal{T}, \mathbf{f}, \eta_S \leq t) \mathrm{d}\mathbf{f}}{\sum_{\mathcal{T} \in \Omega} \int_{\mathbf{I}} \mathbf{P}(\mathcal{T}, \mathbf{f}) \mathrm{d}\mathbf{f}} = \frac{\int_{\mathbf{I}} \sum_{\mathcal{T} \in \Omega} \mathbf{P}(\mathcal{T}, \mathbf{f}, \eta_S \leq t) \mathrm{d}\mathbf{f}}{\int_{\mathbf{I}} \sum_{\mathcal{T} \in \Omega} \mathbf{P}(\mathcal{T}, \mathbf{f}) \mathrm{d}\mathbf{f}}. \tag{2}$$

Since we have not closed formula for the quantities involved in the ratio above, it has to be numerically computed, for instance by uniformly sampling the fossil ages in $\mathbf{I}$. Being given N uniform samples of fossil age vectors $\mathbf{f}_1, \ldots, \mathbf{f}_N$, the estimated ratio is

$$\frac{\sum_{i=1}^{N} \sum_{\mathcal{T} \in \Omega} \mathbf{P}(\mathcal{T}, \mathbf{f}, \eta_S \leq t)}{\sum_{i=1}^{N} \sum_{\mathcal{T} \in \Omega} \mathbf{P}(\mathcal{T}, \mathbf{f}_i)}.$$

Unfortunately, because of its variance, the estimator above requires millions of samples to converge in our dataset. We tackle this issue in the same way as in Didier and Laurin (2020), that is by performing importance sampling (see, e.g., Robert and Casella, 2013). Briefly, an importance sampling procedure estimates the integral $\int_{\mathbf{I}} \sum_{\mathcal{T} \in \Omega} \mathbf{P}(\mathcal{T}, \mathbf{f}, \eta_n = t) \mathrm{d}\mathbf{f}$ by sampling the fossil ages not uniformly but following a density $\mathrm{q}^\star$ biased toward the ages corresponding to high values of $\sum_{\mathcal{T} \in \Omega} \mathbf{P}(\mathcal{T}, \mathbf{f}, \eta_n = t)$. We have basically that

$$\int_{\mathbf{I}} \sum_{\mathcal{T} \in \Omega} \mathbf{P}(\mathcal{T}, \mathbf{f}, \eta_S \leq t) \mathrm{d}\mathbf{f} = \int_{\mathbf{I}} \sum_{\mathcal{T} \in \Omega} \frac{\mathbf{P}(\mathcal{T}, \mathbf{f}, \eta_S \leq t)}{\mathrm{q}^\star(\mathbf{f})} \mathrm{q}^\star(\mathbf{f}) \mathrm{d}\mathbf{f}.$$

It follows that if $\mathbf{f}'_1, \ldots, \mathbf{f}'_N$ are independently sampled from density $\mathrm{q}^\star$, this integral can be estimated by

$$\frac{1}{N} \sum_{i=1}^{N} \frac{\sum_{\mathcal{T} \in \Omega} \mathbf{P}(\mathcal{T}, \mathbf{f}'_i, \eta_S \leq t)}{\mathrm{q}^\star(\mathbf{f}'_i)}.$$

As in Didier and Laurin (2020), by noting that $\sum_{\mathcal{T} \in \Omega} \mathbf{P}(\mathcal{T}, \mathbf{f}', \eta_S \leq t)$ is bounded by (and strongly related to) $\sum_{\mathcal{T} \in \Omega} \mathbf{P}(\mathcal{T}, \mathbf{f}')$, we consider a density $\mathrm{q}^\star(\mathbf{f})$ proportional to $\sum_{\mathcal{T} \in \Omega} \mathbf{P}(\mathcal{T}, \mathbf{f})$ for our importance sampling procedure. Namely, we set

$$\mathrm{q}^\star(\mathbf{f}') = \frac{\sum_{\mathcal{T} \in \Omega} \mathbf{P}(\mathcal{T}, \mathbf{f}')}{\int_{\mathbf{I}} \sum_{\mathcal{T} \in \Omega} \mathbf{P}(\mathcal{T}, \mathbf{f}) \mathrm{d}\mathbf{f}}.$$

Given N iid samples $\mathbf{f}'_1, \ldots, \mathbf{f}'_N$ obtained with the density $\mathrm{q}^\star$ just above, the ratio of Equation 2 is estimated by

$$\begin{aligned}
\frac{\frac{1}{N} \sum_{i=1}^{N} \frac{\sum_{\mathcal{T} \in \Omega} \mathbf{P}(\mathcal{T}, \mathbf{f}'_i, \eta_S \leq t)}{\mathrm{q}^\star(\mathbf{f}'_i)}}{\int_{\mathbf{I}} \sum_{\mathcal{T} \in \Omega} \mathbf{P}(\mathcal{T}, \mathbf{f}) \mathrm{d}\mathbf{f}} &= \frac{1}{N} \sum_{i=1}^{N} \frac{\sum_{\mathcal{T} \in \Omega} \mathbf{P}(\mathcal{T}, \mathbf{f}'_i, \eta_S \leq t)}{\sum_{\mathcal{T} \in \Omega} \mathbf{P}(\mathcal{T}, \mathbf{f}'_i)} \\
&= \frac{1}{N} \sum_{i=1}^{N} \frac{\sum_{\mathcal{T} \in \Omega} \mathbf{P}(\mathcal{T}, \mathbf{f}'_i) \prod_{n \in S} \frac{\mathbf{P}(0, t - \ell_{\mathbf{f}'_i, n})}{\mathbf{P}(0, T - \ell_{\mathbf{f}'_i, n})}}{\sum_{\mathcal{T} \in \Omega} \mathbf{P}(\mathcal{T}, \mathbf{f}'_i)} \\
&= \frac{1}{N} \sum_{i=1}^{N} \prod_{n \in S} \frac{\mathbf{P}(0, t - \ell_{\mathbf{f}'_i, n})}{\mathbf{P}(0, T - \ell_{\mathbf{f}'_i, n})}.
\end{aligned}$$

Since it is not possible to directly sample following the density $\mathrm{q}^\star$, which is known only up to the intractable normalization constant $\int_{\mathbf{I}} \sum_{\mathcal{T} \in \Omega} \mathbf{P}(\mathcal{T}, \mathbf{f}) \mathrm{d}\mathbf{f}$, we proceed in the same way as Didier and Laurin (2020) by using the Metropolis-Hastings algorithm (e.g., Robert and Casella, 2013) to build a Markov chain with target density $\mathrm{q}^\star$. Let us recall the details of this procedure. Given the ith vector of fossil ages $\mathbf{f}'_i$ of this chain, we generate a candidate vector $\tilde{\mathbf{f}}$ first by uniformly selecting a fossil f among all the fossils of the dataset and by replacing the corresponding entry of $\mathbf{f}'_i$ by a uniform sample in the interval centered on this entry of width equal to α time the with of the interval I_f that represents the uncertainty on the age of fossil f (i.e., we use a usual sliding window proposal with reflection; α is a user-defined parameter). This proposal transition kernel is basically symmetrical, i.e., getting the candidate vector $\tilde{\mathbf{f}}$ from $\mathbf{f}'_i$ has the same probability as getting the candidate vector $\mathbf{f}'_i$ from $\tilde{\mathbf{f}}$. The next fossil age vector $\mathbf{f}'_{i+1}$ in the Markov chain is

$$\mathbf{f}'_{i+1} = \begin{cases} \tilde{\mathbf{f}} & \text{with probability } \min\left\{ \dfrac{\sum_{\mathcal{T} \in \Omega} \mathbf{P}(\mathcal{T}, \tilde{\mathbf{f}})}{\sum_{\mathcal{T} \in \Omega} \mathbf{P}(\mathcal{T}, \mathbf{f}'_i)}, 1 \right\} \text{ and} \\ \mathbf{f}'_i & \text{with probability } 1 - \min\left\{ \dfrac{\sum_{\mathcal{T} \in \Omega} \mathbf{P}(\mathcal{T}, \tilde{\mathbf{f}})}{\sum_{\mathcal{T} \in \Omega} \mathbf{P}(\mathcal{T}, \mathbf{f}'_i)}, 1 \right\}. \end{cases}$$

In the case where the speciation, extinction and fossilization rates of the model are not given *a priori* (e.g., estimated by maximum likelihood), we sum over all their possible values by assuming that they all follow the improper uniform distribution over $[0, +\infty]$ and by following a similar importance sampling procedure. In plain English, the importance sampling is now performed over both the fossil age vector $\mathbf{f}$ and the model rates $\Theta = (\lambda, \mu, \psi)$ by using the biased density $\mathrm{q}^\star(\mathbf{f}, \Theta)$ proportional to $\sum_{\mathcal{T} \in \Omega} \mathbf{P}_\Theta(\mathcal{T}, \mathbf{f}')$, where $\mathbf{P}_\Theta(\mathcal{T}, \mathbf{f}')$ is the probability of the phylogeny with fossils $\mathcal{T}, \mathbf{f}'$ under the FBD model with rates $\Theta = (\lambda, \mu, \psi)$.

The proposal mechanism of the Markov chain used to sample according to this density first draws a Bernoulli random variable with a probability β (provided by the user) according to which either a change in a fossil age or a change in the speciation, extinction or fossilization rate is proposed. The change in a fossil age is proposed in the exact same way as just above. The change in a rate is proposed first by selected randomly the speciation, extinction or fossilization rate according to user-defined probabilities then by using a sliding window of fixed length centered on the current selected rate.

The densities presented in the results section were obtained from the MCMC procedure presented just above by discarding the first 10000 iterations, and by keeping only 1 iteration over 200 for the subsequent iterations. We consider a window of width equal to 80% of the length of the age interval I_f to propose a change in the age of fossil f and windows of size 0.5 to propose changes in the speciation, extinction and fossilization rates respectively. We propose a change in a fossil age with probability 0.75 and a change in the speciation, extinction or fossilization rate otherwise (the rate to be changed among the the speciation, extinction and fossilization rate is then drawn uniformly). We run the MCMC procedure until get 10000 iterations. The iterations thus obtained pass all the convergence tests of the `coda` R package with an effective size greater than 1000 for all the parameters (fossil ages and rates, Plummer *et al.*, 2006). All the parameters of the MCMC were manually tuned in order to get a satisfying behavior. They can be modified by calling the corresponding R function in the package.

B Probability that a subset of extinct taxa goes extinct before another

Let $\mathcal{T}$ be a tree with fossils (without its divergence times), $\mathbf{f}$ be a fossil age vector and A and B be two disjoint subsets of extinct taxa of $\mathcal{T}$ (e.g., two mutually exclusive extinct clades of $\mathcal{T}$). We aim here to compute the probability density $\mathbf{P}(\mathcal{T}, \mathbf{f}, \eta_A \leq \eta_B)$, which is the joint probability density of the data $(\mathcal{T}, \mathbf{f})$ and that the taxa of A go extinct before those of B.

We shall consider the same situation as in Appendix A and assume that data are provided as a set Ω of equiparsimonious trees and that the fossil ages are just bounded by geological intervals summarized in the product-set $\mathbf{I}$. The joint probability density above conditioned on the whole dataset is

$$\frac{\sum_{\mathcal{T} \in \Omega} \int_{\mathbf{I}} \mathbf{P}(\mathcal{T}, \mathbf{f}, \eta_A \leq \eta_B) \mathrm{d}\mathbf{f}}{\sum_{\mathcal{T} \in \Omega} \int_{\mathbf{I}} \mathbf{P}(\mathcal{T}, \mathbf{f}) \mathrm{d}\mathbf{f}} = \frac{\int_{\mathbf{I}} \sum_{\mathcal{T} \in \Omega} \mathbf{P}(\mathcal{T}, \mathbf{f}, \eta_A \leq \eta_B) \mathrm{d}\mathbf{f}}{\int_{\mathbf{I}} \sum_{\mathcal{T} \in \Omega} \mathbf{P}(\mathcal{T}, \mathbf{f}) \mathrm{d}\mathbf{f}}. \tag{3}$$

Let us first show how to compute $\mathbf{P}(\mathcal{T}, \mathbf{f}, \eta_A \leq \eta_B)$. We recall that for all extinct taxa n, $\ell_{\mathbf{f},n}$ stands for the most recent fossil age of n and we put $\ell_{\mathbf{f},A \cup B}$ for the most recent age of fossils of $A \cup B$, i.e., $\ell_{\mathbf{f},A \cup B} = \max_{n \in A \cup B} \ell_{\mathbf{f},n}$. The probability density $\mathbf{P}(\mathcal{T}, \mathbf{f}, \eta_A \leq \eta_B)$ can be obtained by integrating over all the possible ages t of extinction of B, the joint probability density that B goes extinct exactly at t and A goes extinct before t. Note that the possible ages t such that B goes extinct exactly at t and A goes extinct before t run from $\ell_{\mathbf{f},A \cup B}$ to the present time T. Moreover, the probability density that B goes extinct exactly at t is the sum over all the taxa m of B of the probability densities that m goes extinct exactly at t and all the other taxa of B go extinct before t. It follows that the probability density

$\mathbf{P}(\mathcal{T}, \mathbf{f}, \eta_A \leq \eta_B)$ can be written

$$\begin{aligned}
\mathbf{P}(\mathcal{T}, \mathbf{f}, \eta_A \leq \eta_B) &= \int_{\ell_{\mathbf{f},A\cup B}}^{T} \mathbf{P}(\mathcal{T}, \mathbf{f}, \eta_A \leq t, \eta_B = t)\mathrm{d}t \\
&= \int_{\ell_{\mathbf{f},A\cup B}}^{T} \sum_{m\in B} \mathbf{P}(\mathcal{T}, \mathbf{f}, \eta_A \leq t, \eta_{B\setminus\{j\}} \leq t, \eta_m = t)\mathrm{d}t \\
&= \frac{\mathbf{P}(\mathcal{T}, \mathbf{f}) \sum_{m\in B} \int_{\ell_{\mathbf{f},A\cup B}}^{T} \left(\prod_{n\in A\cup B\setminus\{m\}} \mathbf{P}(0, t-\ell_{\mathbf{f},n})\right) \mathbf{D}(t-\ell_{\mathbf{f},m})\mathrm{d}t}{\prod_{n\in A\cup B} \mathbf{P}(0, T-\ell_{\mathbf{f},n})} \\
&= \frac{\mathbf{P}(\mathcal{T}, \mathbf{f})\mathcal{I}_{\mathbf{f},A,B}}{\prod_{n\in A\cup B} \mathbf{P}(0, T-\ell_{\mathbf{f},n})},
\end{aligned}$$

where $\mathcal{I}_{\mathbf{f},A,B} = \sum_{m\in B} \int_{\ell_{\mathbf{f},A\cup B}}^{T} \left(\prod_{n\in A\cup B\setminus\{m\}} \mathbf{P}(0, t-\ell_{\mathbf{f},n})\right) \mathbf{D}(t-\ell_{\mathbf{f},m})\mathrm{d}t$.

Since we assume here that the extant sampling parameter $\rho = 1$, Equations 1 give us that

$$\mathcal{I}_{\mathbf{f},A,B} = \sum_{m\in B} \int_{\ell_{\mathbf{f},A\cup B}}^{T} \left(\prod_{n\in A\cup B\setminus\{m\}} \frac{\alpha\beta(1-\mathrm{e}^{\omega(t-\ell_{\mathbf{f},n})})}{\beta-\alpha\mathrm{e}^{\omega(t-\ell_{\mathbf{f},n})}}\right) \frac{\mu(\beta-\alpha)^2\mathrm{e}^{\omega(t-t_m)}}{(\beta-\alpha\mathrm{e}^{\omega(t-t_m)})^2}\mathrm{d}t.$$

A tedious calculation leads to

$$\begin{aligned}
\mathcal{I}_{\mathbf{f},A,B} = \beta(\beta-\alpha) \sum_{m\in B} &\left(\sum_{n\in A\cup B\setminus\{m\}} \frac{(\beta-\alpha)\Upsilon_{\mathbf{f},m,n}}{1-\mathrm{e}^{\omega(\ell_{\mathbf{f},m}-\ell_{\mathbf{f},n})}} - \beta^{|A\cup B|-1}\right) \left[\frac{1}{\beta-\alpha\mathrm{e}^{\omega(t-\ell_{\mathbf{f},m})}}\right]_{\ell_{\mathbf{f},A\cup B}}^{T} \\
&- (\beta-\alpha)^2 \sum_{m\in B} \sum_{n\in A\cup B\setminus\{m\}} \frac{\Upsilon_{\mathbf{f},m,n}\mathrm{e}^{\omega(\ell_{\mathbf{f},m}-\ell_{\mathbf{f},n})}}{(1-\mathrm{e}^{\omega(\ell_{\mathbf{f},m}-\ell_{\mathbf{f},n})})^2} \left[\log\left(1+\frac{\alpha(1-\mathrm{e}^{\omega(\ell_{\mathbf{f},m}-\ell_{\mathbf{f},n})})}{\beta\mathrm{e}^{\omega(\ell_{\mathbf{f},m}-t)}-\alpha}\right)\right]_{\ell_{\mathbf{f},A\cup B}}^{T},
\end{aligned}$$

where

$$\Upsilon_{\mathbf{f},m,n} = \prod_{k\in(A\cup B)\setminus\{m,n\}} \frac{\beta-\alpha\mathrm{e}^{\omega(\ell_{\mathbf{f},k}-\ell_{\mathbf{f},n})}}{1-\mathrm{e}^{\omega(\ell_{\mathbf{f},k}-\ell_{\mathbf{f},n})}}.$$

We thus have an explicit formula to compute $\mathbf{P}(\mathcal{T}, \mathbf{f}, \eta_A \leq \eta_B)$ which can be used in the importance sampling procedure presented in Appendix A to estimate the probability that the subset A goes extinct before B. Namely, if $\mathbf{f}'_1, \ldots, \mathbf{f}'_N$ are N iid samples of density $\mathrm{q}^\star$, the ratio of Equation 3 is estimated by

$$\begin{aligned}
\frac{\frac{1}{N}\sum_{i=1}^{N} \frac{\sum_{\mathcal{T}\in\Omega} \mathbf{P}(\mathcal{T}, \mathbf{f}'_i, \eta_A\leq\eta_B)}{\mathrm{q}^\star(\mathbf{f}'_i)}}{\int_{\mathbf{I}} \sum_{\mathcal{T}\in\Omega} \mathbf{P}(\mathcal{T}, \mathbf{f})\mathrm{d}\mathbf{f}} &= \frac{1}{N}\sum_{i=1}^{N} \frac{\sum_{\mathcal{T}\in\Omega} \mathbf{P}(\mathcal{T}, \mathbf{f}'_i, \eta_A \leq \eta_B)}{\sum_{\mathcal{T}\in\Omega} \mathbf{P}(\mathcal{T}, \mathbf{f}'_i)} \\
&= \frac{1}{N}\sum_{i=1}^{N} \frac{\mathcal{I}_{\mathbf{f}'_i,A,B}}{\prod_{n\in A\cup B} \mathbf{P}(0, T-\ell_{\mathbf{f}'_i,n})}.
\end{aligned}$$

C Simulation study with two other levels of risk

We performed the exact same protocol as that described in Section "Simulation study" by considering confidence upper bounds at two other orders: 50% (Tables 4 and 5) and 75% (Tables 6 and 7). Tables 4 and 6 display the error percentage, that is the percentage of cases where the "real" extinction date is posterior to the upper bound provided by the six assessed methods. Tables 5 and 7 show the mean confidence interval widths obtained with each method at orders 50% and 75% respectively. For both orders 50% and 75%, we observe the same general behavior as that observed with the 95% confidence upper bounds.

Fos. rate	**S&S**	**McI**	**Alr**	**Glo**			**Int**			**Ref**	
0.005	29.29	36.29	11.34	31.24	*32.73*	44.91	*48.15*	**49.59**	47.44	*49.92*	**50.12**
0.01	34.68	42.23	13.97	32.72	*34.67*	49.39	*50.43*	**51.16**	50.10	*50.17*	**50.06**
0.1	44.15	51.75	11.85	43.50	*43.19*	51.31	*51.36*	**51.62**	50.38	*50.15*	**50.15**
1	49.97	54.38	6.32	50.00	*50.02*	50.65	*50.68*	**50.80**	50.49	*50.46*	**50.46**
5	49.79	51.87	2.09	49.59	*49.65*	49.67	*49.74*	**49.78**	49.59	*49.66*	**49.66**

Table 4: Error percentages obtained from the 50% confidence upper bounds provided by the six methods on the simulated extinct taxa where the **S&S**, **McI** and **Alr** computations were feasible (plain text), on those belonging to a simulated tree where a global fossil recovery rate can be estimated (italics; only for methods **Glo**, **Ref** and **Int**) and on all the extinct taxa (bold; only for methods **Ref** and **Int**).

Fos. rate	**S&S**	**McI**	**Alr**	**Glo**			**Int**			**Ref**	
0.005	13.3	9.4	36.0	9.3	*9.5*	4.2	*4.1*	**3.9**	3.7	*3.7*	**3.7**
0.01	10.0	7.1	28.9	7.1	*7.1*	3.6	*3.5*	**3.4**	3.4	*3.4*	**3.4**
0.1	3.2	2.4	14.0	2.4	*2.4*	1.9	*1.8*	**1.8**	1.9	*1.9*	**1.9**
1	0.6	0.5	4.7	0.5	*0.5*	0.5	*0.5*	**0.5**	0.5	*0.5*	**0.5**
5	0.1	0.1	1.6	0.1	*0.1*	0.1	*0.1*	**0.1**	0.1	*0.1*	**0.1**

Table 5: Mean confidence interval width in million years, i.e. the mean difference between the 50% confidence upper bounds provided by the three methods and the most recent fossil, on the simulated extinct taxa where the **S&S**, **McI** and **Alr** computations were feasible (plain text), on those belonging to a simulated tree where a global fossil recovery rate can be estimated (italics; only for methods **Glo**, **Int** and **Ref**) and on all the extinct taxa (bold; only for methods **Int** and **Ref**).

D Posterior probability density distributions of speciation, extinction and fossilization rates from the simulation study

Figure 9 displays the aggregated posterior densities of the speciation, extinction and fossilization rates corresponding to the five runs of simulations presented in the section "Simulation study". Let us recall that we performed five runs of 1000 simulations, all with speciation and extinction rates 0.2 and 0.19 and with fossilization rates 0.005, 0.01, 0.1, 1 and 5 respectively. For each of the five fossilization rates, we plotted the average posterior density of the speciation, extinction and fossilization rates over the 1000 simulated trees.

We observe that each density is roughly centered on the corresponding rate value used to simulate the data and that the distributions of the speciation and extinction rates become tighter as the fossilization rate increases.

E Divergence time and posterior probability density distributions from the empirical dataset

Probability density distributions of divergence times

We applied the approach presented in Didier and Laurin (2020) to the empirical dataset used here to compute the divergence time densities of the tree displayed in Figure 3. Figure 10 displays this tree with the divergence time densities associated to all its internal nodes. The denser fossil record and our revisions of the ages of fossils based on recent literature yields slightly more recent divergence times (Fig. 10) than our previous study (Didier and Laurin, 2020, fig. 6).

Figure 11 presents the probability density of the divergence time between therapsids and sphenacontids. Didier and Laurin (2020) estimated that the divergence between sphenacodontids and therapsids dated from about 313 Ma (early Moscovian), but our current analysis places this divergence closer to 301 Ma, in the mid-Gzhelian (Fig. 11). Again, the more recent estimated age reflects the denser fossil record.

Fos. rate	**S&S**	**McI**	**Alr**	**Glo**			**Int**			**Ref**	
0.005	13.74	23.20	11.34	16.59	*18.59*	22.75	*24.62*	**25.26**	24.95	*25.25*	**25.19**
0.01	17.46	28.65	13.97	16.24	*18.50*	25.56	*26.34*	**26.52**	25.75	*25.56*	**25.46**
0.1	21.06	34.04	11.85	20.18	*20.51*	26.04	*26.17*	**26.26**	25.18	*25.06*	**25.06**
1	24.87	33.23	6.32	24.42	*24.65*	25.08	*25.33*	**25.37**	24.84	*25.03*	**25.03**
5	25.12	29.24	2.09	25.13	*25.19*	25.15	*25.21*	**25.23**	25.08	*25.13*	**25.13**

Table 6: Error percentages obtained from the 75% confidence upper bounds provided by the six methods on the simulated extinct taxa where the **S&S**, **McI** and **Alr** computations were feasible (plain text), on those belonging to a simulated tree where a global fossil recovery rate can be estimated (italics; only for methods **Glo**, **Ref** and **Int**) and on all the extinct taxa (bold; only for methods **Ref** and **Int**).

Fos. rate	**S&S**	**McI**	**Alr**	**Glo**			**Int**			**Ref**	
0.005	39.5	18.7	36.0	18.7	*19.0*	9.8	*9.5*	**9.3**	9.0	*9.0*	**9.0**
0.01	29.5	14.1	28.9	14.3	*14.2*	8.2	*8.0*	**7.9**	8.0	*8.0*	**8.0**
0.1	9.0	4.8	14.0	4.8	*4.8*	3.9	*3.9*	**3.9**	4.0	*4.0*	**4.0**
1	1.5	1.0	4.7	1.0	*1.0*	1.0	*1.0*	**1.0**	1.0	*1.0*	**1.0**
5	0.3	0.3	1.6	0.3	*0.3*	0.3	*0.3*	**0.3**	0.3	*0.3*	**0.3**

Table 7: Mean confidence interval width in million years, i.e. the mean difference between the 75% confidence upper bounds provided by the three methods and the most recent fossil, on the simulated extinct taxa where the **S&S**, **McI** and **Alr** computations were feasible (plain text), on those belonging to a simulated tree where a global fossil recovery rate can be estimated (italics; only for methods **Glo**, **Int** and **Ref**) and on all the extinct taxa (bold; only for methods **Int** and **Ref**).

Posterior probability density distributions of rates and fossil ages

Figure 12 displays some posterior probability density distributions obtained from the MCMC sampling used to compute Figure 11. Since the fossil ages are given as time intervals in the empirical dataset, they have to be sampled as presented in Appendix A. On top of the posterior distributions of the speciation, extinction and fossilization rates, we plotted those of a selection of fossil ages. Among these, we first observe that, though two fossils of *Dimetrodon limbatus* are associated to the time interval $(297, 290)$, the posterior densities of their age are negligible for ages greater than 293 Ma. Since the posterior density associated to an age value is proportional to the density of the phylogenetic tree when the corresponding fossil age is set to this value, this explains why the density of the speciation time of this taxon is negligible before 293 Ma, as shown in Figure 10. The posterior probability density distributions of fossils ages show various shapes. In particular, unlike that of *Dimetrodon limbatus*, some of them are almost uniform (e.g., that of *Tetraceratops insignis*).

Figure 12 shows that the posterior density of the fossilization rate supports rates about ten times higher than those observed with our dataset of Didier and Laurin (2020). This presumably reflects the incorporation of a fairly large amount of fossil occurrence data from the Paleobiology Database, which resulted in the incorporation of many additional occurrences of several taxa (but no additional taxa), as well as the more recent divergence times, which result in shorter branches. The slightly lower speciation and extinction rates observed in Figure 12 may simply reflect the fact that the taxonomic sample is not the same (the current study focuses on a subset of the clades included in our previous studies).

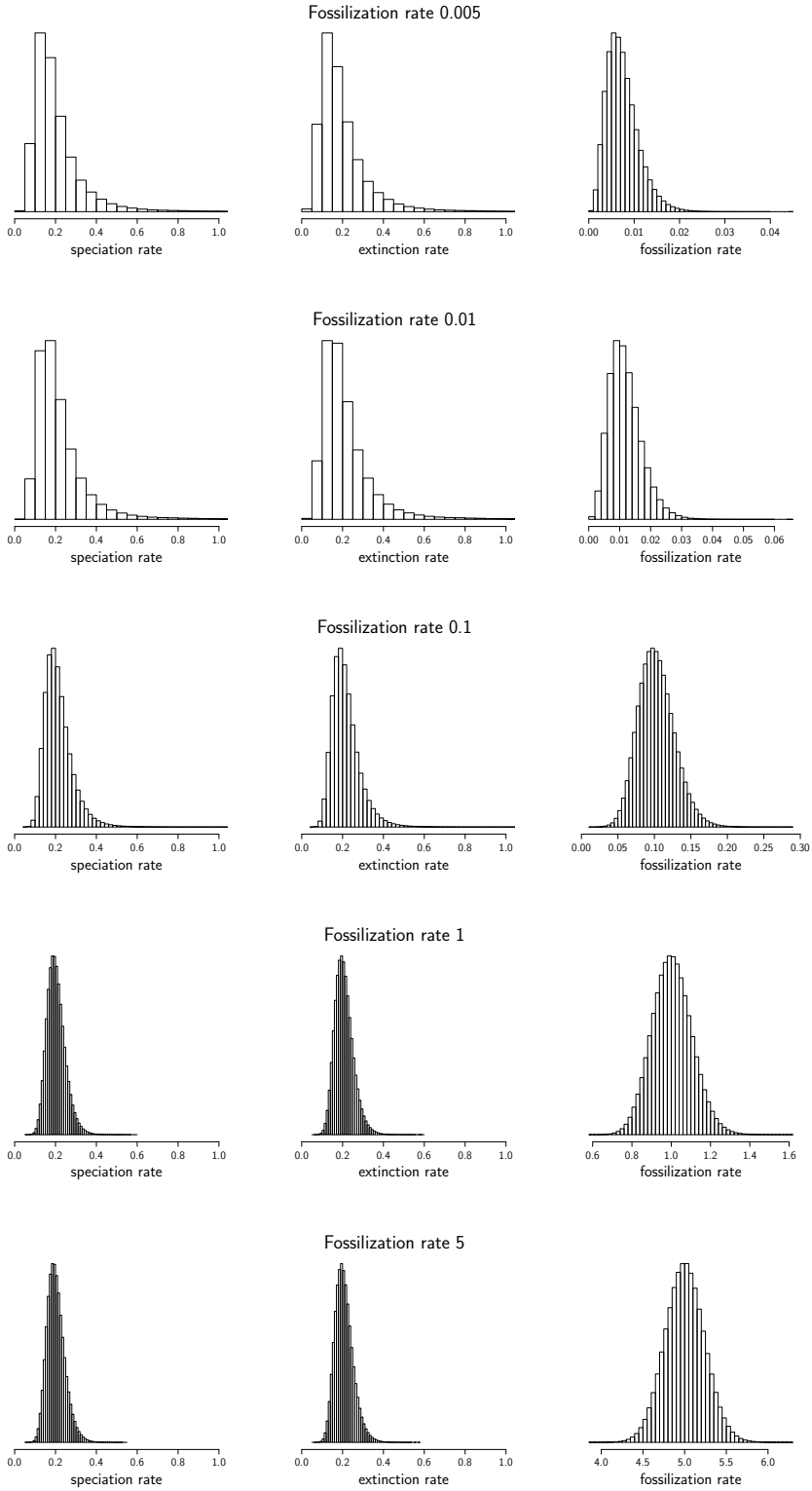

Figure 9: Posterior probability density distributions of the speciation, extinction and fossilization rates obtained from the MCMC procedure used to compute the upper bounds of extinction dates with the method **Int** (i.e., when the FBD rates are unknown) in our simulation study.

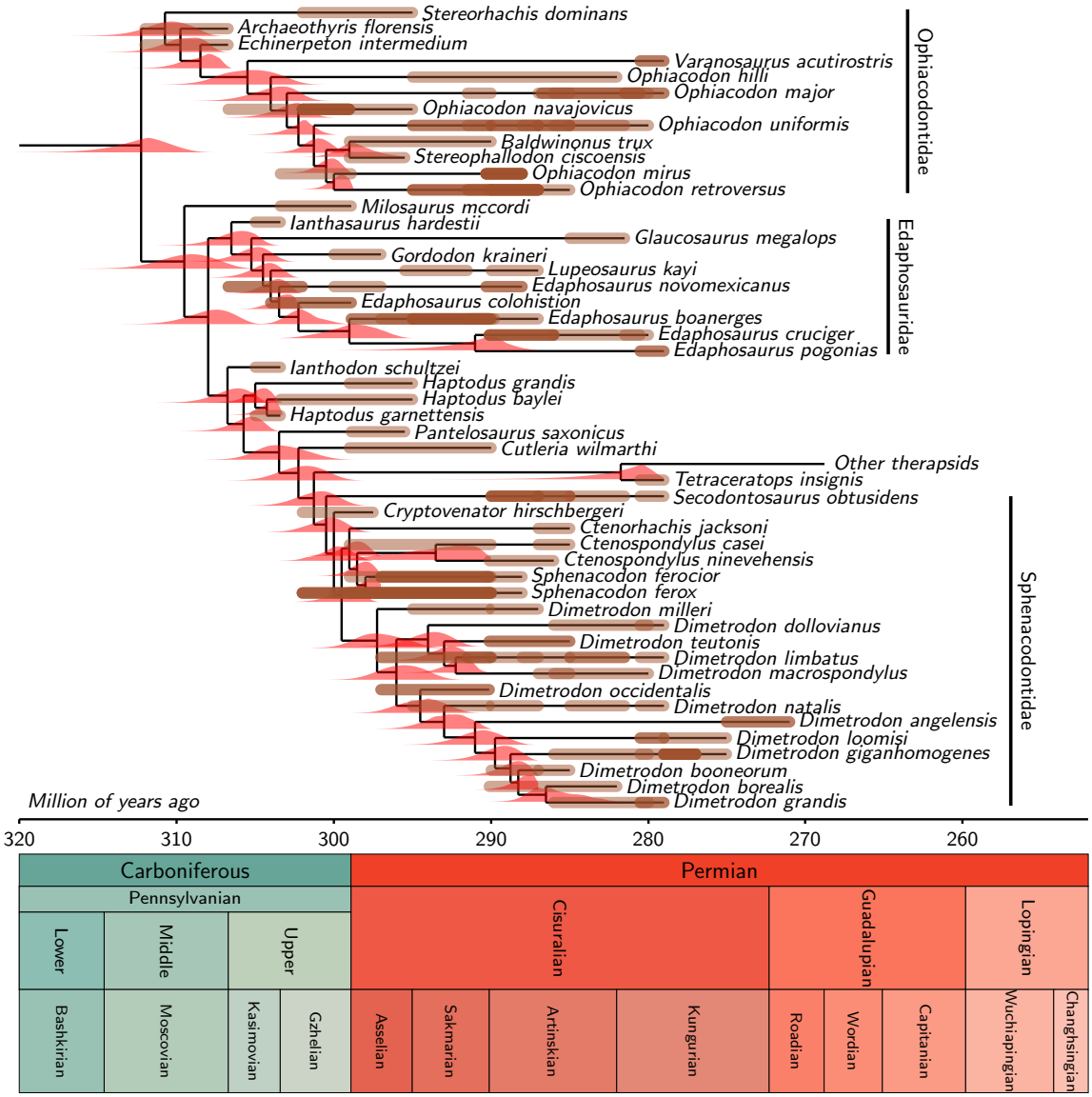

Figure 10: One of the 100 equally parsimonious trees of our dataset with the probability density distributions of its speciation times. Intervals of possible ages for each fossil are represented as thick brown segments with a certain level of transparency (darker brown segments correspond to overlapping intervals). The confidence interval on some fossil records extends below their basal node, most noticeably for *Dimetrodon limbatus*. With this tree topology, the oldest part of the range of possible ages of fossils of some taxa leads to very low probability densities of the whole tree with fossil; thus, it does not contribute significantly to the speciation time density distributions (Fig. 12). All nodes are older than the minimal age of all fossils, as they should be. The denser sample of the fossil record yields younger divergence times for eupelycosaur clades than those of our previous study (Didier and Laurin, 2020). This is because confidence intervals typically become narrower as sample size increases, and this phenomenon also applies to divergence times; as the fossil record becomes denser, our model adds less unobserved lineages below the first occurrence of a fossil on each lineage.

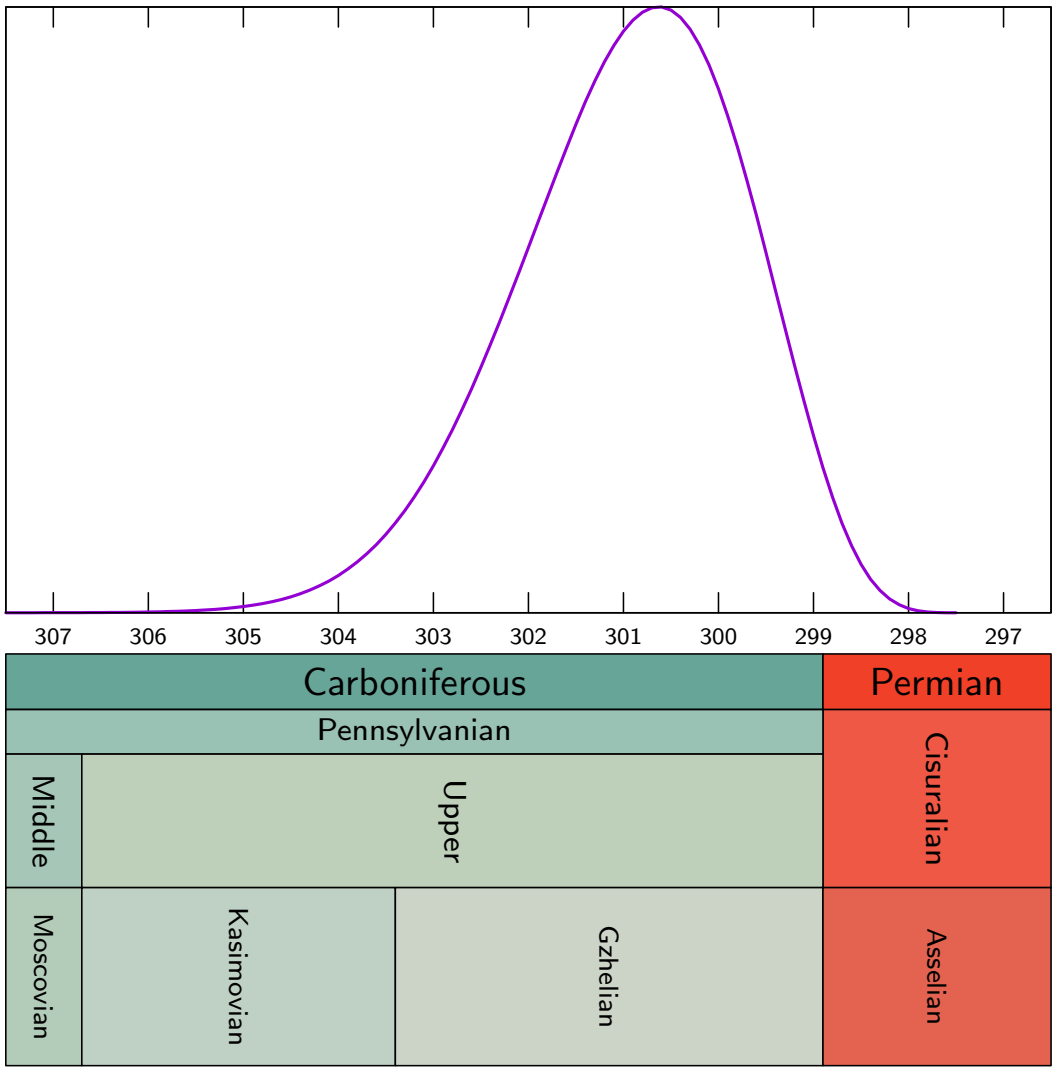

Figure 11: Probability density distribution of the age of the divergence between therapsids and sphenacontids.

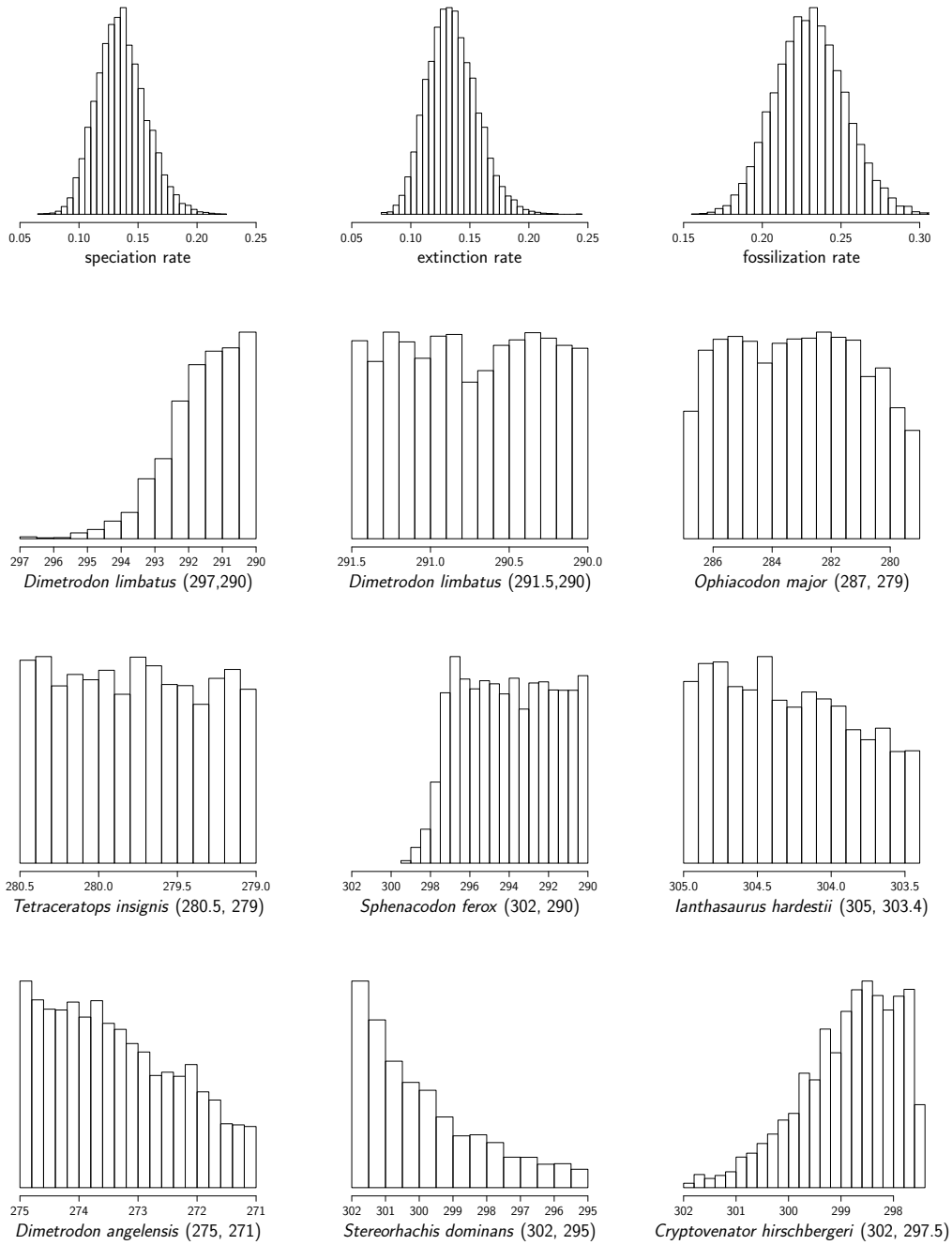

Figure 12: Posterior probability density distributions of the speciation, extinction and fossilization rates and of a selection of fossil ages obtained from the MCMC sampling used to compute the speciation distribution of Fig. 11.