TIMETREES: INCORPORATING FOSSILS AND MOLECULES, 2nd Edition

EDITED BY: Michel Laurin, Gilles Didier and Rachel C. M. Warnock

PUBLISHED IN: Frontiers in Genetics and Frontiers in Ecology and Evolution

frontiers Research Topics

Frontiers eBook Copyright Statement

The copyright in the text of individual articles in this eBook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this eBook is the property of Frontiers.

Each article within this eBook, and the eBook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this eBook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or eBook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-88976-666-6
DOI 10.3389/978-2-88976-666-6

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

TIMETREES: INCORPORATING FOSSILS AND MOLECULES, 2nd Edition

Topic Editors:
Michel Laurin, UMR7207 Centre de recherche sur la paléobiodiversité et les paléoenvironnements (CR2P), France
Gilles Didier, UMR5149 Institut Montpelliérain Alexander Grothendieck (IMAG), France
Rachel C. M. Warnock, ETH Zürich, Switzerland

Publisher's note: In this 2nd edition, the following article has been added: Laurin M, Didier G and Warnock RCM (2022) Editorial: Timetrees: Incorporating fossils and molecules. Front. Genet. 13:937763. doi: 10.3389/fgene.2022.937763

Citation: Laurin, M., Didier, G., Warnock, R. C. M., eds. (2022). Timetrees: Incorporating Fossils and Molecules, 2nd Edition. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88976-666-6

Table of Contents

frontiers | Frontiers in Genetics

TYPE Editorial
PUBLISHED 30 June 2022
DOI 10.3389/fgene.2022.937763

Check for updates

OPEN ACCESS

EDITED AND REVIEWED BY
Carlos G. Schrago,
Federal University of Rio de Janeiro, Brazil

*CORRESPONDENCE
Michel Laurin,
michel.laurin@mnhn.fr

SPECIALTY SECTION
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Genetics

RECEIVED 06 May 2022
ACCEPTED 16 June 2022
PUBLISHED 30 June 2022

CITATION
Laurin M, Didier G and Warnock RCM (2022), Editorial: Timetrees: Incorporating fossils and molecules. *Front. Genet.* 13:937763. doi: 10.3389/fgene.2022.937763

COPYRIGHT
© 2022 Laurin, Didier and Warnock. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Editorial: Timetrees: Incorporating fossils and molecules

Michel Laurin[1]*, Gilles Didier[2] and Rachel C. M. Warnock[3]

[1]UMR7207 Centre de recherche en paléontologie - Paris (CR2P), Paris, France, [2]IMAG, Univ Montpellier, CNRS, Montpellier, France, [3]ETH Zürich, Zurich, Switzerland

KEYWORDS

molecular dating, tip-dating, fossil record, birth-and-death models, geological age

Editorial on the Research Topic
Timetrees: Incorporating fossils and molecules

Calibrating phylogenies to time is central to addressing many questions in evolutionary biology and macroevolution, such as the timing and dynamics of evolutionary radiations (e.g., Brocklehurst, 2017; Ascarrunz et al., 2019; Didier and Laurin, 2020) and of mass extinction events and their possible environmental causes (e.g., Allen et al., 2019; Didier and Laurin, 2021). The fossil record once provided our only source for establishing a timeline for evolution (Romer, 1966), but the incompleteness of this record and its non-uniformity in space and time limit the precision of divergence time estimates (Laurin, 2012; Heath et al., 2014; Warnock et al., 2017; Didier and Laurin, 2020). Molecular dating, which combines evidence from the geological and molecular records, can generate a much more complete and precise timeline of events (e.g., Sauquet, 2013; Magallón, 2020). This Research Topic focuses on recent advances in methodology, outstanding challenges, and the application of molecular and paleontological dating methods to empirical case studies across the Tree of Life.

Marshall reviews paleontological approaches to estimate divergence times, pointing out the many difficulties arising from this task. Though minimum ages are quite straightforward to infer from the fossil record, maximum age constraints are not so easy to establish. A first point to keep in mind is that the fossil record informs only about the first (fossilized) apomorphy and not the actual divergence time. Other major issues arise from the fact that the fossil recovery rate is not homogeneous and varies substantially over time and space. Marshall discusses various approaches to deal with these difficulties and shows some examples of paleontological dating.

Matschiner performs simulations in order to assess the influence of selective sampling of fossils or extant species on the accuracy of divergence times inferred under the Fossilized Birth-Death (FBD) model. He observes that non-uniform sampling of fossils or extant taxa leads to biased estimates of node ages obtained from the FBD model, notably in the case where the fossil record is reduced to the oldest fossil of each

branch. Another node dating approach, called CladeAge (Matschiner et al., 2016), shows better behavior in the presence of selective sampling of taxa in simulated data (but see Zhang et al., 2016).

Barido-Sottani et al. use simulations to examine the impact of fossil age uncertainty on trees recovered using the FBD model for fully extinct clades. They show that fixing fossil ages to a point age within the known range of stratigraphic age uncertainty produces incorrect estimates of both topology and divergence times. They also illustrate the impact of different approaches to handling fossil age uncertainty on parameter estimates among a group of Paleozoic crinoids. They further demonstrate that best solution is to explicitly model fossil age uncertainty.

Guindon provides a general presentation of molecular dating methods based on various assumptions (namely, strict, not-so-strict, uncorrelated and autocorrelated, relaxed clock models). He next reviews several approaches to calibrate clock models, mainly based on fossil records. After a brief presentation of how to process fossil for use in this context, he presents and discusses various model-based calibration methods, pointing out some issues in using the FBD model.

Powell et al. assess the advantages and drawbacks of secondary calibrations (which are molecular estimates of divergence times obtained in previous studies) compared to more distant primary (i.e., paleontological or geological) calibrations. This is timely because for many taxa with a poor fossil record (typically those containing organisms lacking a mineralized skeleton), calibration can be performed only through one of these alternatives. They find that distant primary calibrations provide better precision, but note that secondary calibrations remain useful.

Lozano-Fernandez et al. explore hypotheses about the geological context surrounding the colonization of land by arachnids. They generate a large dated tree of arachnids based genome-scale sequence data and a suite of rigorously assessed node calibrations. The origin of arachnids is dated to the Cambrian or Early Ordovician, indicating that terrestrialization occurred within this interval. This is followed by a rapid radiation of the group, coincident with elevated rates of molecular evolution. The authors suggest that the outstanding discrepancy between molecular estimates for the origin of crown group arachnids and the first appearance of body fossils belonging to this group can be attributed to incompleteness of the early terrestrial record.

Marjanović highlights problems associated with obtaining reliable time calibrations for node dating, caused by rapid progress in paleontology, thus rendering the few compilations (e.g., Benton et al., 2015) of such calibration constraints soon out of date, as more fossils are discovered or the information is updated. But worse, some molecular studies copy such constraints from previous molecular studies that had not necessarily used the most recent paleontological literature. These problems are illustrated through a detailed analysis of the 30 calibrations used to produce the largest available vertebrate timetree (Irisarri et al., 2017).

Pardo et al. assess the problems in obtaining reliable ages for three main crown-clades of limbed vertebrates (Tetrapoda, Lissamphibia and Amniota) to calibrate molecular clocks. They show that whereas much emphasis has been placed recently on documenting the age of fossils and providing synapomorphies that prove that they belong to a given clade (Parham et al., 2012), the main problem with deep tetrapod nodes is that the phylogeny is controversial and that various alternatives imply different ages for these clades.

Springer et al. review evolutionary models for the diversification of placental mammals, which differ from each other in the proposed timing of the evolutionary radiation of crown-placentals relative to the K/Pg boundary. At one extreme, this whole radiation may have started soon after the K/Pg boundary and proceeded very quickly, whereas at the other end of the spectrum, this radiation started around the mid-Cretaceous. Many problems (e.g., establishing homology of molecular sequences, taxonomic affinities of fossils and validity of the morphological clock) affect some or all of the three main dating methods (node-, tip-, and fossilized birth-death dating).

Celik and Phillips examine incongruence in the phylogeny of mammals based on different anatomical regions. This incongruence is attributed to convergent and correlated character evolution within ecologically similar but phylogenetically distinct groups. The authors develop a metric (the maximum parsimony disadvantage score) that allows us to identify homoplasy within anatomical partitions. They find that within mammals, cheek teeth and shoulder girdle characters have high potential to mislead phylogenetic inference due to non-phylogenetic covariance within these regions. These results have implications for assessing the placement of mammal fossils and consequently their inclusion in molecular dating studies.

Finally, Paterson et al. re-examine the monophyly of pinnipeds, which were widely believed to be diphyletic from the 1960s to the 1980s, and to assess parallel evolution within the group. Their Bayesian (as well as parsimony) analyses confirm pinniped monophyly but also demonstrate a surprising amount of parallel evolution in characters that had previously been interpreted as pinniped synapomorphies. These

include dental and limb bone characters relating to homodonty and aquatic locomotion, respectively. New tip-dating analyses date the divergence between pinnipeds and musteloids to about 45 Ma.

Together, these studies illustrate the utility of timetrees in addressing fundamental questions about evolution, as well as underscoring the need to apply a rigorous approach to select calibrations, models and prior parameters.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Allen, B. J., Stubbs, T. L., Benton, M. J., and Puttick, M. N. (2019). Archosauromorph extinction selectivity during the triassic-jurassic mass extinction. *Palaeontology* 62, 211–224. doi:10.1111/pala.12399

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 Evol. Biol.* 19, 117. doi:10.1186/s12862-019-1436-x

Benton, M., Donoghue, P., Vinther, J., Asher, R., Friedman, M., and Near, T. (2015). Constraints on the timescale of animal evolutionary history. *Palaeontol. Electron* 18, 1–106. doi:10.26879/424

Brocklehurst, N. (2017). Rates of morphological evolution in captorhinidae: an adaptive radiation of permian herbivores. *PeerJ* 5, e3200. doi:10.7717/peerj.3200

Didier, G., and Laurin, M. (2021). Distributions of extinction times from fossil ages and tree topologies: the example of mid-permian synapsid extinctions. *PeerJ* 9, e12577. doi:10.7717/peerj.12577

Didier, G., and Laurin, M. (2020). Exact distribution of divergence times from fossil ages and tree topologies. *Syst. Biol.* 69, 1068–1087. doi:10.1093/sysbio/syaa021

Heath, T. A., Huelsenbeck, J. P., and Stadler, T. (2014). The fossilized birth-death process for coherent calibration of divergence-time estimates. *Proc. Natl. Acad. Sci. U.S.A.* 111, E2957–E2966. doi:10.1073/pnas.1319091111

Irisarri, I., Baurain, D., Brinkmann, H., Delsuc, F., Sire, J.-Y., Kupfer, A., et al. (2017). Phylotranscriptomic consolidation of the jawed vertebrate timetree. *Nat. Ecol. Evol.* 1, 1370–1378. doi:10.1038/s41559-017-0240-5

Laurin, M. (2012). Recent progress in paleontological methods for dating the tree of life. *Front. Gene.* 3. doi:10.3389/fgene.2012.00130

Magallón, S. (2020). "Principles of molecular dating," in *The Molecular Evolutionary Clock*. Editor S. Y. Ho (Cham, Switzerland: Springer Nature Switzerland), 67–81. doi:10.1007/978-3-030-60181-2_5

Matschiner, M., Musilová, Z., Barth, J. M. I., Starostová, Z., Salzburger, W., Steel, M., et al. (2016). Bayesian phylogenetic estimation of clade ages supports trans-atlantic dispersal of cichlid fishes. *Syst. Biol.* 66, syw076–22. doi:10.1093/sysbio/syw076

Parham, J. F., Donoghue, P. C. J., Bell, C. J., Calway, T. D., Head, J. J., Holroyd, P. A., et al. (2012). Best practices for justifying fossil calibrations. *Syst. Biol.* 61, 346–359. doi:10.1093/sysbio/syr107

Romer, A. (1966). *Vertebrate paleontology*. University of Chicago press Chicago.

Sauquet, H. (2013). A practical guide to molecular dating. *Comptes Rendus Palevol* 12, 355–367. doi:10.1016/j.crpv.2013.07.003

Warnock, R. C. M., Yang, Z., and Donoghue, P. C. J. (2017). Testing the molecular clock using mechanistic models of fossil preservation and molecular evolution. *Proc. R. Soc. B* 284, 20170227. doi:10.1098/rspb.2017.0227

Zhang, C., Stadler, T., Klopfstein, S., Heath, T. A., and Ronquist, F. (2016). Total-evidence dating under the fossilized birth-death process. *Syst. Biol.* 65, 228–249. doi:10.1093/sysbio/syv080

frontiers in Genetics

ORIGINAL RESEARCH
published: 31 October 2019
doi: 10.3389/fgene.2019.01064

Selective Sampling of Species and Fossils Influences Age Estimates Under the Fossilized Birth–Death Model

Michael Matschiner[1,2]*

[1] Department of Palaeontology and Museum, University of Zurich, Zurich, Switzerland, [2] Centre of Ecological and Evolutionary Synthesis, Department of Biosciences, University of Oslo, Oslo, Norway

OPEN ACCESS

Edited by:
Michel Laurin, UMR7207 Centre de recherche sur la paléobiodiversité et les paléoenvironnements (CR2P), France

Reviewed by:
David Marjanović, Museum für Naturkunde, Germany
David Buckley, Autonomous University of Madrid, Spain
Gilles Didier, UMR5149 Institut Montpelliérain Alexander Grothendieck (IMAG), France

***Correspondence:**
Michael Matschiner, michaelmatschiner@mac.com

Specialty section:
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Genetics

Received: *06 June 2019*
Accepted: *03 October 2019*
Published: *31 October 2019*

Citation:
Matschiner M (2019) Selective Sampling of Species and Fossils Influences Age Estimates Under the Fossilized Birth–Death Model. Front. Genet. 10:1064. doi: 10.3389/fgene.2019.01064

The fossilized birth–death (FBD) model allows the estimation of species divergence times from molecular and fossil information in a coherent framework of diversification and fossil sampling. Some assumptions of the FBD model, however, are difficult to meet in phylogenetic analyses of highly diverse groups. Here, I use simulations to assess the impact of extreme model violations, including diversified sampling of species and the exclusive use of the oldest fossils per clade, on divergence times estimated with the FBD model. My results demonstrate that selective sampling of fossils can produce dramatically overestimated divergence times when the FBD model is used for inference, due to an interplay of underestimates for the model parameters net diversification rate, turnover, and fossil-sampling proportion. In contrast, divergence times estimated with CladeAge, a method that uses information about the oldest fossils per clade together with estimates of sampling and diversification rates, are accurate under these conditions. Practitioners of Bayesian divergence-time estimation should therefore ensure that the dataset conforms to the expectations of the FBD model, or estimates of sampling and diversification rates should be obtained *a priori* so that CladeAge can be used for the inference.

Keywords: phylogeny, bayesian inference, divergence-time estimation, fossil, diversified sampling, BEAST 2, fossilized birth–death, CladeAge

INTRODUCTION

With increases in the sizes of molecular datasets and improvements to inference methodology, our understanding of the timeline of evolution has grown tremendously over the past two decades. One of the most significant methodological developments for the estimation of divergence times has been the fossilized birth–death (FBD) model (Stadler, 2010; Heath et al., 2014), a phylogenetic framework that combines the two processes of species diversification and fossil sampling. By using fossils as tips or sampled ancestors in the phylogeny, the FBD model is able to estimate the probability that species fossilize before their extinction and it accounts for this probability in the inference. The FBD model thus overcomes a limitation of the commonly applied "node dating" approach in which only the oldest fossils of some clades are used to define constraints on the divergence times among these clades: As fossils can provide reliable evidence for the minimum age of a clade but are only vaguely informative about maximum ages when the sampling process is not included in the model (Benton and Donoghue, 2007), the placement of maximum ages in node dating is often controversial, even though it is essential

for the inference (Marjanović and Laurin, 2007; Warnock et al., 2012). Because age constraints with minimum and maximum ages are not required with the FBD model, estimates obtained with this model do not depend on the controversial specification of those constraints and may thus be generally more reliable.

The FBD model was first available for inference in the program DPPDIV (Heath et al., 2014), allowing the estimation of divergence times from a molecular dataset and a user-provided tree with a fixed topology. The dependence on a known topology has been relaxed in subsequent implementations of the model in BEAST 2 (Gavryushkina et al., 2014; Gavryushkina et al., 2017; Bouckaert et al., 2019), MrBayes (Ronquist et al., 2012; Zhang et al., 2016), and RevBayes (Höhna et al., 2016) all of which also allow the inference of fossil positions based on morphological information instead of requiring the user to know their positions *a priori*. The FBD model has further matured with the integration of stratigraphic-range information and different speciation modes (Silvestro et al., 2018; Stadler et al., 2018), time-variable diversification and sampling (Gavryushkina et al., 2014), coalescent processes (Ogilvie et al., 2018), and the estimation of divergence times without assuming molecular or morphological clocks (Didier and Laurin, 2018).

The accuracy of age estimates obtained with the FBD model has been tested with simulations in multiple studies that all confirmed reliable inference (Gavryushkina et al., 2014; Heath et al., 2014; Zhang et al., 2016; Matschiner et al., 2017). The simulations in these studies mostly did not violate the assumptions of the FBD model, which include that either all simulated species and fossils or a randomly selected subset of these are used for the inference. A "complete sampling" scheme (or at least nearly complete sampling) for species and fossils was also applied in a number of empirical studies using the FBD, to estimate divergence times among, e.g., bears (Heath et al., 2014), penguins (Gavryushkina et al., 2017), and beech trees (Renner et al., 2016); however, most empirical datasets may not be completely sampled. Instead, extant species may often be missing from phylogenetic datasets, for example due to limited availability of sequence data. The incompleteness of taxon sets is usually amplified in phylogenetic analyses of larger clades, where the inclusion of all species would be computationally infeasible or the generation of molecular data for all species would be too costly. Moreover, the selection of species for such analyses may rarely be uniformly random and instead "diversified" sampling of extant species may be more common (Höhna et al., 2011; Höhna, 2014), because researchers often aim to include representatives of each major group within the studied clade (e.g. Meredith et al., 2011; Jarvis et al., 2014; dos Reis et al., 2015; Barba-Montoya et al., 2018; Musilova et al., 2019).

Like the sampling of extant species in empirical analyses, the inclusion of fossils may also often be neither complete nor random. For larger clades with high preservation potential, complete sampling of fossils may not be possible due to their sheer numbers, and instead of applying random sampling of fossils as an alternative, researchers may want to ensure that the earliest records are included in the dataset. Thus, both the sampling of extant species and of fossils is probably selective in most empirical datasets used in analyses with the FBD model; however, the degree to which the FBD model is robust to these model violations has so far not been tested with simulations.

As another alternative to node dating, Matschiner et al. (2017) developed CladeAge, an approach that estimates divergence times based on information about the oldest fossils of clades, in combination with estimates of sampling and diversification rates. Specifically, CladeAge uses this information to derive probability distributions for the ages of individual clades under a model of time-homogeneous diversification and fossil sampling, and those probability distributions are then used as calibration densities in phylogenetic divergence-time estimation. By assuming Poisson processes for diversification and fossil sampling, the derivation of probability distributions in CladeAge is essentially based on the FBD model, but the processes are truncated at the first sampling event. As the term "FBD model" is commonly understood to describe the process continued to the present, I will use this term as a synonym for the implementations of this untruncated model (e.g., in BEAST 2, MrBayes, or RevBayes), and I will use the term "CladeAge" to refer to the combined approach of model-based quantification of calibration densities per clade and the use of these densities for divergence-time estimation. Like the FBD model, the performance of CladeAge has been tested with simulations that confirmed reliable inference (Matschiner et al., 2017); however, as in the case of the FBD model, these simulations matched the expectations of the method in terms of sampling of extant species and fossils.

Here, I test the performance of both the FBD model and CladeAge in scenarios of model violations that include a strict diversified sampling scheme for extant species and the exclusive use of the oldest fossils per clade (the latter only violates the FBD model but matches the assumptions of CladeAge). While these scenarios are probably more extreme than the model violation in almost all empirical analyses, I expect that the results will provide valuable clues about the robustness of the inference with the two approaches.

METHOD

Simulations

I used forward simulation to generate phylogenetic trees as in Matschiner et al. (2017), with branch lengths corresponding to time. In these simulations, I set the age of the first divergence to 100 time units in the past and applied a constant-rate birth–death process (Gernhard, 2008) with cladogenetic speciation (hereafter: "speciation") rate $\lambda = 0.12$ and extinction rate $\mu = 0.06$ to generate the tree. The net diversification rate ($\lambda-\mu$) was thus 0.06 and the turnover (λ/μ) was 0.5, and these rates applied to all branches of the tree. I repeated this simulation until 20 trees were found that had between 4,000 and 5,000 extant species, and I discarded all trees that did not fulfill this condition. Age and species richness of the simulated phylogenies were thus roughly comparable to those of placental mammals (Meredith et al., 2011; Stadler, 2011) if we assume one time unit to correspond to one million years. I added a simulated fossil record to all branches of the trees, assuming a homogeneous Poisson process of fossil sampling with sampling rate $\psi = 0.01$ and thus a fossil-sampling proportion of $\psi/(\mu+\psi) = 0.143$ (Gavryushkina et al., 2014). The recorded fossil ages were assumed to be known without error.

To mimic the information content of empirical datasets in which fossils are assigned to extant clades, but no morphological

data are available to infer interrelations, extinct branches were pruned from all simulated phylogenies and their fossil records were transferred to the ancestral branch in the reconstructed phylogeny from which they had diverged. Each internal branch thus represented the stem of an extant clade and the fossils assigned to the branch can be interpreted as the stem group of that clade. The ages of these fossils did not necessarily fall into the time period covered by the branch, but, as is the case for stem-group fossils, could postdate the origin of the crown group and thus the end of the stem branch. I then selected 50 extant species from each tree according to the strict diversified sampling scheme of Höhna et al. (2011), meaning that first, the time point in the phylogeny was identified at which 50 branches with extant descendants existed, and second, one of these descendants is sampled at random for each of these 50 branches. The tree was then reduced to the branches connecting these 50 species, the "diversified tree" (**Figure 1**). As a consequence of this sampling scheme, the diversified tree is guaranteed to include the 49 oldest divergences among extant species but none of the divergences that are younger than those 49. In addition to the diversified sampling scheme, I separately applied the random sampling scheme, sampling 50 extant species uniformly at random with the only requirement that at least one extant species was sampled from both sides of the root. For both the diversified tree and the randomly sampled tree, the fossil records of pruned branches were once again transferred to the corresponding ancestral branches remaining in the phylogeny. As in Matschiner et al. (2017), nucleotide sequences of a length of 3,000 base pairs (bp) were simulated along each tree according to the unrestricted empirical codon model of Kosiol et al. (2007) with a mean substitution rate set to $3 \times 10\text{-}3^{-3}$ and a rate-variance parameter of $9 \times 10\text{-}6^{-6}$.

Divergence-Time Estimation With the FBD Model and Cladeage

For each of the 20 datasets simulated with the diversified sampling scheme and the 20 datasets generated with random sampling, I estimated divergence times among the 50 species with both the FBD model (Gavryushkina et al., 2014; Heath et al., 2014) and CladeAge (Matschiner et al., 2017). I performed these analyses either with the FBD model implementation in the SA package v.1.1.7 or the CladeAge implementation in the CA package v.1.3.0, both of which are add-ons for the software BEAST 2 (Bouckaert et al., 2019) (of which I used v.2.4.2). As starting trees in analyses with the FBD model, I prepared two modified versions of each diversified and randomly sampled tree in which either only the oldest fossil per branch or all fossils of each branch were inserted as extinct tips and connected to their respective branches *via* newly added branches. The topology of extant species was fixed to their true topology; however, as in Matschiner et al. (2017), this was done with "CladeConstraint" topology constraints (Gavryushkina et al., 2014) in the case of FBD analyses, so that fossils were allowed to attach either on the stem or in the crown of the clade to which they were assigned.

Using four different settings ("Set1" to "Set4") in FBD analyses, the parameters of the FBD model implementation in BEAST 2, net diversification rate ($\lambda-\mu$), turnover (μ/λ), and fossil-sampling proportion ($\psi/(\mu+\psi)$) (Gavryushkina et al., 2014), were either fixed (Set1,3) to the true values used in simulations (0.06, 0.5, and 0.143; see above) or estimated (Set2–4; **Table 1**), and all three parameters were assumed constant throughout the tree. When these parameters were estimated, uniform priors were used as constraints for each of them, centered on the true values and with lower and upper boundaries corresponding to 50% and 150% of

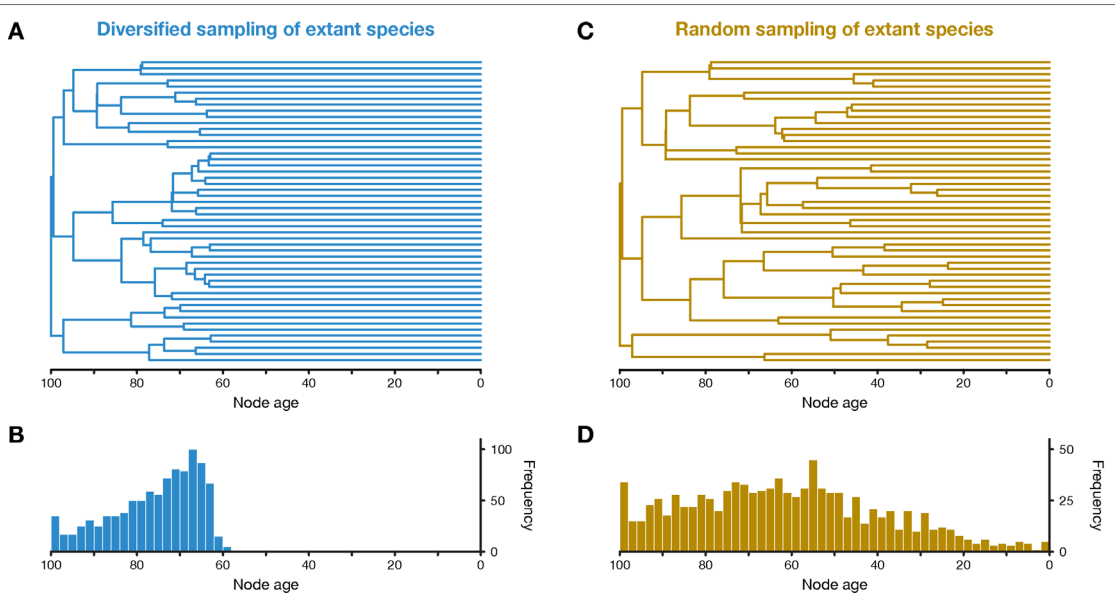


FIGURE 1 | Simulation of phylogenetic trees. **(A)** One out of 20 diversified trees simulated for this study, after applying the diversified sampling scheme to sample 50 out of 4,000 to 5,000 extant species. **(B)** Distribution of node ages in all 20 diversified trees. **(C)** One out of 20 randomly sampled trees. **(D)** Distribution of node ages in all 20 randomly sampled trees.

the true value, respectively. The probability of sampling extant species, ρ, was fixed to the true proportion of sampled species (thus, 50 divided by the number of extant species in the full simulated tree; a value between 0.01 and 0.0125). In most analyses with the FBD model (Set1–3), the fossil records were reduced to the oldest fossil per branch, but an additional set of analyses (Set4), in which diversification and sampling parameters were estimated, was also conducted with all fossils of each branch (**Table 1**). For computational reasons, the setting Set4 was only applied to datasets generated with diversified sampling of extant species.

In contrast to the FBD model, CladeAge does not assume that the diversification parameters of the tree-generating process are identical to those of the fossil-generating process; thus, these parameters need to be specified separately for the fossil-generating process. While they could also be specified differently for each clade, I here always used the same values for all clades. The fossil-generating process is parameterized as in the FBD model implementation in BEAST 2 with net diversification rate ($\lambda-\mu$) and turnover (μ/λ), but instead of the fossil-sampling proportion ($\psi/(\mu+\psi)$), the sampling rate (ψ) is used. Analogous to the analyses with the FBD model, I specified the three parameters either exactly according to their true values (Set5) or applied confidence intervals with lower and upper boundaries set to 50% to 150% of the true parameter values (Set6). The tree-generating process, on the other hand, was in all CladeAge analyses assumed to be the birth–death process (Gernhard, 2008) and uninformative uniform priors were used for the two parameters of the birth–death process, the net diversification rate ($\lambda-\mu$; constrained to $\lambda-\mu\in[0,1{,}000]$) and the turnover ($\mu/\lambda$; constrained to $\mu/\lambda\in[0,1]$). Conforming to the assumptions of CladeAge, fossil records were reduced to the oldest fossil per branch in all CladeAge analyses and fossils were reused for parental branches if these did not have a fossil record on their own (see **Figure 2A** in Matschiner et al., 2017).

In all analyses, sites of the sequence alignment were grouped into three different partitions according to codon position, and the reversible-jump-based substitution model of Bouckaert et al. (2013) was applied to each of them. Branch-rate variation was modeled with the uncorrelated clock model of Drummond et al. (2006). All analyses were set to use 100 million Markov-chain Monte Carlo iterations but were resumed after finishing if the chain had not reached stationarity. To assess stationarity, effective sample sizes of all model parameters (ESS) were calculated with the coda R package v.0.19 (Plummer et al., 2006) and considered sufficient if all of them were above 200. After reaching stationarity, the length of the burnin period was determined visually from trace plots generated with Tracer v.1.7.1 (Rambaut et al., 2018) and the minimum number of iterations required to reach stationarity post-burnin was calculated, again with the coda R package. Finally, the results of each set of 20 analyses that shared identical simulation (diversified or random sampling) and inference (Set1–Set6) settings were pooled before interpretation. The accuracy of divergence-time estimates was quantified as the proportion of 95% highest-posterior-density (HPD) intervals (across the 20 analyses) that included the true node age. Without model violations, an accuracy of 95% would be expected. All BEAST 2 analyses made use of the BEAGLE computing library v.4.1 (Ayres et al., 2012) and were carried out with three threads on dual eight-core Intel Xeon E5-2670 (Sandy Bridge-EP) CPUs running at 2.6 GHz.

RESULTS

Simulations

The 20 simulated trees had on average 4,490.6 (standard deviation, sd: 226.0) extant species. After applying the diversified sampling scheme, all terminal branches were longer than 57.9–68.3 time units (mean across trees: 63.2; sd: 2.2) and all divergences were thus concentrated within the first 31.7–42.1 time units of the diversified tree (**Figures 1A, B**). This was reflected by the γ statistic of the constant-rates test of Pybus and Harvey (2000), which was highly negative for all 20 diversified trees (mean: −10.1; sd: 0.2). Qualitatively, the diversified trees appeared similar in shape to time-calibrated phylogenies of larger clades based on genomic datasets, such as the phylogeny of birds by Jarvis et al. (2014) or the phylogeny of spiny-rayed fishes by Alfaro et al. (2018). In contrast, random sampling of extant species produced trees with a wider distribution of node ages and shorter terminal branches (**Figures 1C, D**). The γ statistic

TABLE 1 | Settings, results, and run statistics for analyses of simulated datasets with the FBD model and CladeAge. The simulated datasets were either based on diversified or random sampling of extant species. Accuracy (percentage of estimates within 95% HPD interval), root-mean-square-deviation between true ages and mean node-age estimates (RMSD), iterations to stationary, time per iteration, and time to stationarity are averaged over the analyses of 20 simulated datasets.

Species sampling	Inference setting	Method	Diversification parameters	Fossil-sampling parameters	Fossils used	Accuracy	RMSD	Iterations to stationary	Time per iteration	Time to stationarity
Diversified	Set1	FBD	Fixed	Fixed	Oldest	78.5%	6.0	76.7M	0.643 ms	13.6 h
Diversified	Set2	FBD	Estimated	Estimated	Oldest	0.6%	55.8	101.1M	1.012 ms	25.4 h
Diversified	Set3	FBD	Estimated	Fixed	Oldest	5.5%	38.1	54.1M	0.967 ms	13.8 h
Diversified	Set4	FBD	Estimated	Estimated	All	89.4%	6.6	>1,200.0M	3.224 ms	>1,074.5 h
Diversified	Set5	CladeAge	Fixed	Fixed	Oldest	90.3%	4.6	49.2M	0.594 ms	8.5 h
Diversified	Set6	CladeAge	Estimated	Estimated	Oldest	91.0%	4.6	58.9M	0.551 ms	8.6 h
Random	Set1	FBD	Fixed	Fixed	Oldest	90.7%	4.8	80.8M	0.578 ms	12.9 h
Random	Set2	FBD	Estimated	Estimated	Oldest	2.9%	69.2	53.8M	0.825 ms	13.1 h
Random	Set3	FBD	Estimated	Fixed	Oldest	6.3%	52.1	57.2M	0.627 ms	9.9 h
Random	Set5	CladeAge	Fixed	Fixed	Oldest	91.6%	4.6	39.3M	0.550 ms	5.8 h
Random	Set6	CladeAge	Estimated	Estimated	Oldest	91.5%	4.7	35.2M	0.523 ms	5.0 h

was closer to zero but still negative for randomly-sampled trees (mean: −6.8; sd: 0.6), as expected due a decline in the "pulled speciation rate" near the present in cases of incomplete sampling of extant species (Louca and Pennell, 2019). The simulated fossil records included between 672 and 800 fossils (mean across trees: 737.4; sd: 37.8) that attached to 52–59 branches of the diversified trees or 62–72 branches of the randomly-sampled trees. This increased number of branches with fossils in randomly-sampled trees can be explained by the smaller number of very short branches compared to diversified trees (see **Figures 1A, C**). The sequence alignments of 3,000 bp simulated for diversified trees contained between 2,741 and 2,890 (mean: 2,824.8; sd: 34.6) variable sites, out of which 1,737–2,141 (mean: 1,962.4; sd: 94.0) sites were parsimony-informative. For randomly-sampled trees, between 2,526 and 2,857 (mean: 2,722.1; sd: 73.9) sites were variable, including 1,579–2,183 (mean: 1,913.6; sd: 1,35.1) parsimony-informative sites.

Divergence-Time Estimation With the FBD Model and Cladeage

When all diversification and fossil-sampling parameters were fixed to the true values used in simulations (Set1), age estimates obtained with the FBD model were relatively accurate despite the model violations of diversified sampling and the reduction of the fossil record to the oldest fossils per branch. With these settings, 78.5% of the 95% HPD intervals contained the true node age (**Table 1**), and the mean age estimates appeared close to the true ages (root-mean-square deviation, RMSD: 6.0 time units) (**Figures 2A, 3A-C**). However, when diversification and fossil-sampling parameters were estimated instead of fixed to the true values (Set2), almost all ages were substantially overestimated (RMSD: 55.8). In this case, only 0.6% of the 95% HPD intervals included the true node age and every single mean node-age estimate was older than the true age (**Figure 2B**). The low accuracy of node ages was reflected by the estimates of the net diversification rate, the turnover, and the fossil-sampling proportion, all of which appeared at the lower boundaries of the uniform prior intervals used as constraints (**Figures 3D–F**). Fixing only the fossil-sampling proportion to the true value while estimating the diversification parameters (Set3) led to a moderate improvement in the node-age estimates, resulting in an accuracy of 5.5% and slightly lower mean node ages (RMSD: 38.1) (**Figure 2C**). The estimates of the net diversification rate and the turnover, however, remained near the lower prior boundary (**Figures 3G–I**). In contrast, the use of all fossils instead of only the oldest per branch (Set4) resulted in a much better accuracy of node-age estimates, namely 89.2% (RMSD: 6.6) (**Figure 2D**). In this set of analyses, estimates of the net diversification rate were centered close to the true value ($\lambda-\mu = 0.06$) (**Figure 3J**), and while the turnover and the fossil-sampling proportion appeared to be under- and overestimated, respectively, the posterior distributions of these estimates included the true values (**Figures 3K, L**).

The computational requirements of FBD analyses with all fossils, however, were far larger than those of all other analyses. Whereas the FBD analyses with only the oldest fossils per branch of the diversified tree required between 54.1 and 76.7 million MCMC iterations to stationarity and these completed within 13.6 to 25.4 h, all but one of the 20 FBD analyses with all fossils had not reached stationarity even after 1.2 billion MCMC iterations that took 1,074.5 h (45 days) (**Table 1**). The lowest ESS value after this number of iterations was 24.2, suggesting that around 10 billion iterations and a run time of around a year might be necessary to reach ESS values greater than 200 for all parameters in all analyses of Set4.

The analyses of datasets generated with random sampling of extant taxa produced results similar to those based on diversified sampling (**Figures 2E–G**). When all diversification and fossil-sampling parameters were fixed to their true values, node ages estimated with the FBD model were largely accurate (**Figure 2E**), with 90.7% of the 95% HPD intervals containing the true node age (**Table 1**). In contrast, allowing all model parameters, or only the diversification parameters, to be estimated led to a degree of node age overestimation that was even larger than in the results based on diversified sampling of extant species (RMSD: 69.2 and 52.1 without and with fixing the fossil-sampling proportion, respectively) (**Figures 2F, G**). Due to their great computational requirements, I did not conduct FBD analyses with all fossils for the randomly sampled trees. However, as these FBD analyses with all fossils of randomly sampled trees would not suffer from model violations due to selective sampling of species or fossils, I assume that they would provide accurate estimates of node ages and model parameters.

Analyses with CladeAge were not affected by the reduction of the fossil record to the oldest fossils as this reduction is already expected by CladeAge. Regardless of whether datasets had been generated with diversified or random sampling of extant species and whether the diversification and sampling parameters were specified exactly (Set5) or considered uncertain within intervals ranging from 50% to 150% of the true values (Set6), the accuracy of node-age estimates was above 90% (90.3–91.6%) and mean node-age estimates were close to the true values (RMSD: 4.6–4.7) (**Figure 4**). Run times with CladeAge were comparatively short; on average, between 35.2 and 58.9 million iterations were required to reach stationarity and these numbers of iterations completed within 5.0 to 8.6 h (**Table 1**).

DISCUSSION

My analyses of simulated data show that the FBD model can produce highly inflated age estimates when sampling of species and fossils is not complete or random but selective. Because of their great computational demand, I did not perform analyses in which species were sampled randomly (or completely) and all fossils were used; however, based on the results of previous studies (Gavryushkina et al., 2014; Matschiner et al., 2017), I assume that these analyses would have resulted in high accuracy close to 95%. This would mean that the decrease in accuracy of node-age estimates (to 89.2%) was minor when only the diversifed sampling of extant species was applied but the entire fossil record was used for calibration in analyses with setting Set4. A far more dramatic decrease in accuracy (down to 0.6%), along with substantial overestimation of node ages, resulted from

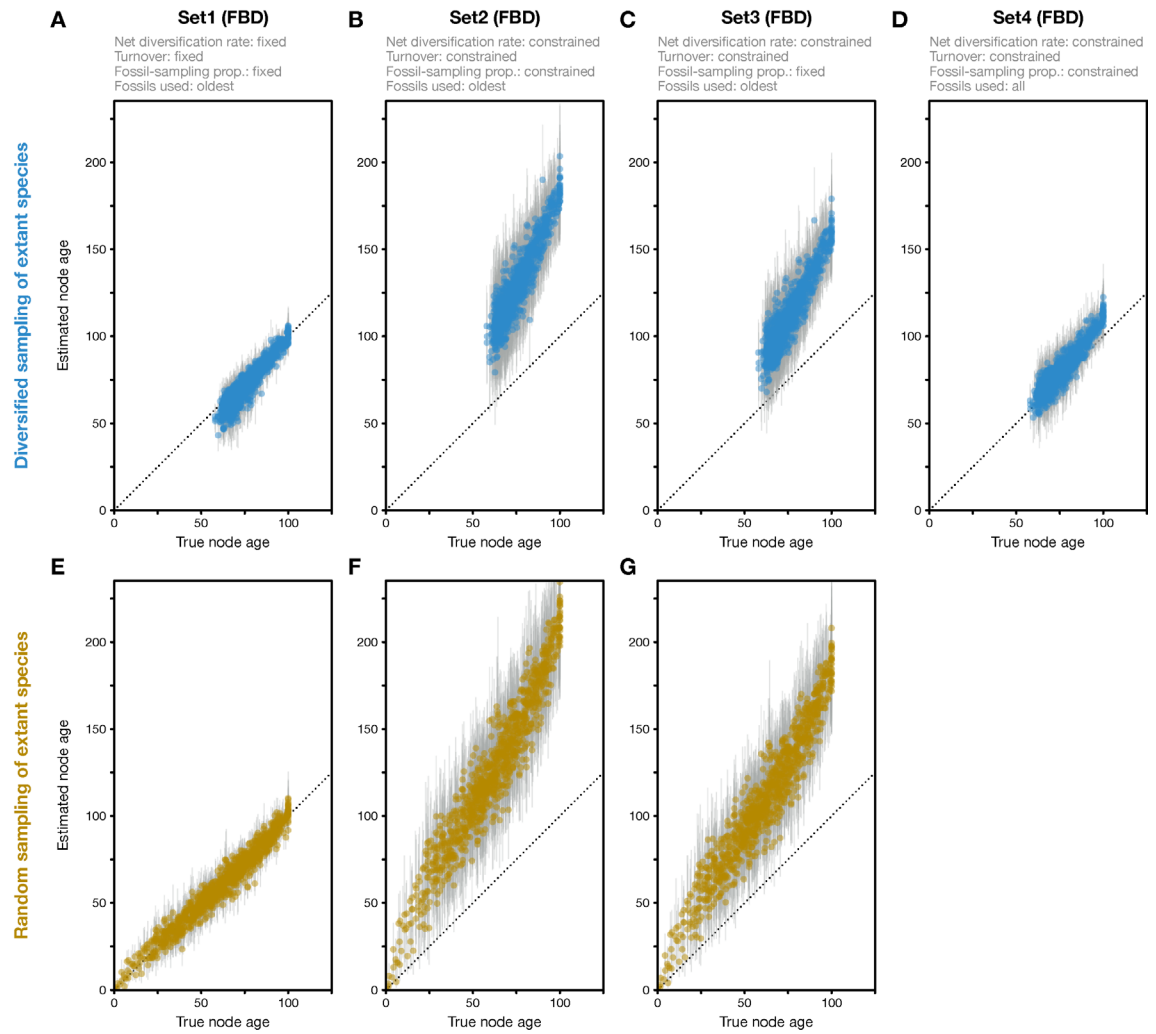


FIGURE 2 | Divergence times estimated with the FBD model. **(A)** Comparison of true node ages and node ages estimated in FBD analyses of simulated diversified trees, using inference setting Set1 in which diversification and sampling parameters were fixed to their true values. The dotted line marks the diagonal. **(B)** Node-age estimates for FBD analyses with setting Set2 in which net diversification, turnover, and fossil-sampling proportion were estimated. **(C)** Node-age estimates for FBD analyses with setting Set3 in which net diversification and turnover were estimated but the fossil-sampling proportion was fixed to its true value. **(D)** As B but for FBD analyses with setting Set4 in which all fossils of each branch were used instead of reducing the fossil records to the oldest fossils per branch. **(E–G)** As A–C but for datasets generated based on random sampling of extant species. Analyses with all fossils were conducted only for diversified trees, not for randomly sampled trees.

the reduction of the fossil record to the oldest fossils per branch in analyses with setting Set2. The comparison of the results obtained with settings Set2 and Set4 thus allows us to disentangle the effects of the two types of model violation and interpret how they may have led to the observed node-age overestimation.

First, the underestimated turnover observed in the FBD analyses with setting Set4 may be explained by the bottom-heavy shape of the diversified trees (**Figure 1**), a pattern that is opposite to that expected from high turnover, a concentration of divergences among extant species near the present. Underestimates of turnover, in turn, imply that the number of extinct branches is also underestimated, which could be responsible for overestimation of the fossil-sampling proportion in the analyses with setting Set4 based on unreduced fossil records. In the FBD analyses with setting Set2, however, the reduction of fossil records to the oldest fossils per branch may have counteracted the overestimation of the fossil-sampling proportion, leading even to strong underestimation of this proportion. The more accurate estimates of turnover in analyses of Set4 with all fossils, compared to those of Set2 with only the oldest fossils, are likely explained by the large number of additional extinct branches in the phylogenies of Set4 that support a higher extinction rate (μ) and thus a higher turnover (μ/λ).

As the comparison of results obtained with settings Set4 and Set2 shows, it is the selective sampling of the oldest fossils per branch that is responsible for most of the overestimation of node ages in analyses with setting Set2. Thus, it might be surprising that by fixing the sampling proportion to its true value in

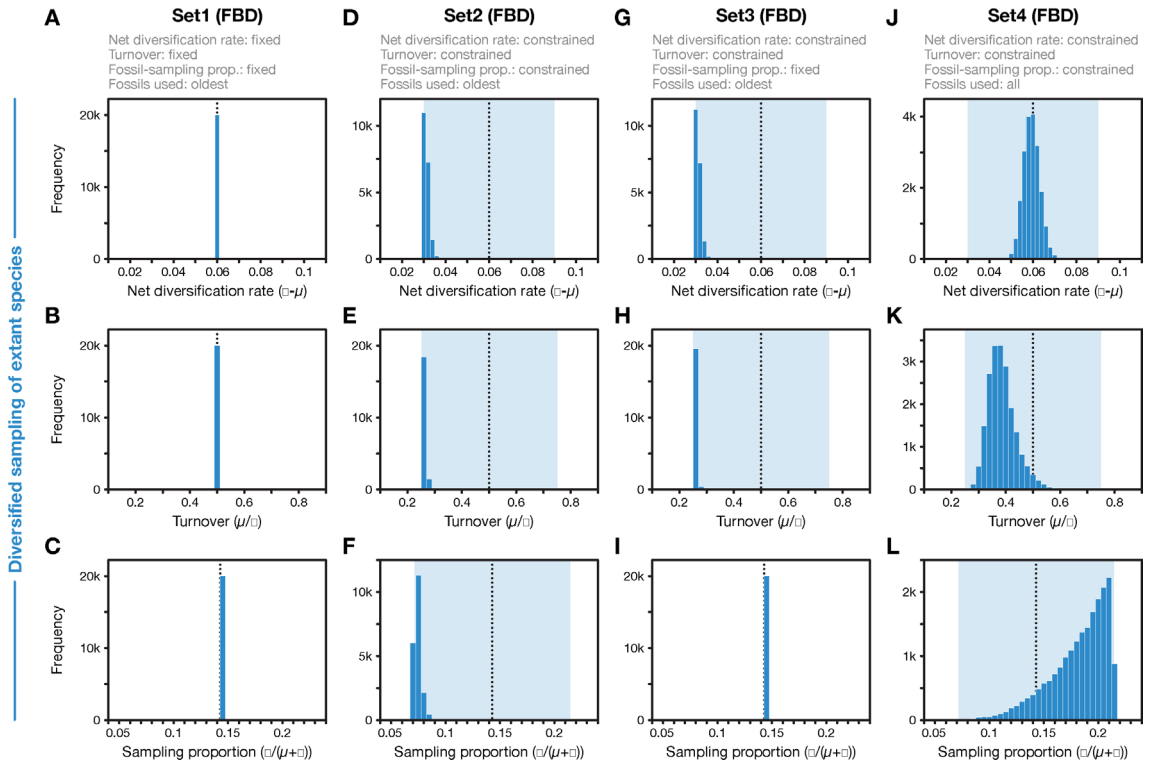


FIGURE 3 | Parameter values for net diversification rate, turnover, and fossil-sampling proportion, estimated with the FBD model. **(A–C)** In FBD analyses with inference setting Set1, diversification and sampling parameters were fixed to their true values, indicated with vertical dotted black lines. **(D–F)** In FBD analyses with setting Set2, net diversification, turnover, and fossil-sampling proportion were estimated with uniform prior intervals shown in light blue; the histograms show the posterior distributions. **(G–I)** As D–F but for FBD analyses with setting Set3 in which the fossil-sampling proportion was fixed to its true value. **(J–L)** As D–F but for FBD analyses with setting Set4 in which all fossils of each branch were used instead of reducing the fossil records to the oldest fossils per branch. All results shown here were obtained with the 20 simulated diversified trees; those obtained with random sampling of extant species were nearly identical (for **A–I**).

analyses with setting Set3, only moderate improvements in age estimates are gained. How can the selective sampling of fossils impact age estimates if not through the sampling proportion? The answer probably lies in the indirect relationship between fossil-sampling proportion and fossil-sampling rate ψ, which is influenced by the extinction rate μ, as the fossil-sampling proportion is ψ/(μ+ψ). If, as is roughly the case in the analyses with settings Set2 and Set3, both the net diversification rate (λ−μ) and the turnover (μ/λ) are estimated as half of their true values (0.06 and 0.5, respectively), this means that the extinction rate is implicitly estimated as μ = 0.01 (and the speciation rate is implicitly estimated as λ = 0.04). The estimated extinction rate is thus only a sixth of the true value used in the simulations (μ = 0.06; see above). With an estimated extinction rate μ = 0.01 and the fossil-sampling proportion fixed at ψ/(μ+ψ) = 0.143, the fossil-sampling rate is ψ = 0.00167, also a sixth of the true value used in the simulations. Thus, despite fixing the fossil-sampling proportion in setting Set3, the fossil-sampling rate ψ that is implicit in the model remains substantially underestimated. With underestimated fossil-sampling rates, the expected waiting times between clade ages and their first fossil records increase, and as a result, older trees become more probable under the FBD model, leading to the observed overestimated node ages. However, the underestimation of the fossil-sampling rate with setting Set3 is not as severe as in the analyses with setting Set2 (with the fossil-sampling proportion estimated around ψ/(μ+ψ) = 0.071 in those analyses, the implicitly estimated fossil-sampling rate is ψ = 0.00077), which likely explains the modest improvements in node-age estimates between analyses with setting Set2 and those with Set3. In the analyses with setting Set1, on the other hand, fixing of all three explicit model parameters net diversification rate (λ−μ), turnover (μ/λ), and fossil-sampling proportion (ψ/(μ+ψ)) also fixes the implicit model parameters speciation rate (λ), extinction rate (μ), and fossil-sampling rate (ψ) to the true values used in the simulations, explaining the largely accurate age inference in those analyses.

The issues highlighted by my analyses of simulated data suggest that the application of the FBD model to larger empirical datasets may often be problematic. To investigate divergence times of species-rich clades with large fossil records like placental mammals, birds, or teleost fishes with the FBD model, researchers would need to decide between the options of complete, random, or selective fossil sampling, all of which are not ideal. Complete sampling of the fossil records of these clades would entail the use of thousands of fossils, but as my analyses with setting Set4 showed, even hundreds of fossils, in combination with rather small molecular datasets, require

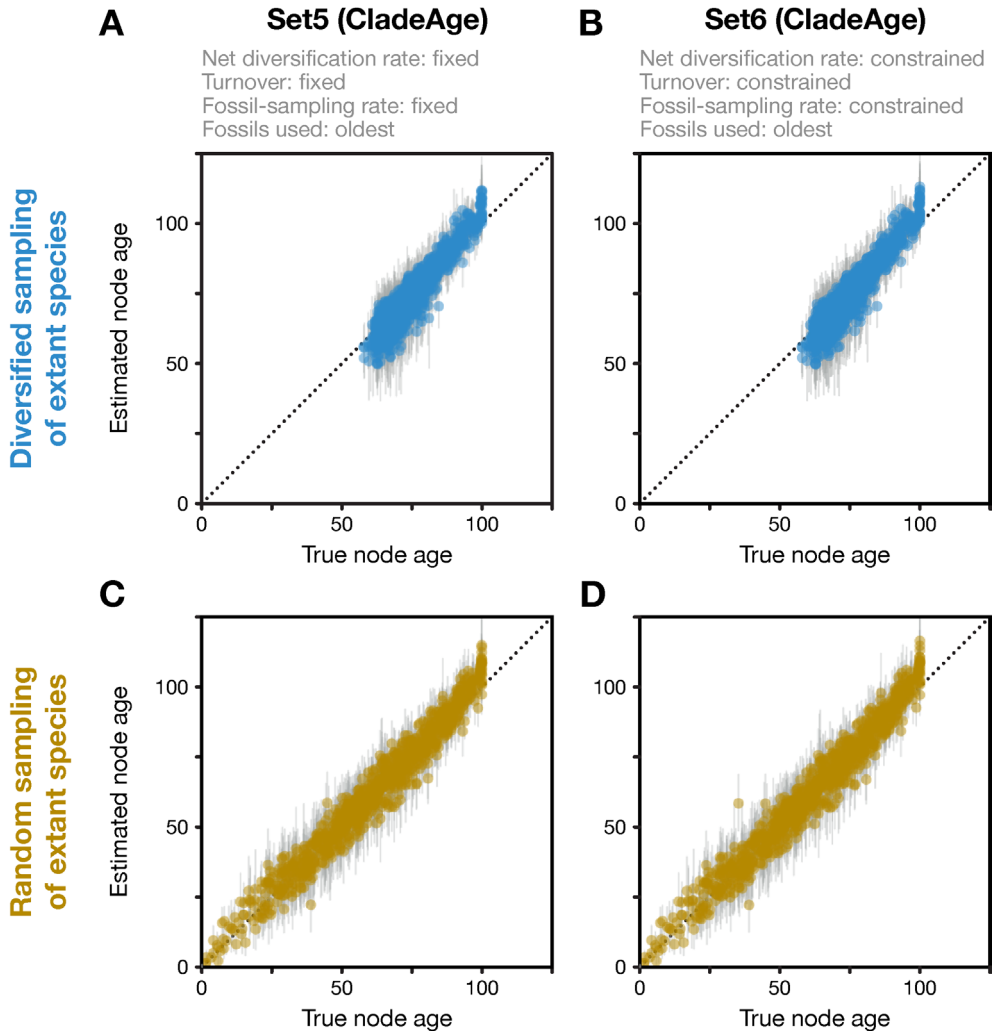


FIGURE 4 | Divergence times estimated with CladeAge. Comparison of true node ages and estimated node ages for the 20 diversified trees **(A**, **B)** and the 20 randomly sampled trees **(C**, **D)**. In the inference, diversification and sampling parameters were either specified exactly **(A**, **C)** or as confidence intervals **(B**, **D)**. The dotted lines mark the diagonals.

prohibitive run times of months or years. Whereas future improvements to FBD implementations may shorten these run times to some extent, it is questionable whether analyses with thousands of fossils and large molecular datasets will ever become computationally feasible (note, however, that by not using molecular data, the FBD implementation of Didier and Laurin (2018) allows rapid inference with larger numbers of fossils). On the other hand, random sampling of the fossil record may provide feasible run times and largely unbiased age estimates, but as the random sampling scheme may often exclude the oldest fossils of clades, the estimated ages of these clades may sometimes be younger than their oldest fossils if the molecular data do not permit sufficiently precise estimates. In contrast, as shown by my FBD analyses with settings Set2 and Set3, the sampling scheme in which only the oldest fossils per clade are used results in strongly overestimated node ages when the values of diversification and fossil-sampling parameters are not known exactly.

The results obtained with setting Set4 further suggest that even when all fossils are used in FBD analyses, diversified sampling of extant species leads to moderately inaccurate estimates of node ages, turnover, and fossil-sampling proportion. When the empirical sampling of extant species is in fact strictly according to the diversified sampling scheme, this issue could be solved with FBD model implementations that explicitly account for this scheme. An FBD model implementation with this feature is available in the program MrBayes (Zhang et al., 2016) but is so far missing from BEAST 2 (Bouckaert et al., 2019). However, even though many empirical datasets may be designed to include a diverse set of species, they are unlikely to follow the diversified sampling scheme strictly (Höhna, 2014). The reason for this is that the strict diversified sampling scheme expects

all nodes up to a certain age, but no nodes with younger ages, to be sampled, but this age information is not usually available to researchers prior to the analysis (Höhna et al., 2011). As a result, the use of FBD implementations that account for strict diversified sampling may result in node-age estimates that are biased in the opposite direction, towards underestimation, when empirical datasets are compiled with a sampling scheme that is intermediate between diversified and random sampling (Harrington and Reeder, 2017). A "semi-diversified" sampling scheme that could often be more appropriate for empirical datasets has been described and used for simulations by Colombo et al. (2015), but is not available for inference. In cases where each sampled species represents a clade with known species richness and no clades are missing from the phylogeny, the "empirical taxon sampling" scheme, which is available in RevBayes and accounts for varying fossil-sampling proportions across clades, might allow unbiased inference with RevBayes' FBD implementation. Testing this assumption with simulations should be the focus of a future study.

In contrast to the FBD model, CladeAge produced largely accurate node-age estimates after short run times, regardless of whether datasets had been generated with diversified or random sampling of extant taxa and regardless of whether diversification and sampling parameters were fixed or constrained within intervals. This difference between the models likely has two reasons: First, CladeAge explicitly assumes that only the oldest fossils per clade are used for calibration whereas the reduction of fossil records to the oldest fossils violates the FBD model. Second, whereas the FBD model assumes that the same diversification and sampling parameters apply to the fossil-generating process and the tree-generating process, this is not the case for CladeAge and thus, this method may be better able to buffer the model violation of diversified sampling by adjusting the diversification parameters of the tree-generating process without affecting the way in which fossils calibrate the tree. However, unlike the FBD model, CladeAge is unable to estimate the parameters of the fossil-generating process from the data, and these parameters therefore need to be specified by the user. While rough estimates of these parameters may be available from published literature (see, e.g., references in Supplementary Table 1 of Matschiner et al., 2017), separate analyses of clade-specific fossil records may in many cases be required to obtain these estimates, for example with the programs PyRate (Silvestro et al., 2014) TRiPS (Starrfelt and Liow, 2016), or Diversification (Didier et al., 2017). To obtain accurate divergence-time estimates, users should thus ensure that either their dataset conforms to the expectations of the FBD model—then this model will allow accurate estimation—or that estimates for diversification and sampling parameters are available *a priori*—then CladeAge can be used for the inference.

DATA AVAILABILITY STATEMENT

Code to reproduce this study can be found on GitHub: https://github.com/mmatschiner/fbd_test.

AUTHOR CONTRIBUTIONS

MM conceived and designed the study, performed simulations, analyzed results, prepared figures, and wrote the manuscript.

FUNDING

This research was funded by the Norwegian Research Council (FRIPRO 275869).

ACKNOWLEDGMENTS

I thank Michel Laurin, Gilles Didier, and Rachel Warnock for organizing the Timetrees symposium at the 5th International Palaeontological Congress in Paris and all participants of this symposium for stimulating discussions. I also thank Michel Laurin, Rachel Warnock, David Marjanović, David Buckley, and Gilles Didier for insightful comments that helped to improve the manuscript.

REFERENCES

Alfaro, M. E., Faircloth, B. C., Harrington, R. C., Sorenson, L., Friedman, M., Thacker, C. E., et al. (2018). Explosive diversification of marine fishes at the Cretaceous-Palaeogene boundary. *Nat. Ecol. Evol.* 2, 688–696. doi: 10.1038/s41559-018-0494-6

Ayres, D. L., Darling, A., Zwickl, D. J., Beerli, P., Holder, M. T., Lewis, P. O., et al. (2012). BEAGLE: an application programming interface and high-performance computing library for statistical phylogenetics. *Syst. Biol.* 61, 170–173. doi: 10.1093/sysbio/syr100

Barba-Montoya, J., dos Reis, M., Schneider, H., Donoghue, P. C. J., and Yang, Z. (2018). Constraining uncertainty in the timescale of angiosperm evolution and the veracity of a Cretaceous Terrestrial revolution. *New Phytol.* 38, 7. doi: 10.1111/nph.15011

Benton, M., and Donoghue, P. (2007). Paleontological evidence to date the tree of life. *Mol. Biol. Evol.* 24, 26–53. doi: 10.1093/molbev/msl150

Bouckaert, R. R., Alvarado-Mora, M. V., and Rebello Pinho, J. R. (2013). Evolutionary rates and HBV: issues of rate estimation with Bayesian molecular methods. *Antivir. Ther.* 18, 497–503. doi: 10.3851/IMP2656

Bouckaert, R. R., Vaughan, T. G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., et al. (2019). BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* 15, e1006650. doi: 10.1371/journal.pcbi.1006650

Colombo, M., Damerau, M., Hanel, R., Salzburger, W., and Matschiner, M. (2015). Diversity and disparity through time in the adaptive radiation of Antarctic notothenioid fishes. *J. Evol. Biol.* 28, 376–394. doi: 10.1111/jeb.12570

Didier, G., Fau, M., and Laurin, M. (2017). Likelihood of tree topologies with fossils and diversification rate estimation. *Syst. Biol.* 66, 964–987. doi: 10.1093/sysbio/syx045

Didier, G., and Laurin, M. (2018). Exact distribution of divergence times from fossil ages and tree topologies. *bioRχiv*. doi: 10.1101/490003

dos Reis, M., Thawornwattana, Y., Angelis, K., Telford, M. J., Donoghue, P. C. J., and Yang, Z. (2015). Uncertainty in the timing of origin of animals and the limits of precision in molecular timescales. *Curr. Biol.* 25, 2939–2950. doi: 10.1016/j.cub.2015.09.066

Drummond, A. J., Ho, S. Y. W., Philips, M. J., and Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4, e88. doi: 10.1371/journal.pbio.0040088

Gavryushkina, A., Heath, T. A., Ksepka, D. T., Stadler, T., Welch, D., and Drummond, A. J. (2017). Bayesian total-evidence dating reveals the recent crown radiation of penguins. *Syst. Biol.* 66, 57–73. doi: 10.1093/sysbio/syw060

Gavryushkina, A., Welch, D., Stadler, T., and Drummond, A. J. (2014). Bayesian inference of sampled ancestor trees for epidemiology and fossil calibration. *PLoS Comput. Biol.* 10, e1003919. doi: 10.1371/journal.pcbi.1003919

Gernhard, T. (2008). The conditioned reconstructed process. *J. Theor. Biol.* 253, 769–778. doi: 10.1016/j.jtbi.2008.04.005

Harrington, S. M., and Reeder, T. W. (2017). Phylogenetic inference and divergence dating of snakes using molecules, morphology and fossils: new insights into convergent evolution of feeding morphology and limb reduction. *Biol. J. Linn. Soc.* 121, 379–394. doi: 10.1093/biolinnean/blw039

Heath, T. A., Huelsenbeck, J. P., and Stadler, T. (2014). The fossilized birth-death process for coherent calibration of divergence-time estimates. *Proc. Natl. Acad. Sci. U. S. A.* 111, E2957–E2966. doi: 10.1073/pnas.1319091111

Höhna, S. (2014). Likelihood inference of non-constant diversification rates with incomplete taxon sampling. *PLoS One* 9, e84184. doi: 10.1371/journal.pone.0084184

Höhna, S., Landis, M. J., Heath, T. A., Boussau, B., Lartillot, N., Moore, B. R., et al. (2016). RevBayes: Bayesian phylogenetic inference using graphical models and an interactive model-specification language. *Syst. Biol.* 65, 726–736. doi: 10.1093/sysbio/syw021

Höhna, S., Stadler, T., Ronquist, F., and Britton, T. (2011). Inferring speciation and extinction rates under different sampling schemes. *Mol. Biol. Evol.* 28, 2577–2589. doi: 10.1093/molbev/msr095

Jarvis, E. D., Mirarab, S., Aberer, A. J., Li, B., Houde, P., Li, C., et al. (2014). Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science* 346, 1320–1331. doi: 10.1126/science.1253451

Kosiol, C., Holmes, I., and Goldman, N. (2007). An empirical codon model for protein sequence evolution. *Mol. Biol. Evol.* 24, 1464–1479. doi: 10.1093/molbev/msm064

Louca, S., and Pennell, M. W. (2019). Phylogenies of extant species are consistent with an infinite array of diversification histories. *bioRχiv*. doi: 10.1101/719435

Marjanović, D., and Laurin, M. (2007). Fossils, molecules, divergence times, and the origin of Lissamphibians. *Syst. Biol.* 56, 369–388. doi: 10.1080/10635150701397635

Matschiner, M., Musilová, Z., Barth, J. M. I., Starostová, Z., Salzburger, W., Steel, M., et al. (2017). Bayesian phylogenetic estimation of clade ages supports trans-Atlantic dispersal of cichlid fishes. *Syst. Biol.* 66, 3–22. doi: 10.1093/sysbio/syw076

Meredith, R. W., Janecka, J. E., Gatesy, J., Ryder, O. A., Fisher, C. A., Teeling, E. C., et al. (2011). Impacts of the Cretaceous terrestrial revolution and KPg extinction on mammal diversification. *Science* 334, 521–524. doi: 10.1126/science.1211028

Musilova, Z., Cortesi, F., Matschiner, M., Davies, W. I. L., Stieb, S. M., de Busserolles, F., et al. (2019). Vision using multiple distinct rod opsins in deep-sea fishes. *Science* 364, 588–592. doi: 10.1126/science.aav4632

Ogilvie, H. A., Vaughan, T. G., Matzke, N. J., Slater, G. J., Stadler, T., Welch, D., et al. (2018). Inferring species trees using integrative models of species evolution. *bioRχiv*. doi: 10.1101/242875

Plummer, M., Best, N., Cowles, K., and Vines, K. (2006). CODA: convergence diagnosis and output analysis for MCMC. *R News* 6, 7–11.

Pybus, O. G., and Harvey, P. H. (2000). Testing macro-evolutionary models using incomplete molecular phylogenies. *Proc. R. Soc. B* 267, 2267–2272. doi: 10.1098/rspb.2000.1278

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., and Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* 67, 901–904. doi: 10.1093/sysbio/syy032

Renner, S. S., Grimm, G. W., Kapli, P., and Denk, T. (2016). Species relationships and divergence times in beeches: new insights from the inclusion of 53 young and old fossils in a birth-death clock model. *Phil. Trans. R. Soc. B* 371, 20150135. doi: 10.1098/rstb.2015.0135

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., et al. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61, 539–542. doi: 10.1093/sysbio/sys029

Silvestro, D., Salamin, N., and Schnitzler, J. (2014). PyRate: a new program to estimate speciation and extinction rates from incomplete fossil data. *Method Ecol. Evol.* 5, 1126–1131. doi: 10.1111/2041-210X.12263

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

Stadler, T. (2010). Sampling-through-time in birth-death trees. *J. Theor. Biol.* 267, 396–404. doi: 10.1016/j.jtbi.2010.09.010

Stadler, T. (2011). Mammalian phylogeny reveals recent diversification rate shifts. *Proc. Natl. Acad. Sci. U. S. A.* 108, 6187–6192. doi: 10.1073/pnas.1016876108

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

Starrfelt, J., and Liow, L. H. (2016). How many dinosaur species were there? Fossil bias and true richness estimated using a Poisson sampling model. *Phil. Trans. R. Soc. B* 371, 20150219. doi: 10.1098/rstb.2015.0219

Warnock, R. C. M., Yang, Z., and Donoghue, P. C. J. (2012). Exploring uncertainty in the calibration of the molecular clock. *Biol. Lett.* 8, 156–159. doi: 10.1098/rsbl.2011.0710

Zhang, C., Stadler, T., Klopfstein, S., Heath, T. A., and Ronquist, F. (2016). Total-evidence dating under the fossilized birth-death process. *Syst. Biol.* 65, 228–249. doi: 10.1093/sysbio/syv080

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Matschiner. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genetics

REVIEW
published: 12 November 2019
doi: 10.3389/fgene.2019.01049

Using the Fossil Record to Evaluate Timetree Timescales

Charles R. Marshall[1,2]*

[1] Department of Integrative Biology, University of California, Berkeley, Berkeley, CA, United States, [2] University of California Museum of Paleontology, University of California, Berkeley, Berkeley, CA, United States

The fossil and geologic records provide the primary data used to established absolute timescales for timetrees. For the paleontological evaluation of proposed timetree timescales, and for node-based methods for constructing timetrees, the fossil record is used to bracket divergence times. Minimum brackets (minimum ages) can be established robustly using well-dated fossils that can be reliably assigned to lineages based on positive morphological evidence. Maximum brackets are much harder to establish, largely because it is difficult to establish definitive evidence that the absence of a taxon in the fossil record is real and not just due to the incompleteness of the fossil and rock records. Five primary methods have been developed to estimate maximum age brackets, each of which is discussed. The fact that the fossilization potential of a group typically decreases the closer one approaches its time of origin increases the challenge of estimating maximum age brackets. Additional complications arise: 1) because fossil data actually bracket the time of origin of the first relevant fossilizable morphology (apomorphy), not the divergence time itself; 2) due to the phylogenetic uncertainty in the placement of fossils; 3) because of idiosyncratic temporal and geographic gaps in the rock and fossil records; and 4) if the preservation potential of a group changed significantly during its history. In contrast, uncertainties in the absolute ages of fossils are typically relatively unimportant, even though the vast majority of fossil cannot be dated directly. These issues and relevant quantitative methods are reviewed, and their relative magnitudes assessed, which typically correlate with the age of the group, its geographic range, and species richness.

Keywords: timetree, calibration, phylogeny, cladogram, fossil record, absolute time

OPEN ACCESS

Edited by:
Michel Laurin, Centre de recherche sur la paléobiodiversité et les paléoenvironnements, France

Reviewed by:
Walter Joyce, Université de Fribourg, Switzerland
Juliana Sterli, National Council for Scientific and Technical Research (CONICET), Argentina
Philip Donoghue, University of Bristol, United Kingdom

***Correspondence:**
Charles R. Marshall
crmarshall@berkeley.edu

Specialty section:
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Genetics

Received: *20 June 2019*
Accepted: *30 September 2019*
Published: *12 November 2019*

Citation:
Marshall CR (2019) Using the Fossil Record to Evaluate Timetree Timescales. Front. Genet. 10:1049. doi: 10.3389/fgene.2019.01049

1 INTRODUCTION

Developing rigorous methods for using paleontological and geological data to estimate divergence times between lineages has proven challenging. Yet, these methods are needed for both the construction and evaluation of timetrees (Donoghue and Yang, 2016), trees where the relative branch lengths are largely derived from DNA sequence data but have been converted into units of absolute time. Timetrees consist of a topology, branch lengths proportional to time, and an absolute timescale. Here, I am specifically interested in the paleontological evaluation of the timescales, the estimates of lineage divergence times—that is, I focus on how paleontologists estimate divergence times, not on how a given timetree might have been generated. Nonetheless, some of my discussion has bearing on the construction of timetrees, especially those derived from node-dating methods where the fossil record is used to provide priors on divergence times, including the difficult-to-establish maximum age constraints (Yang and Rannala, 2006; Ho and Phillips, 2009). Some of my discussion is also relevant to non-node-dating

methods for constructing timetrees (see Donoghue and Yang, 2016 for a review), even though these do not require *a priori* maximum estimates of divergence times, for they still need to make assumptions about the rates of fossil recovery (Warnock et al., 2017). These methods include the Fossilized Birth Death (FBD) process (Heath et al., 2014; Stadler et al., 2018), total evidence methods that simultaneously estimate the phylogenetic position of the extant taxa and relevant fossils (Pyron, 2011; Ronquist et al., 2012), and integration of the FBD and total evidence methods (Zhang et al., 2016; Gavryushkina et al., 2017).

1.1 The Three Components of the Paleontological Estimation of Divergence Times

The first component is the simplest, establishing the minimum estimate of the divergence time. This consists of identifying the oldest fossil of the focal lineage, its First Appearance Datum (*FAD*) (**Figure 1A**). As paleontologists are typically limited to working with morphological data, the minimum age constraint corresponds to the age of the oldest appearance in the fossil record of the first fossilizable apomorphy of the focal lineage.

Given the incompleteness of the fossil record, a literal reading of the fossil record is biased in that the age of the *FAD* will post-date the divergence time—we need to estimate the size of this temporal gap, that is, provide a maximum age constraint. However, because paleontologists must deal with morphological data, the statistical methods paleontologists have developed for estimating maximum age constraints actually pertain to the estimation of the true time of origin of the first fossilizable apomorphy (ΔT_{Gap} in **Figure 1A**) not the actual divergence time itself. Thus, estimating maximum age constraints consist of two steps. The first step, and second component of the paleontological estimation of divergence times, consists of estimating the size of the temporal gap between the *FAD* and the true time of origin of the first fossilizable apomorphy (ΔT_{Gap} in **Figure 1A**). The second step, and third component of the paleontological estimation of divergence times, consists of estimating the size of the gap between the true time of origin of this first apomorphy and the actual divergence time between the focal lineage and its extant sister clade ($\Delta T_{Div\text{-}1stApo}$ in **Figure 1A**). This last factor is often ignored, although it has long been recognized (e.g., see Marshall 1990b; Magallon, 2004; Steiper and Young, 2008; Marshall and Valentine, 2010). It is the hardest to quantify because there will typically be a lag between the time of genetic separation of two lineages, their divergence time, and the time of origin of the first fossilizable diagnosable morphological feature, the first autapomorphy, in the focal lineage. As discussed below, one or both of ΔT_{Gap} and $\Delta T_{Div\text{-}1stApo}$ can be large depending on the taxon.

1.2 Coalescence Times

Turning for a moment to the DNA component of timetrees, note that DNA data, when properly calibrated, provide a measure of the divergence time ($T_{Divergence}$) plus the coalescence time for the loci being compared (e.g., see **Figure 1** in Edwards and Beerli, 2000) (**Figure 1A**):

$$T_{DNA} = T_{Divergence} + T_{Coalescence} \quad (1)$$

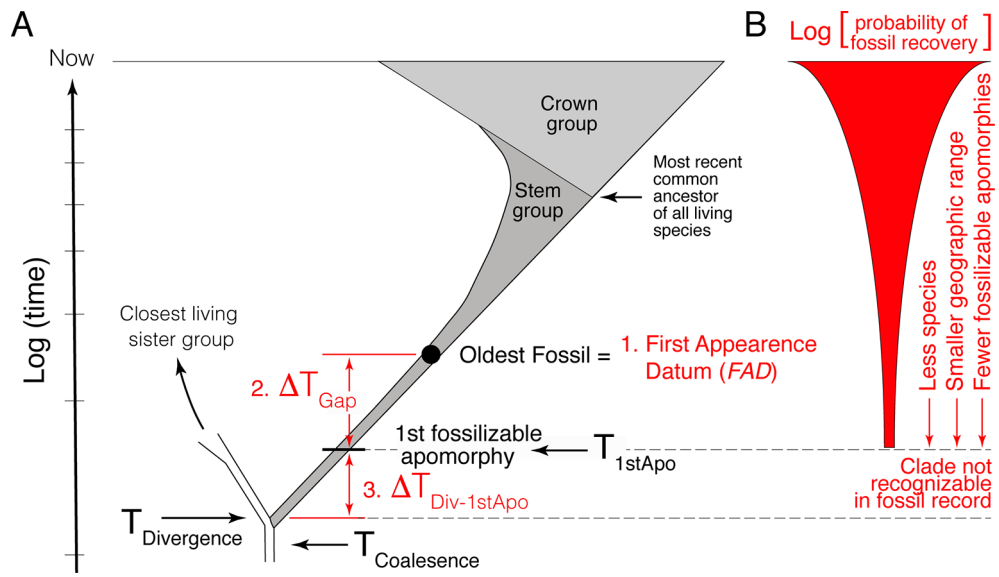


FIGURE 1 | Schematic representation of the challenges encountered when using the fossil record to estimate divergence time between lineages. **(A)** The three primary challenges are: 1) the easiest, identifying and determining the age of the oldest fossil (*FAD*); 2) estimating the size of the temporal gap between the *FAD* and the time of origin of the first fossilizable apomorphy of the lineage (ΔT_{Gap}); and 3) estimating the size of the gap between the true time of origin of the first fossilizable apomorphy and the actual divergence time ($\Delta T_{Div\text{-}1stApo}$). $\Delta T_{Div\text{-}1stApo}$ cannot be directly addressed with the fossil record because fossils that belong to this part of the clade's history are not recognizable. **(B)** Estimating the size of ΔT_{Gap} is made difficult by the fact that the probability of recovering fossils for most lineages decreases as one approaches its time of origin, as well as the fact that the fossil and rock records are idiosyncratically incomplete temporally and spatially (not depicted).

Thus, even with accurate and precise temporal calibration with metronomically evolving DNA sequences, estimated divergence times will be too deep if one fails to take into account the standing polymorphism that was present at the time of the population divergence of the lineages of interest unless a correction has been made (e.g., Marshall and Swift, 1992). This issue is most important for shallower divergence times, typically less than a few million years, where the magnitude of the coalescence time can be a significant proportion of the divergence time (Edwards and Beerli, 2000). Note that this issue may be compounded by the fact that different loci may yield different topologies, which in turn may lead to incorrect branch lengths, which can impact inferred divergence times. But even if all loci yield the same topology, equation (1) still holds—DNA data bear directly on the coalescent time between the loci analyzed, not the divergence time *per se*.

1.3 The Challenge of Dealing With the Temporally Biased Fossil Record

Temporal information in the fossil record is biased, with correctly identified well-dated *FAD*s being younger than their respective divergence times. Quantifying how much older divergence times are than *FAD*s is challenging because there is no positive evidence that a taxon existed a given temporal distance beyond its know temporal (stratigraphic) range; it is hard to establish whether the absence of the taxon is real or just due to the incompleteness of the fossil record. Statistical approaches can be used, but the rigor of these approaches is made difficult by the fact that the probability of finding fossils of a clade generally decreases beyond its *FAD* (**Figure 1B**) given that: 1) at the time of inception of a clade there is only one lineage; 2) they likely lived in a limited geographic area; and; 3) typically, there are fewer and fewer diagnosable morphological features with, which to recognize fossils of the focal clade as one approaches its time of inception, a factor exacerbated by the fact that fossils are often fragmentary.

In the discussion that follows, I concentrate on quantitative methods where they exist. Note that the best practices depend in part on: 1) the richness of the fossil record within the focal group; 2) the richness of the fossil records of clades that lie outside the focal group with similar fossilization potentials; 3) the phylogenetic scope of the study; and, (4) the depth in geologic time over which the focal group evolved, which is typically correlated with the phylogenetic scope of the study.

1.3.1 Heterogeneity in the Incompleteness of the Fossil Record

The stochastic nature of the fossil record means that the gap size between *FAD*s and true divergence times will be heterogeneous in size, which becomes relevant when generating timetrees with methods that use uncorrelated rates of molecular evolution (see section 1.4. below), and when contemplating the use of cross-validation approaches (see section 3.6 below). This heterogeneity has long been recognized (Jaanusson, 1976; Marshall, 1995), and its importance for the temporal calibration of molecular phylogenies was highlighted by Springer (1995). Springer (1995) showed using the Australian marsupial fossil record that a literal reading of the fossil record led to an estimate of the average rate of singly copy DNA evolution of 1% per million years, with a 17-fold difference from one lineage to the next (this was before DNA-branch lengths were used as part of timetree estimation). However, once Springer (1995) took into account the incompleteness of the fossil record using a confidence interval approach (see section 3.1 below), the data were shown to be consistent with a constant rate of DNA evolution at a much slower rate of 0.4% per million years.

1.4 Timetree Construction Is Especially Sensitive to Paleontological Data

It is well recognized that timetree timescales are very sensitive to the paleontological data used for calibration [e.g., see Barba-montoya et al. (2018) for a succinct summary]. Part of the reason is that when constructing timetrees there is typically no further explicit information on absolute time beyond the paleontological data used; thus, in Bayesian analysis for example, there is no direct data within the analysis to update the priors on the divergence times, and thus, those priors tend to constrain the range of dates in the resulting timetree. This dependence on the paleontological data means that timetree construction with uncorrelated rates of molecular evolution with priors that favor a literal reading of the fossil record (i.e., exponential priors; see Section 3.1.2 below) will tend to collapse the nodes onto the ages of the *FAD*s.

The sensitivity to the paleontological data itself stems from: 1) the difficulty in establishing rigorous maximum age constraints on divergence times [relevant to node-dating approaches (Yang and Rannala, 2006; Ho and Phillips, 2009)]; 2) the uncertainty in the phylogenetic placement of fossils either due to missing data or conflicting characters (e.g., see Sterli et al., 2013) (which effects almost all approaches); and, 3) uncertainties in the actual dating of fossils [which can have a large effect on total evidence approaches (O'Reilly et al., 2015)].

If this review has any simple take home message, it is that it is crucial that the utmost care be taken in specifying divergence time priors (Warnock et al., 2015).

2 ESTIMATING ROBUST MINIMUM DIVERGENCE TIMES

The best practices for establishing minimum age estimates from fossil data, the oldest fossil securely assignable to the focal lineage, are well established (e.g., Benton and Donoghue, 2007; Donoghue and Benton, 2007; and especially Parham et al., 2012). Below, I outline the key points. I do not consider the use of paleobiogeographic constraints, except to note that they often lack precision both because the emergence of a land-bridge or the opening of a seaway is often a protracted event, and because most organisms have a dispersal capacity which means that divergence times can predate the formation of biogeographic barriers, often by an unknown magnitude. See Ho et al. (2015) and De Baets et al. (2016) for synoptic summaries of issues associated

with using biogeographic constraints on divergence times, Loeza-Quintana and Adamowicz (2018) for an iterative approach for dealing with the complexity typical of biogeographic constraints, and Landis (2017) for a method for integrating information from multiple biogeographic constraints.

2.1 Minimum Times of Origin (*FAD*s) Must Be Apomorphy Based

This principle has now been well articulated (e.g., see Benton and Donoghue, 2007; Donoghue and Benton, 2007; Parham et al. 2012; Sauquet et al., 2012). Parham et al. (2012) also emphasize the importance of explicit listing of relevant museum numbers for the specimens that show the chosen apomorphies, as well as reconciling any discordance between molecular and morphological phylogenies that might impact which node the calibration fossil calibrates. It is important to take into account uncertainties in the phylogenetic position of calibration fossils, as these can greatly impact timetree calibration (Sterli et al., 2013). Careful selection of apomorphy-rich calibration fossils helps ameliorate the impact of this factor. I will not discuss here the interesting approaches designed to co-estimate the phylogeny of all the taxa, both the fossil and living (Pyron, 2011; Ronquist et al., 2012), where the phylogenetic placement of fossils is part of the process of generating the timetree, except to note that the richer and more accurate the morphological description of the fossils, the less ambiguity there will be in where those fossils join the phylogeny.

The reason *FAD*s must be apomorphy based is easily demonstrated. Imagine two closely related living taxa *X* and *Y*, where *Y* has morphological autapomorphies with respect to *X*, and where there is an oldest fossil that belongs unequivocally to morphospecies *X*, thus constituting the *FAD* of X (**Figure 2A**). For simplicity, I assume there are no fossils assignable to morphospecies *Y*. The age of *FAD X* represents a robust minimum estimate on the divergence time between the species *if* taxon *X* (including its fossils) has at least one morphological autapomorphy with respect to *Y*, as it implies that *Y* diverged from *X* before the first appearance of that autapomorphy (otherwise, we would expect *Y* to also have that character state) (**Figure 2A**).

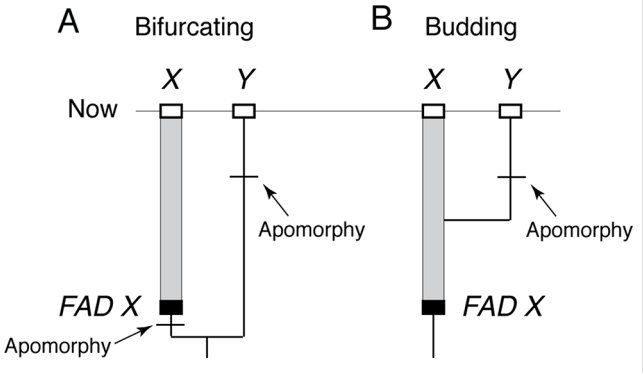


FIGURE 2 | The importance of apomorphy-based *FAD*s when establishing minimum age constraints. **(A)** When lineage *X* has an apomorphy with respect to lineage *Y*, then oldest fossil of *X*, *FAD X*, can be used as minimum age constraint on the divergence of *X* and *Y*. **(B)** However, if *X* lacks apomorphies with respect to *Y*, then the age of *FAD X* does not constrain the divergence time of *X* and *Y* because *Y* could have budded off *X* at any time. Note: without a fossil record of *Y*, it is not possible to place a minimum age constraint on when its apomorphy evolved.

However, if morphospecies *X* (including its fossils) has no morphological autapomorphies with respect to taxon *Y*, then *Y* could have budded off from lineage *X* at any time (**Figure 2B**) with the possibility that *FAD X* predates the emergence of lineage *Y*. In this case, the age of *FAD X* is not an unequivocal minimum estimate on the divergence time between the two lineages as it could either post-date of pre-date their time of divergence. Note, further, that character state reversals are commonly observed in morphological data, so there will be some probability that even if *X* has autapomorphies with respect to *Y*, which *Y* might still have budded off lineage *X*, having subsequently lost those characters (Wagner, 1998).

While the notion of budding has been part of paleontological reasoning for decades (e.g., see Raup, 1985) and underpins the FBD method of incorporating fossils (Heath et al., 2014), only recently have its implications for integrating neontological and paleontological data begun to be explored (Silvestro et al., 2018; Stadler et al., 2018). Below, I give two examples where the fact that the morphology of "hosts" of diverging DNA sequences might be subject to stasis can affect the way one interprets, and in the second case, calibrates DNA trees.

2.1.1 The Dentist Who Infected Several Patients With HIV

In a case that gained international notoriety, a DNA tree derived from a portion of the HIV genome verified that an HIV-positive dentist in Florida had accidently infected several of his patients (Ou et al., 1992). The DNA tree itself shows the dentist at the top of the tree (**Figure 3A**), which might suggest that the dentist acquired HIV recently from patient C, although without additional data, there is no way of knowing. However, once one recognizes that budding is possible, in the sense that the dentist remained unchanged as the host to the virus that was evolving within him, then a budding tree can be drawn (**Figure 3B**) where it is immediately obvious that the dentist sequentially infected several patients.

2.1.2 Potential Example of Budding Cladogenesis—Multiple Invasions of Riverine Gastropods Into Lake Tanganyika?

The endemic thalassoid gastropods in Lake Tanganyika represent one of the many species flocks in the major East African lakes. Surprisingly, a widely distributed riverine and putative outgroup, *Cleopatra*, lies high in DNA trees of the group, buried deeply within the endemic Lake Tanganyikan clade (**Figure 3C**). A natural explanation for this topology is that *Cleopatra* is not an outgroup but had its origin in the lake to later invade the adjacent rivers (West and Michel, 2000; Wilson et al., 2004). Under this scenario, the oldest fossil morphologically assignable to *Cleopatra*, i.e., its *FAD*, might be used as a minimum age constraint on *Cleopatra*'s divergence from its closest relatives, the lake endemics *Stormsia* and *Spekia* (**Figure 3C**) or *Reymondia* (West and Michel, 2000).

However, another possibility, consistent with the DNA tree, is one of pervasive morphological budding cladogenesis

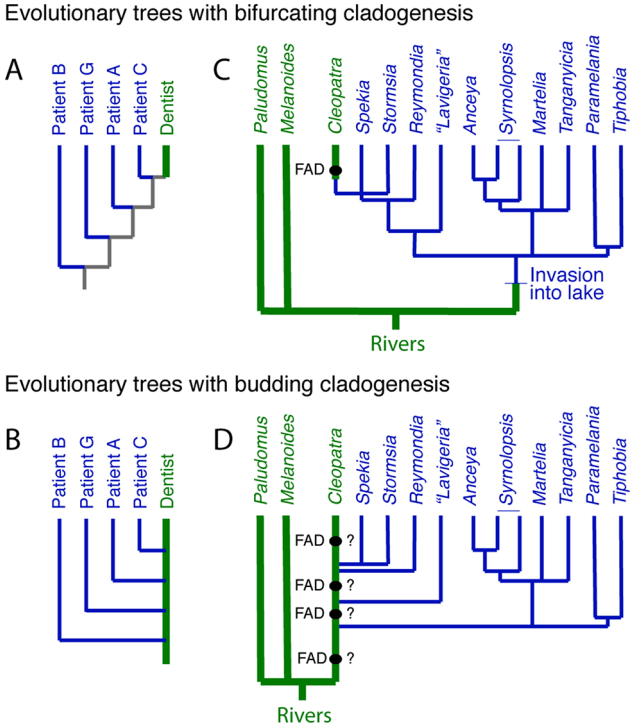


FIGURE 3 | The distinction between bifurcating and budding cladogenesis matters for DNA trees. **(A)** Part of the evolutionary tree (*y* clone of the HIV-1 *env* V3 region) of the dentist and some of his HIV infected patients from Ou et al. (1992). The gray internal branches indicate that, without additional information, we do not know if the virus in the patients was derived from the dentist or not. **(B)** The tree from (A) redrawn with budding cladogenesis, making obvious the sequential infection of patients with HIV by the dentist. **(C)** Bifurcating tree at the genus level of three East African riverine and multiple endemic Lake Tanganyika gastropods, which suggests the riverine *Cleopatra* evolved from a lake endemic (Wilson et al., 2004)—the *FAD* of *Cleopatra* constrains the divergence of *Cleopatra* from its nearest relatives. **(D)** However, if the lineage that lead to the living *Cleopatra* invaded the lake several times (West and Michel, 2000), then the *FAD* of *Cleopatra* does not offer a reliable constraint on the divergences within the lake taxa, unless it can be shown to have synapomorphies with a subset of the lake endemics.

(**Figure 3D**) (West and Michel, 2000). Under this scenario, the endemic lake fauna were derived from multiple invasions into the lake of populations of riverine snails that might have been morphologically indistinguishable from the living *Cleopatra* (or some allied forms [see Van Damme and Pickford, 2003]). In the bifurcating DNA tree, this would have left *Cleopatra* as sister group to the last lake lineage it gave rise to, as is observed. If this scenario is correct, then the oldest fossil *Cleopatra* will lack morphologic autapomorphes with respect to its lake descendent lineages, and thus, these fossils offer no minimum age constraint on the time of origin of the lake lineages.

This scenario has yet to be formally tested but highlights the fact that morphospecies-level interpretations of DNA-based topologies could be inaccurate if one ignores the possibility of morphological budding. It also highlights the importance of apomorphy-based minimum age constraints. Intriguingly, the budding scenario finds support in the fact that fossil *Cleopatra* are known to at least 12.5 million years ago (Van Damme and Pickford, 2003), older than the onset of rifting that led the formation of the lake ~9–12 million years ago (Cohen et al., 1993).

2.2 Most Groups Have Problematic Potential FADs

Given that the number of diagnostic features drops as one approaches the origin of a group, and given that most fossils are morphologically incomplete, most groups have problematic fossils that might conceivably be *FAD*s, but where there is insufficient morphology preserved to be sure. If one is simply trying to establish reliable brackets on divergence times, then the best practice is to only use morphologically secure FADs, which are typically younger than older potential *FAD*s (Donoghue and Benton, 2007). This approach also ameliorates to some degree the sensitivity of temporal calibrations to the phylogenetic uncertainty in the placement of key fossils.

2.3 Dating *FAD*s

2.3.1 The Basis for the Dating of *FAD*s Needs to Be Explicit

Parham et al. (2012) deal with the need for explicit statements about how the absolute age constraints on an *FAD* have been established, including the locality and stratigraphic level the specimen(s) came from, and the basis of the absolute time assigned to that stratigraphic level. Here, for those not familiar with how ages are assigned to fossils, is the reason for their insistence.

2.3.2 Only the Youngest Fossils can be Dated Directly

There are two standard ways of directly dating fossil material *via* radioisotopes. The first and more versatile is ^{14}C dating, but its half life is so short (5,730 years) that reliable dates can only be obtained for fossils up to about 40,000–60,000 years old (Taylor and Bar-Yosef, 2014). The second is uranium series disequilibrium dating of carbonates (which biologically includes corals) including $^{238}U/^{234}U/^{230}Th$ and $^{235}U/^{231}Pa$ datings (Edwards et al., 2003). But it also can only be applied to very young fossils, just over 600,000 years (for ^{230}Th dating, see Stirling et al. (2001)).

2.3.3 The Dating of the Vast Majority of *FAD*s Is Indirect

Ultimately, all absolute dates in the rock record are derived from radiometric dates. These typically provide an estimate of when the minerals that contain a relevant radioisotope crystallized out of molten rock, either in a magma chamber (most commonly zircons, which trap ^{235}U and ^{238}U) or as a volcanic ash is erupted from a volcano (most commonly sanidine feldspar, which traps ^{40}Ar). In an ideal case, a key fossil will lie in sediments that are bracketed by younger and older dateable volcanic ash layers. Even better is the rare case where a fossil is actually embedded in a datable rock—for example, the rhinocerotid skull found in a 9.2 million year old ignimbrite flow erupted from a volcano in Turkey (Antoine et al., 2012). The worst case scenario is where the fossil of interest lies in sediments with no nearby igneous rocks,

and without fossils that can be temporally correlated with similar fossils elsewhere (i.e., using biostratigraphy) where constraining radiometric dates are available. Thus—for example, this is the case for the famous Ediacaran localities of enigmatic latest pre-Cambrian fossils in the Ediacara Hills in South Australia that have yet to be dated with any sort of precision.

The more normal situation is where local radiometric dates (or other well dated events, such as switches in the Earth's magnetic polarity) are not available, but where biostratigraphy can be used to correlate with places that have some age control. The dating of the famous Cambrian Burgess Shale fauna is one of these—its age assignment is based on biostratigraphy on the assumption that its trilobite fauna, specifically *Ehmaniella*, lived at about the same time at other localities where radiometric dates are available (see p.441 in Peng et al., 2012). However, while the order in which species appear and disappear in the fossil record is pretty consistent in different geographic areas, species typically take time to reach their maximum geographic range and are often extirpated (become locally extinct) heterogeneously on their way to extinction (Foote, 2007; Foote et al., 2007; Liow and Stenseth, 2007)—thus, paleontologists assume that the presence of a species in two geographic areas only indicates approximately the same point in geologic time (see Figure 7.9 in Taylor, 1987, reproduced in Donoghue and Benton, 2007).

In absolute terms, spatial asynchrony in times of first and last appearances of a species is typically less than an average species duration [~2 million years for Cenozoic mammals, for example (Marshall, 2017)], perhaps no larger than a few hundred thousand years, but conservatively ± 1 million years. Radiometric dating errors are typically less than 1% of the age of the rock (Cohen et al. 2013) but can be as low as 0.1% for Ar-40/Ar-39 dating (Sprain et al., 2019) and down to almost 0.01% for U-Pb dating (Burgess et al., 2014), which is probably smaller than the uncertainty in pre-eruptive residence time of zircons in magma chambers (where zircons form before they are erupted). However, in some cases, the age uncertainties can be large—for example, the Dominican amber is very poorly constrained with an age range from 15 to 20 million years ago (Iturralde-Vinent and MacPhee, 1996; Ramírez et al., 2007).

Finally, there is variety of other indirect methods available for dating fossils. For example, the ratio of $^{87}Sr/^{86}Sr$ can be used to estimate the age of deposition of sediments, or hard tissues such as the shells of fossil brachiopods, as long as one knows roughly how old the fossil is [in some cases, this prior knowledge can be very imprecise—for example, for younger fossils, it is often sufficient to simply know that it is Cenozoic in age to make use of the approach (McArthur et al., 2012)]. The precision can be as high as ± 0.1 million years (McArthur et al., 2012).

2.3.4 Dating Uncertainties Are Typically Relatively Small

Generally speaking, if good dates are available for *FAD*s (see Benton and Donoghue (2007); Clarke et al. (2011); and especially Benton et al. (2015) for many examples across the tree of life), the dating errors are small compared with the approximate divergence time, in the order of a million years, typically shorter than a species duration. However, sometimes, the dating of key *FAD*s is imprecise—for example, the age of the oldest fossil hominin, *Sahelanthropus*, lacks precision, somewhere between 6.5 and 7.5 million years old [see discussion in Benton and Donoghue (2007) and Reis et al. (2018)], an uncertainty that amounts to ~14% the total age of the fossil.

3 MAXIMUM AGE CONSTRAINTS—STEP 1: ESTIMATING THE SIZE OF ΔT_{GAP}

While establishing robust minimum age constraints is relatively straightforward (using well-dated, well-diagnosed apomorphy-rich fossils), there are no well established procedures for establishing robust maximum age constraints, a challenge that has plagued node-dating approaches and has led some to favor alternative approaches (for discussion, see Donoghue and Yang (2016)). In this section, I deal with constraining the time of origin of the oldest fossilizable apomorphy, which is estimating the size of ΔT_{Gap} (**Figure 1A**). In the following section (Section 4), I then deal with estimating the size of the gap between the true time of origin of the oldest fossilizable apomorphy and the actual divergence time ($\Delta T_{Div\text{-}1stApo}$). Note that total evidence approaches (Pyron, 2011; Ronquist et al., 2012), by ducking the need to estimate maximum age constraints, simply ignore the fact that fossil age estimates of divergence times are too young.

Below, I discuss five approaches to estimating the size of ΔT_{Gap}: 1) *confidence interval* approaches, which use quantitative measures of the richness of the fossil record of individual lineages within the focal taxon to estimate how much of the fossil record might be missing; 2) the *taphonomic control group* approach which uses the ages of non-focal-taxon fossils that are older than the focal taxon's *FAD* to provide evidence that the absence of the focal taxon is real; 3) the *super-taxon* approach which uses an un-calibrated timetree to combine all the *FAD*s across into a "super-taxon," which is then analyzed using a confidence interval approach; 4) *clade diversity dynamics* approaches that model the stratigraphic ranges of species not preserved in the fossil record; and 5) *FADs of successive outgroups* approach, which can be used when the fossil record is quite rich. The first two methods can be applied to multiple lineages with the focal clade, or to entire clades. The latter three were designed for the analysis of entire clades (**Table 1**).

Note that all these methods place soft maxima on the target time of origin (Marshall, 1990a, Marshall, 1990b; Yang and Rannala, 2006; Benton and Donoghue,2007; Donoghue and Benton, 2007), that is, they provide a maximum age constraint at some level of confidence or probability, typically, 95%.

It has also been suggested that ancestral fossils can also be used as hard maxima (e.g., see Marshall, 1990b). But even if a taxon has a morphology consistent with being ancestral to some target species, and is older than the target species as it should be (Smith, 1994, chapter 6), unless the proportion of all species preserved is very high (see Section 4.1.1 for some examples), the chances are that the putative ancestral fossil will be sister to the focal taxon not ancestral (Foote, 1996) and thus is unlikely to represent a valid maximum constraint. The reason is simply that, for any given taxon, the proportion of ancestral lineages is small compared with the number of older non-ancestral lineages (Foote, 1996).

TABLE 1 | Methods that use fossil data to place soft maxima on times of origin of first fossilizable apomorphies, that is, to estimate the size of ΔT_{Gap} (**Figure 1A**). See text for citations.

Method	Can be applied to:		Ancillary data needed
	Multiple lineages in focal clade	Whole clade	
1. Confidence intervals	Yes	Yes	Fossil record richness within focal clade
2. Taphonomic control groups	Yes	Yes	Fossil record outside of focal clade
3. Super-taxon confidence intervals	No	Yes	Un-calibrated ultrametric tree; multiple *FAD*s
4. Clade diversity dynamics	No	Yes	Extant and fossil species richness
5. *FAD*s of successive outgroups	No	Yes	*FAD*s of successive outgroups

3.1 Method 1—Confidence Intervals to Constrain ΔT_{Gap}

This approach uses the density of fossil finds through time within the focal taxon to quantify ΔT_{Gap}.

3.1.1 Simplest Approach

Intuitively, the richer the fossil record the closer the *FAD* will be to the true time of origin of the first fossilizable apomorphy of the group (T_{1stApo}) (Strauss and Sadler, 1989; Marshall, 1990a; Marshall, 1990b). For an extant taxon known from n distinct fossil localities, the sizes of the temporal gaps between successive localities will be exponentially distributed if fossilization and fossil recovery were random, and the confidence interval T_C on T_{1stApo} at a confidence level, C, is given by

$$T_C = FAD(1-C)^{-1/n} \tag{2}$$

The unbiased estimate of the T_{1stApo} is the average gap size, FAD/n. In the paleontological literature, equation (2) has an exponent of $-1/(n-1)$, not $-1/n$ (see equation [8]), reflecting the fact that the equation was developed for extinct species where one needs to condition on the last fossil occurrence, which also results in the gaps sizes between successive localities having a Dirichlet rather than exponential distribution (Strauss and Sadler, 1989).

3.1.2 Likelihood Formulation of the Simplest Approach

Bayesian approaches for generating timetrees typically require specification of priors on the times of origin, which need to be expressed as likelihoods. Thus, the frequentist formulation described above needs to be translated into a likelihood framework. This has been done by Strauss and Sadler (1989)—the fact that stratigraphic gap sizes will be distributed exponentially under random fossilization (Strauss and Sadler, 1989) suggests that the most appropriate prior will be exponentially distributed (**Figure 4A**). However, maximum likelihood analysis of monotonic distributions such as the exponential are biased, with the maximum likelihood estimate corresponding to a zero temporal range extension (Strauss and Sadler, 1989)—the maximum likelihood estimate for the time of origin of a clade (T_{ML}) is the age of its oldest known fossil (**Figure 4A**):

$$T_{ML} = FAD \tag{3}$$

However, a correction for the maximum likelihood estimate for the time of origin, T^*_{ML}, given a finite sample size (n), can be found by multiplying T_{ML} by $(n+1)/n$:

$$\begin{aligned} T^*_{ML} &= FAD(n+1)/n \\ &= FAD + FAD/n \end{aligned} \tag{4}$$

$$= FAD + \textit{average temporal gap between fossil localities} \tag{5}$$

Thus, as for the frequentist analysis above, the most likely time of origin is the average gap size below the *FAD*.

3.1.2.1 The Paleontologically Most Appropriate Prior

This analysis suggests that the paleontologically most appropriate prior on the time of origin of a clade, at least that can be detected with the fossil record, will have a mode that extends an average gap size below the oldest fossil, and, using equation (2), a 95% tail extending $\sim FAD(0.05)^{-1/n}$ beyond the age of the oldest fossil locality (the *FAD*) (**Figure 4B**). Of the priors currently available for the construction of timetrees, the lognormal distribution has this shape (Ho and Phillips, 2009), as do the gamma (Yang and Rannala, 2006) and truncated Cauchy (Inoue et al. 2010) distributions depending on how they are parameterized. Given a *FAD* and the number of distinct localities (n) for a taxon, the mean and variance of the corresponding lognormal prior can be found (see Appendix A for derivation):

$$\mu = \ln\left(FAD\right) - \ln\left(n\right) + \sigma^2 \tag{6}$$

$$\sigma^2 = \left(-0.8224 + 0.5\sqrt{2.7055 - 4\left(-2.9957/n - \ln\left[n\right]\right)}\right)^2 \tag{7}$$

3.1.2.2 Comparison With an FBD Process Timetree

Interestingly, Heath et al. (2014) analysis of extant bear divergence times, using their "fossilized birth–death" (FBD) process for constructing timetrees, provides uncertainty estimates on their divergence times that are broadly congruent with priors developed with the procedure outlined above (see **Figure 5**). This is perhaps not surprising given that the same

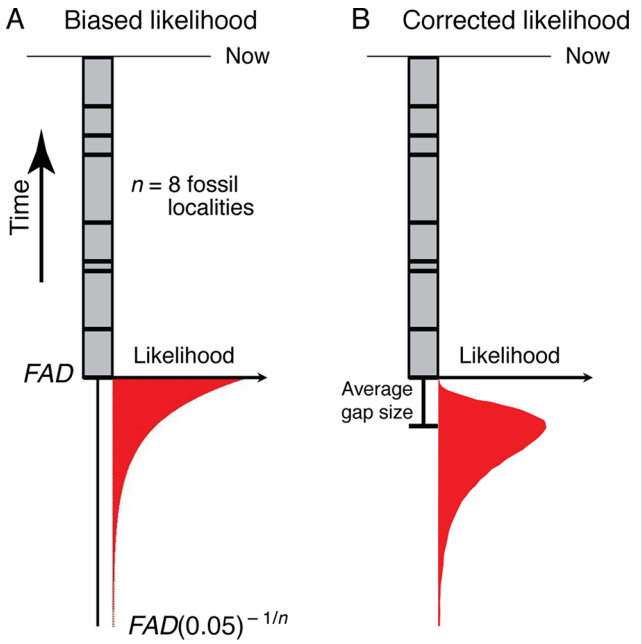


FIGURE 4 | (A) The exponential likelihood density function for how much older the true time of origin of a morphospecies is than its *FAD* (ΔT_{Gap}) under the assumption of random fossilization. **(B)** The likelihood density function in (A) corrected for a finite sample size (see Section 3.1.2).

fossil data were used in both the computation of the confidence intervals and in the FBD analysis, which takes into account the average preservation rate based on the fossils incorporated into the analysis. Nonetheless, it is heartening that there is broad agreement between a purely paleontological approach (the paleontological parameterization of the lognormal distribution) and an approach that incorporates fossil and DNA data, as well as an explicit branching model as part of its inference engine (the FBD process). As a methodological aside, note that modifications of the FBD model can accommodate variation in diversification and fossil recovery rates [e.g., see Gavryushkina et al. (2014)].

3.1.3 Difficulties With the Simplest Approach

Generally speaking, we do not expect the probability of finding fossils to be stochastically constant through a lineage's temporal range; instead, we expect the probability of finding fossils to decrease the closer we approach the time of origin (**Figure 1B**) [see Marjanović and Laurin (2008) for an empirical example]: 1) the number of separate lineages will approach one, the initiating lineage; 2) the geographic range is likely to be smaller; and, 3) there will be progressively less apomorphies of the group, making it progressively harder to diagnose incomplete fossils and thus unequivocally assign taxa to the clade of interest. Moreover, (4) the rock and fossil records are spotty both temporally and geographically [e.g., see Wagner and Marcot (2013) and Marjanović and Laurin (2008) for some empirical

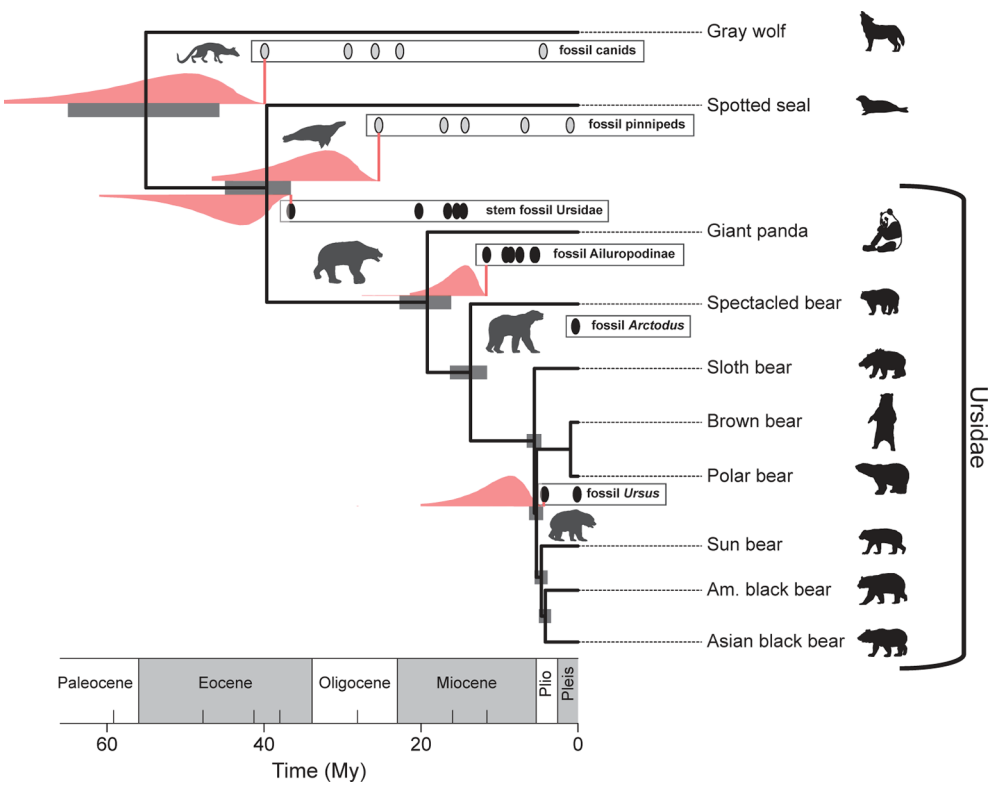


FIGURE 5 | Broad agreement in the estimated uncertainties in the divergence times of selected bear and outgroup lineages based on the FBD (fossilized birth–death) process (gray bars) (Heath et al., 2014) and those based on lognormal-like likelihood analysis of the fossil records depicted (pink distributions; see also Figure 4). Pink vertical lines correspond to the *FAD*s. The pink range extensions have been added to a reproduction of Figure 4 from Heath et al. (2014).

examples], which adds further uncertainty, although we now know a great deal about the controls and therefore the structure of the sedimentary rock record (Patzkowsky and Holland, 2012; Holland, 2016). Below is a simple example that illustrates the impact of the incompleteness of the rock record has in computing confidence intervals, and then I discuss ways in which the decrease in preservation potential can be accommodated.

3.1.3.1 Example—Trying to Date the Time of Origin of a Sand Dollar

As part of my Ph.D., I attempted to quantify the uncertainty in the divergence time between several sand dollar species. Multiple confounding difficulties made this difficult, which are exemplified here by my analysis of the time of origin of the genus *Mellita*.

The fossil record of *Mellita* is relatively strong with at least 10 localities alone from the Atlantic Coastal Plain of the USA (Lindberg, 1983) (**Figure 6**). The *FAD* in this region is ~ 4 million years old, and the 95% confidence interval extends to about ~5 million years old. However, while the rock record always appears complete in outcrop, it is typically riddled with temporal gaps (Sadler, 1981; Holland, 1995; Patzkowsky and Holland, 2012; Holland, 2016). Further, even when rocks are present in a given time interval, they might not represent suitable environment for the taxa of interest. For example, *Mellita* only lives on sandy substrates, and thus, its fossils are only found in sandstones—the St. Mary's Formation (formation "f" in **Figure 6**) is a muddy unit, and so, we don't expect fossil *Mellita* to be found there. When confidence intervals are calculated by rock thickness of the sandstone formations, a way of taking into account the major temporal gaps in appropriate deposition, the 95% confidence interval extends into the Choptank Formation (formation "g" in **Figure 6**), about ~10 million years in age, doubling the soft maximum estimate.

Moreover, *Mellita* and its sister genus *Leodia* have current geographic ranges that extend to Uruguay (Mooi and Peterson, 2000; Martínez and Mooi, 2005), and the now-extinct basal members of the clade to which *Mellita* and *Leodia* belong are only known from Uruguay, Argentina, and Chile (Mooi et al. 2000). But there are no fossil *Mellita* or *Leodia* known south of Caribbean, and so, it is quite possible that these genera had their origins in geographic region from which there has been very little paleontological effort exerted, the Atlantic coast of South America; the fossil record might be giving us a record of when *Mellita* and *Leodia* migrated into Caribbean (and then into the Pacific), not when they originated. Using the fossil record of these now-extinct basal members of the clade as taphonomic controls (see Section 3.2), a soft maximum limit on the time of origin of the genus *Mellita*, is the Middle to Upper Miocene boundary (Mooi et al., 2000), 11.6 million years ago.

Relatively few analyses of this kind have been undertaken [although see Marjanović and Laurin (2008)], but it exemplifies how the temporal and geographic spottiness of the rock and fossils record, and our spotty knowledge of them, adds considerable uncertainty to when a taxon originated.

3.1.3.2 Dealing With a Non-Random Fossil Record

In **Figure 1**, I have emphasized the fact that the probability of fossil recovery generally drops the further back we go in the history of a clade. The importance of accommodating this is illustrated *via* a thought experiment—if one assumes that crown group birds existed in the Cretaceous, but at, say, 1/10th Cenozoic preservation rates, the 95% confidence on the crown group time of origin of a relatively fossiliferous group, the Caprimulgiformes, increases to ~90 million years from the ~70 million year estimate under the assumption of random fossilization (Marshall, 1999).

Marshall (1997) developed a method for accommodating decreasing probabilities of fossil recovery with time (or in fact any non-random distribution of fossil recovery potential). However, we do not yet have standard methods for developing the required empirical non-random fossilization potential curves (Marshall, 2010). Nonetheless, Marjanović and Laurin (2008) provide an

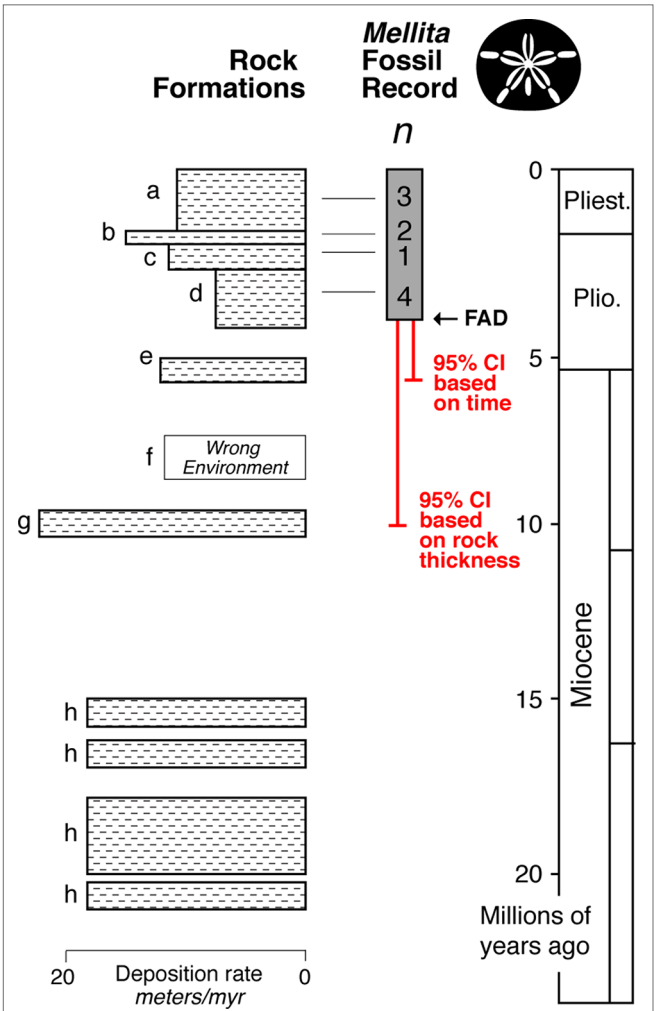


FIGURE 6 | Idealized representation of the rock formations on the Atlantic Coastal Plain region of the USA (Lindberg, 1983), with the fossil record and minimum number of localities (*n*) for each formation for the fossil sand dollar *Mellita* (silhouette). The *FAD* in this region is ~4 million years old, with the end of the 95% confidence interval (equation 1) extending to ~5 million years ago. However, when one takes into account the substantial gaps in the rock record, and the fact that some rock formations were deposited in environments inimical to *Mellita* (formation f), the 95% confidence interval extends to ~10 million years.

example of a compound model that used sedimentary rock outcrop area through time coupled with an exponential model of diversification punctuated by mass extinctions to model the diversity trajectory of the living amphibians (the Lissamphibia). They parameterized (and tested the efficacy of) their model with the Lissamphibian fossil record (1,207 localities across the global history of the group) to establish confidence intervals on its time of the origin using the generalized confidence interval approach of Marshall (1997). Marjanović and Laurin's (2008) study may serve a good model for realistic fossil-based confidence intervals for higher taxa, and thus also for establishing priors on divergence times.

There are also simpler analytic methods for accommodating trends of decreasing fossil recovery within the known stratigraphic range to approximate the assumed further decrease in fossil recovery beyond the known stratigraphic range. The first methods used the Weibull distribution, which assumes a decreasing rate of preservation (Roberts and Solow, 2003; Solow and Roberts, 2003). However, these methods tend to overestimate the true temporal endpoint (Rivadeneira et al., 2009). The most recent and best performing method is the flexible beta method of Wang et al. (2016), but all of these methods assume simple monotonic change in fossil recovery potential, unlike the empirically richer (and demanding) approach of Marjanović and Laurin (2008).

3.1.3.3 Another Way of Ameliorating the Difficulties—the Origin of Hominins

Another way of trying to ameliorate the difficulties associated with the decreasing probably of encountering fossils the further back we go in time (**Figure 1B**) is to work only with the oldest part of a lineage's fossil record where the fossil recovery rate is likely to have been relatively constant (however that is determined). Thus—for example, from the divergence of our own species from chimpanzees to about 4 million years ago, our own evolutionary branch may well have consisted of just one lineage. Fossils come from just $n = 4$ fossiliferous places in Africa for this early part of our history, one for *Sahelanthropus* [which provides our lineage's *FAD* of between 7.5 and 6.5 million years ago; see Benton and Donoghue (2007); Reis et al. (2018) for discussion of the age uncertainty], one for *Orrorin* [which has yielded fossils from ~6.0 to 5.7 million years ago (Sawada et al. 2006)] and two for *Ardipithecus* (~5.8–5.2 (Haile-Selassie 2001) and ~4.4 million years ago (White et al. 2009)]. This yields a stratigraphic range (R) of 2.1–3.1 million years. For paleontological data, the confidence interval T_C on the *FAD* at a confidence level, C, is given by (Marshall, 1990a):

$$T_C = R[(1-C)^{(-1/n-1)} - 1] + FAD \quad (8)$$

The 95% confidence interval on this part of the hominin fossil record using equation (8) extends to between 12.8 and 10.1 million years ago, with an unbiased estimate (the average gap size added to the *FAD*) of between 8.5 and 7.2 million years ago. This is in good agreement with the taphonomic control group approach (see Section 3.2 immediately below), where there are several fossil localities that yield hominin fossils at about 10 million years ago with no evidence of fossils assignable to either the chimpanzee or hominin lineages (Benton and Donoghue, 2007; Reis et al., 2018).

3.2 Method 2—Constraining ΔT_{Gap} With Taphonomic Control Groups

The difficulty in quantifying the fossil recovery potential of a taxon beyond its known stratigraphic range has led many to rely on a more qualitative approach, the age of taphonomic control groups found beyond the *FAD* of the focal taxon as a maximum estimate for the time of origin. Taphonomy is the study of how organisms decay and become fossilized (Behrensmeyer et al., 2000), and taphonomic control groups are groups that are frequently found preserved in the same rocks, or at least the same environments, as the focal taxon—thus, their preservation in rocks older than the *FAD* of the focal taxon is taken as an indication that the focal taxon had not yet evolved (Bottjer and Jablonski, 1988; Marshall 1990b; Benton and Donoghue, 2007; Donoghue and Benton, 2007). Of course, the co-occurrence of taxa is typically never 100%, and so the first appearance of the taphonomic control group, *per se*, may not be a fully robust maximum bound on the time of origin of the focal taxon, a supposition that finds some empirical support (Clarke et al. 2011). To control for geographic incompleteness (see section 3.1.3.1 above), the control group must be found in the same broad geographic region where we think the focal group originated.

3.2.1 Example—The Origin of the Sand Dollar *Clypeaster*

The clypeasteroid echinoids, the sand dollars, and sea biscuits (e.g., *Clypeaster*) have a relatively rich fossil record. Ali (1983) documents 397 species in the fossil record known from 768 localities (so on average, each species is known from about two localities). With this quality of fossil record, we have reasonable confidence that the genus had its origin in the equatorial Tethys Sea (**Table 2**), now seen in the rock record around the Mediterranean and in the Middle East. The oldest fossils are in Middle Eocene. Other irregular echinoids are found in the region in the Lower Eocene and in the older Paleocene [see Souto et al. (2019) and also the Paleobiology Database (PBDB), although for most groups the PBDB only documents a portion of the known fossil record (Marshall et al., 2018)]. Thus, a reasonable maximum estimate for the time of origin of *Clypeaster* was by beginning of the Eocene, and we can be even more certain that it had its time of origin somewhere in the interval bracketed by its Middle Eocene *FAD* and the beginning of the Paleocene.

3.3 Method 3—The Super-Taxon Confidence Interval Approach

Confidence intervals on stratigraphic ranges have long been used to assess likely times of extinctions, especially mass extinctions (Marshall, 1995a; Marshall and Ward, 1996; Jin et al., 2000; Marshall, 2010; Wang and Marshall, 2016). The most powerful approach for mass extinction victims is to combine all the data, effectively collapsing all the species into a single super-taxon

TABLE 2 | Species occurrences of the 397 fossil species in Ali's (1983) compilation of the genus *Clypeaster* by time and geographic region.

Duration (myr)	Epoch	E. Pacific	Caribbean	Tethys Ocean (Mediterranean)	Indian Ocean	West Pacific
2.6	Pleistocene	3	6		10	4
2.8	Pliocene	10	14	16	20	10
17.7	Miocene	7	44	508	31	3
10.9	Oligocene	1	35	23	10	1
7.3	Upper Eocene		2	6	2	
6.6	Middle Eocene			2		
8.2	Lower Eocene					

The intensity of the shading is proportional to the log(# occurrences/myr). The rich fossil record of Clypeaster *indicates that the genus originated in the region of what is now the Mediterranean Sea, which was the Tethys Ocean at the time of* Clypeaster*'s FAD.*

(Wang et al., 2009). The same logic has been applied in reverse, using all the *FAD*s to form a super-taxon, to estimate the time of origin of a clade, as well as the internal branches, where the relative positions of the *FAD*s are adjusted by the relative length of their branches on an un-calibrated ultrametric tree (timetree) (Marshall, 2008)(**Figure 7**). This approach has the advantage of not requiring estimates of maximum divergence times but has the disadvantage of the potentially unrealistic assumptions about the fossilization process (although see Marshall (2008) for discussion). It differs from most approaches for constructing time trees in that it is sequential in nature—an ultrametric tree is constructed first in the absence of any absolute time constraints, and then the scaling of that tree is established using the super-taxon paleontological approach.

3.3.1 Congruence Between the Taphonomic Control Group and Super Taxon Methods?

Several analyses of turtle divergence times (Joyce et al., 2013; Pereira et al., 2017; Shaffer et al., 2017) have employed the best practices for establishing *FAD*s and used taphonomic control groups for establishing soft maxima (Joyce et al., 2013) (**Table 3**). All three studies used the same fossil calibrations, updated from (Near et al., 2005), except for Pereira et al. (2017) who used a updated minimum paleontological date for the root node. Marshall (2008) also used the Near et al. (2005) data and the super-taxon approach to estimate turtle divergence times. When the super-taxon approach is adjusted by eliminating the three *FAD*s identified as being questionable by Joyce et al. (2013) [*"Aspideretes" maortuensi* (calibration lineage 6), *Proterochersis robusta* (calibration lineage 1), and *Santanachelys gaffneyi* (calibration lineage 5), which Marshall's method also indicated as being problematic], leaving lineage 10 (*Baltemys*) as the calibration lineage, the new super-taxon results are broadly congruent with the taphonomic control group studies (**Table 3**).

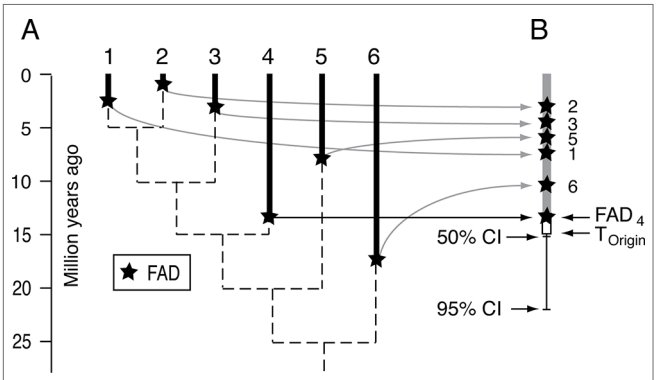


FIGURE 7 | Schematic for the super-taxon approach for using multiple *FAD*s to constrain the time of origin of a clade. **(A)** Hypothetical ultrametric tree (dashed lines) with the *FAD*s for each lineage. **(B)** The method uses the branch lengths of the ultrametric tree to map the *FAD*s onto a single lineage, or super-taxon, and then uses confidence intervals (equation [1]) to bracket the time of origin, that is estimate the size of ΔT_{Gap} for the entire clade. See Marshall (2008) for further explanation. Adapted from **Figure 1** in Marshall (2008).

3.4 Method 4—Using *FAD*s of Successively More Inclusive Clades

As one examines successively older rocks focal taxa disappear with only successively more plesiomorphic sister groups being found (from the point of view of the focal group). Thus, using a taphonomic control group type reasoning, the presence of these plesiomorphic taxa without taxa from the focal group gives the sense that the focal group had not yet evolved, providing a maximum age estimate for the focal taxon. Following this logic—for example, Gustafsson et al. (2010) used the *FAD* of the entire monocot clade of plants as a maximum age constraint on the time of origin of the orchids within the monocots.

This approach has been formalized for cases where the order in which a series of clades appear on a cladogram is matched by the temporal order of those clades' *FAD*s (Hedman, 2010). The method is developed in a Bayesian framework and is implemented in R (Lloyd et al., 2016). It can been adjusted for groups that violate this requirement by leaving out inconsistent groups (Friedman and Brazeau, 2011; Friedman et al., 2013) and has the virtue that it does not require any estimate of preservation and fossil recovery rates.

The method has been applied to the origin of digit bearing tetrapods, with a 95% credible interval from~396 to 427 million years ago (Friedman and Brazeau, 2011), although this example highlights the potential discrepancies between times of origin of fossilizable apomorphies (digits in this case) and lineage divergence times—the relatively rich fossil record of the first tetrapods and their precursors indicates that the time of origin of digit bearing tetrapods considerably post-dates the time of divergence of tetrapods from the nearest living relatives, the lungfish or coelacanths (Marshall and Schultze, 1992; Min and Schultze, 2001).

TABLE 3 | Taphonomic control group and super-taxon confidence interval approaches to estimating maximum (and minimum) age constraints on turtle divergence times are broadly congruent when they all use the same fossil *FAD*s.

Study	Method for assigning maxima	Turtle crown group age (million years ago)	
		Mean	95% HPD/ confidence
Joyce et al. (2013)	Taphonomic control group	212	195–231
Pereira et al. (2017)	Taphonomic control group[1]	199.5	179–225
Shaffer et al. (2017)	Taphonomic control group[1]	220	194–251
Marshall (2008)[2]: H-bar = 1	Super-taxon CI	220	209–259
Marshall (2008)[2]: H-bar = 2	Super-taxon CI	214	209–232

[1]*Following Joyce et al. (2013).*
[2]*With problematic FADs eliminated (see text).*

3.5 Method 5—Modeling the Stratigraphic Ranges of Missing Species

Typically, only a small proportion of all species that have ever existed are found in the fossil record. For example, less than 7% of all living primate species are found as fossils (Martin, 1993; Tavaré et al., 2002; Wilkinson et al., 2011). Martin (1993) noted that when the proportion of species preserved is small, and especially for clades that have been steadily expanding, the oldest fossil species found in the group might be several species durations younger than the very first species. For example, he simulated the expansion of clade over 16 species durations from one to 48 extant species, which resulted in a total of ~380 species, ~330 of which were extinct. A random sample of 3% of the extinct species yielded on average an oldest fossil that was five species durations younger than the base of the clade—his fossil record was missing about the first 30% of the true stratigraphic range of the group.

This approach was formalized by Tavaré et al. (2002) and applied to the crown group primate fossil record where the oldest fossil is ~55 million years old (that is, we ignored the extinct stem-group primates, the plesiadapiformes). This paleontological method yielded a mean estimate of 81.5 million years for crown group primate origins, a date compatible with the molecular clock estimates of the divergence of primates from their nearest relatives [which in 2002 was about ~90 million years ago (Tavaré et al., 2002)]. More recently, Wilkinson et al. (2011) developed an extension of Tavaré et al.'s (2002) approach for integrating paleontological and molecular data, obtaining a very similar result, with a mean estimate of 84.5 million years ago for crown group primates.

Foote et al. (1999) employed a different analytic approach but also used a branching process as well as explicit preservation rates to determine how deep into the Cretaceous several mammalian orders likely extended. Foote et al. (1999) found that the fossil preservation rates for the better preserved mammalian orders give much younger times of divergence, much closer to the end of the Cretaceous, which is dated to 66 million years ago, in conflict with the older crown group primate date.

The reason for the discrepancy has not been determined, but while it is not unreasonable that a relatively poorly preserved group of mammals, crown group primates, for example, might have a very deep time of origin; it is harder to believe that all the other better preserved mammalian orders, which all diverged from each other at about the same time as primates, also had a similarly deep time of origin. Note, however, that Wilkinson et al. (2011) 95% confidence interval on the time of origin of crown primates ranges from 69.2 to 103.5 million years ago, its upper limit compatible with Foote et al.'s (1999) analysis.

Primates continue to be a test case for combined DNA-paleontological timetree construction. For example, Reis et al. (2018) find that crown group primates originated toward the younger end of the range established by Wilkinson et al. (2011), somewhere between 79.2 and 70.0 million years ago for their preferred analysis (using autocorrelated rates of molecular evolution), or perhaps 71.4–63.9 million years ago (with uncorrelated rates). They also find that primary sources of uncertainty in the analysis are associated with fossil calibration uncertainty.

Tavaré et al. (2002) approach assumed logistic diversification, conditioned on the number of living species, the oldest fossil crown group primates, and the richness of the crown group primate fossil record. Newer paleontological approaches have been developed to estimate the amount of time missing history prior to *FAD*s that explicitly use the fossil record to calculate average speciation, extinction, and preservation rates from the fossil record (Bapst, 2013; Nowak et al., 2013). Most recently, Wagner (2019) has developed a method for estimating branch durations and stratigraphic gaps in phylogenies when rates of speciation, extinction, and fossil sampling vary with time.

3.6 Multiple Calibration Points and Cross Validation

The first two methods (confidence intervals and taphonomic control groups) make use of multiple calibrations across the tree (**Table 1**). This has the advantage that the process of calibration is not so dependent on difficulties that might be associated with any one specific lineage (e.g., see the discussion of the sand dollar *Mellita* for an example of these difficulties, section 3.1.3.1).

The use of multiple calibrations also allows for the possibility of cross validation, that is, the search for consistency between the temporal calibrations of one calibration to the next.

3.6.1 Cross Validation in Light of the Bias in the Fossil Record

The initial idea of cross validation was to see if various subsets of calibration points (*FAD*s) yielded similar absolute ages for the timetree, with the goal of eliminating calibrations that yielded anomalously young or old divergence time estimates (Near and Sanderson, 2004; Near et al., 2005). However, given that *FAD*s all underestimate the divergence times they are being used to estimate, simple cross-validation on *FAD*s will provide divergence time estimates that are too young, eliminating the best calibrations (Marshall, 2008), as well as the worst. To overcome this shortfall, rather than cross-validating on the *FAD*s, cross-validation may be performed using the temporal ranges between the minima (*FAD*s) and soft maxima (Clarke et al., 2011), however those soft maxima are established. However, cross-validation is not generally

recommended because the sequential use of single calibrations does not have the same effect as the simultaneous use of all the calibrations (Warnock et al., 2015).

4 HOW MUCH OLDER ARE DIVERGENCE TIMES THAN THE TRUE TIMES OF ORIGIN OF THE FIRST FOSSILIZABLE APOMORPHIES—HOW BIG IS $\Delta T_{Div\text{-}1stApo}$?

So far, we have ignored the second step in the estimation of maximum age constraints on divergence times, the fact that the fossil record calibration methods discussed above give estimates of the true time of origin of the first diagnosable fossilizable morphological feature of a lineage, not the divergence time from its sister group (**Figure 1A**). The reason is that the fossil record can only be used to constrain the time of origin of taxa where those taxa can be morphologically recognized as belonging to that lineage. Thus, paleontological methods only provide data on the morphologically diagnosable portion of a lineage's history.

Often $\Delta T_{Div\text{-}1stApo}$ is probably relatively small, but it can be very large indeed. Below, I examine situations that span this range (see **Table 4** for a synopsis). I begin with the expectation with a complete fossil record.

4.1 Size of $\Delta T_{Div\text{-}1stApo}$ if All Extinct Species Were Preserved

Generally speaking, the appearance of a new fossilizable autapomorphy results in the recognition of a new paleontological species. Thus, if all extinct species have been preserved, and if there was no drop in fossilization potential prior to the emergence of the first apomorphy, the lag between the actual divergence time and the first appearance of the first diagnosable autapomorphy ($\Delta T_{Div\text{-}1stApo}$) will be in the order of an average paleontological species' duration, because that is (very roughly) about how long it takes to evolve the first fossilizable diagnosable morphology:

$$\Delta T_{Div-1stApo} \leq \text{average paleontological species duration} \quad (9)$$

However, if at the inception of the lineage there was rapid morphological change, then $\Delta T_{Div\text{-}1stApo}$ would be much shorter than a species duration; so, equation (9) should be viewed as an upper limit.

There are relatively few extant groups for which average species durations have been calculated, but for Cenozoic North American mammals, the average species duration is ~2.3 million years (based on an analysis of >3,000 fossil species), while for planktic foraminifera, it is 5–10 million years, depending on their morphology; for angiosperms it is ~3 million years, Coniferales at just over 5 million years, pteridophytes at ~12 million years, and cycads at ~15 million years [see compilation in Marshall, 2017].

Turning to the size of ΔT_{Gap}, even if all species were preserved, there would still be a gap between the *FAD* of the founding species and the true time of origin of the first fossilizable apomorphy, but in all probability ΔT_{Gap} would be relatively small, less than an average species duration. Thus, to a first approximation, $\Delta T_{Div\text{-}1stApo}$ and ΔT_{Gap} would probably be of a similar magnitude if all extinct species were preserved.

4.1.1 Near-Perfect Fossil Records do Exist

A simple metric for measuring the quality of the fossil record is given by:

$$Q = \text{proportion of extant taxa found in the fossil record} \quad (10)$$

Probably the richest fossil record is the marine skeletonized single-celled eukaryotic microplantkon. In particular, the Cenozoic macroperforate planktonic foraminifera are so abundant in the fossil record that not only does $Q \approx 1$, but it is estimated that each species has at least an 81% of being sampled per million year interval (Ezard et al., 2011). In some geographic regions, marine macroinvertebrates are also well represented in the fossil record. For example, 77% of the 698 living species of bivalves and gastropods living at shelf depths in the California Province are found in the Pleistocene of the same region, with perhaps 85% ($Q = 0.85$) of all durable species captured in the fossil record (Valentine, 1989).

There is no comprehensive database of Q values, but within mammals, which overall have a strong fossil record, cetaceans are considered to have a good fossil record, with $Q = 0.54$ to 0.59 at the genus level (Quental and Marshall, 2010), while the primate fossil record is relatively weak, with $Q < 0.07$ at the species level (Tavaré et al., 2002). Didier et al. (2017) using an explicit diversification model estimate that about 14% ($Q = 0.14$) of all lineages of Permo-Carboniferous stem group mammal (synapsids) are currently known from the fossil record.

From a calibration standpoint, even if only 1% of extant taxa are preserved ($Q = 0.01$), there will still be many calibration points across a reasonably large phylogeny, given that species-turnover rates are sufficiently high that, on geologic timescales, the number

TABLE 4 | Relative magnitudes of $\Delta T_{Div\text{-}1stApo}$ (the difference between the paleontological estimate of the true time of origin of the first fossilizable apomorphy and the paleontologically unobservable lineage divergence time) and ΔT_{Gap} [the size of the gap between the first appearance in the fossil record (*FAD*) and the true time of origin of that first fossilizable apomorphy] (see **Figure 1A**).

Nature of the fossil record	$\Delta T_{Div\text{-}1stApo}$	Importance of $\Delta T_{Div\text{-}1stApo}$ compared with ΔT_{Gap}
Preservation potential ~constant		
Perfect—all species preserved	≤ a species duration	$\Delta T_{Div\text{-}1stApo} \approx \Delta T_{Gap}$ (section 4.1)
Good—many species preserved	≤ a species duration	$\Delta T_{Div\text{-}1stApo} < \Delta T_{Gap}$ (section 4.2)
Poor—very few species preserved	≤ a species duration	$\Delta T_{Div\text{-}1stApo} << \Delta T_{Gap}$ (section 4.3)
Long, poorly preserved stem groups	Can be many species durations	ΔT_{Gap} (section 4.4)
Preservation decreases near base of clade	Can be many species durations	$\Delta T_{Div\text{-}1stApo}$ can swamp ΔT_{Gap} (section 4.5)

of fossil species greatly exceeds the number of extant species. For example, for Cenozoic mammals, for each of the ~5,500 living species, it has been estimated that there were ~ 26 times that number that are now extinct (Marshal, 2017). Thus, for Q = 0.01, the expected number of preserved mammal species would be ~5,500 x 0.01 x 26, or some 1,400 fossil taxa. If only one of the ~5,500 extant mammal species was known in the fossil record (Q = 1/5,500 = 0.0002), we would still expect some 26 calibration fossils. Even if Q = 0.0, there may be fossils available for calibration (see orchid example, where Q = 0, Section 5.1 below).

4.2 Size of $\Delta T_{Div\text{-}1stApo}$ for Realistic (but Good) Fossil Records

As discussed in section 3.5 above, many species durations can be missing from the base of a taxon's observed fossil record. Thus, for groups that are considered to have pretty good fossil records, the expected gap between the *FAD* and the time of appearance of the first fossilizable apomorphy (ΔT_{Gap}) is likely to be many species durations, swamping the size of $\Delta T_{Div\text{-}1stApo}$, which is probably less than a species duration (equation [9]). For example, as discussed above, the oldest fossil crown group primate is about 55 million years old, but it is possible that the actual time of origin is ~85 million years ago (for primate species, Q < 0.07). This difference of ~30 million years is large compared with the ~2.3 million years of an average mammalian species duration. Even if primates diverged from their nearest relatives at the upper limit of Wilkinson et al. (2011) range, at ~69 million years ago, that is still ΔT_{Gap} of 14 million years, six average species-durations, much larger than <2 million years guesstimate for $\Delta T_{Div\text{-}1stApo}$.

Generally speaking, $\Delta T_{Div\text{-}1stApo}$ is likely to be small compared to ΔT_{Gap} for groups with good fossil records (except for most recently diverged clades, those that diverged on the order of a species duration ago): using the first fossilizable apomorphy as a proxy for the desired divergence time, i.e., ignoring $\Delta T_{Div\text{-}1stApo}$, will not typically add substantial error.

4.3 Size of $\Delta T_{Div\text{-}1stApo}$ for Weaker Fossil Records

For groups with poor fossil record, where very few species are preserved, the size of ΔT_{Gap} is even larger than with better fossil records, further decreasing the importance of $\Delta T_{Div\text{-}1stApo}$ over groups with better fossil records. Nonetheless, sometimes $\Delta T_{Div\text{-}1stApo}$ can be very large, as discussed below.

4.4 Size of $\Delta T_{Div\text{-}1stApo}$ With Long-Lived Poorly Preserved Stem Groups

4.4.1 Neontological Data Have Significant Blind Spots

The pervasiveness of extinction has left large lacunae in the record of cladogenic events that can be accessed *via* the living biota. Those lacunae, unbroken branches on molecular phylogenies, can be very long and typically represent stem groups (diagrammed in **Figure 1A**). For example, the last common ancestor of all living birds, the base of the crown group, dates to the late Cretaceous, perhaps 66–87 million years ago (Benton et al., 2015), while the divergence between birds and their living sister group (the crocodiles) dates to deep in the Triassic or into the late Permian (247–260 million years ago) (Benton et al., 2015). Thus, there are no living lineages that connect to the bird lineage over the first ~70% of its history, since it diverged from its nearest living relatives, the Crocodilia. Similarly, for angiosperms, where the fossil record is more difficult to work with (Coiro et al., 2019), the uncontested *FAD* for crown group angiosperms is ~126 million years old (Coiro et al., 2019) with a maximum estimate of 256 million years ago (Barba-montoya et al., 2018), while their divergence from their living sister group is anywhere from 306 to 367 million years ago (Clarke et al., 2011)—perhaps 40% of the history of the angiosperm lineage is not accessible *via* living species.

4.4.2 For Groups With Long Poorly Preserved Stem Groups $\Delta T_{Div\text{-}1stApo}$ Can Be Very Long

With relatively poor preservation potential, and with long stem groups, it can be very difficult in the fossil record to determine the size of the gap between the time it diverged from its living sister group and when the first diagnosable apomorphy of the group originated ($\Delta T_{Div\text{-}1stApo}$). For example, the primary paleontological diagnostic feature of angiosperms, tricolpate pollen, is first seen in the fossil record 125 ± 1 million years ago (Clarke et al., 2011; Coiro et al., 2019), while the *FAD* of its living sister group is known from rocks 307 ± 1 million years ago (Clarke et al., 2011). We don't when the first species that we would recognize as an angiosperm by neontological criteria first appeared with in this interval, but $\Delta T_{Div\text{-}1stApo}$ could be well over a 100 million years.

4.4.3 Mammalian Radiation and the End-Cretaceous Discontinuity

While fossil mammals were abundant before the end-Cretaceous mass extinction (Luo, 2007), it appears that there was an increase in the importance of mammals in terrestrial ecosystems after the mass extinction, accompanied by the relatively rapid evolution of new morphologies (Alroy, 1999). Thus, the ability to recognize members of the living mammalian orders may have been reduced in the Cretaceous if they were present—the ecological discontinuity across the end-Cretaceous mass extinction adds to the size of $\Delta T_{Div\text{-}1stApo}$ for the living mammalian orders, but we do not know how to quantify its magnitude.

4.5 Size of $\Delta T_{Div\text{-}1stApo}$ With a Radical Change in Preservability at the Base of the Clade

The size of $\Delta T_{Div\text{-}1stApo}$ can be very large if one or more of the earliest diagnostic features of the group dramatically increased the preservability of the lineage. For example, it appears that the last common ancestor of the well-skeletonized animal phyla was un-skeletonized—the first representatives of the animal phyla were probably not readily diagnosable in the

fossil record (e.g., see Marshall, 2006; Marshall and Valentine, 2010). Thus, it is difficult to use the fossil record to assay how much before skeletonization the actual divergences between the phyla really were. Nonetheless, exceptional soft bodied preservation in rocks older than the first skeletonized phyla offers some maximum age constraints, although the difference between the minima and maxima is in the order of ~85 million years (Benton et al., 2015).

Another group whose preservation potential appears to have changed dramatically during its history are the Scleractinian corals. Based on molecular clock data, it appears that their crown group extends in the Carboniferous, perhaps some 300 million years ago, well before the oldest fossils in the Middle Triassic, some 240 million years ago—the inference is that there was a substantial history where they were unskeletonized and therefore invisible in the fossil record (Romano and Palumbi, 1996), with more than one independent skeletonization event much later in the Triassic (Stanley, 2003).

5 ESTIMATING DIVERGENCE TIMES FOR GROUPS WITH NO, OR VIRTUALLY NO, FOSSIL RECORD—THE VALUE OF HYPOTHESIS TESTING

The entire discussion above on using the fossil record to constrain the absolute divergence times between lineages is predicated upon the assumption that the focal clade and, for some of the methods, the outgroups are known from at least several well diagnosed and dated fossils. However, for many groups, there is virtually no fossil record, or no fossil record at all. In these cases, well constrained calibrated timetrees are obviously difficult to obtain. Nonetheless, I want to make the case that hypothesis testing is still possible, especially if the minimum age estimate for a divergence time leads to a timetree that yields older dates than that proposed by the hypothesis—sometimes testing hypotheses is much less demanding of the data than trying to reconstruct the actual history of a group.

5.1 Virtually No Fossil Record

Paleobiologists typically work with groups with tens to tens of thousands of fossil species (e.g., trilobites are known from some 20,000 species). However, some groups are known from just a few species. Thus—for example, none of the 20,000–30,000 living species of orchid are known from the fossil record, and only three unequivocal extinct species are known from the fossil record (Ramírez et al., 2007; Conran et al., 2009). With such an awful fossil record, it is difficult to estimate reasonable maxima for the divergences within the orchids, or for the group as a whole (but see Section 3.4). Yet, even with the first fossil described, which has poor age constraints [anywhere from 15 to 20 million years, the degree of uncertainty associated with the difficulty of dating amber from the Dominican Republic (Iturralde-Vinent and MacPhee, 1996)], it was possible to establish that the orchid crown group extends into the Cretaceous, refuting the hypothesis of Cenozoic origins (Ramírez et al., 2007). This result was obtained by simply using the fossil to date one node in an ultrametric tree, a result further supported in a Bayesian analysis using all three fossils (Gustafsson et al., 2010).

5.2 With No Fossil Record

Particularly at lower taxonomic ranks, many groups have no fossil record, neither do their immediate outgroups. Nonetheless, despite the lack of direct temporal data, average rates of molecular evolution estimated for closely related groups can sometimes provide valuable temporal data. For example, using an insect-wide molecular rate of ~1.5% change/million years for the mitochondrial COI gene, Quek et al. (2007) were able to refute the hypothesis that diversification of mitochondrial lineages of *Crematogaster* ants from the Sunda Shelf (peninsular Malaysia, Borneo, Sumatra, Java, etc.)

TABLE 5 | Relative magnitude of major factors that challenge our ability to estimate robust soft maxima on divergence times.

	Temporal, geographic, and taxonomic scale		
	Shallow time	**Deeper time**	**Deepest time**
	< ~2 million yrs Quaternary ~Local scale ~Genus	~2 to ~540 million yrs Pliocene–Cambrian Regional to global scale ~Family, order, class	> ~540 million yrs Precambrian ~Global scale ~Phylum, kingdom
Coalescence	Often large	Typically unimportant, less than 1% the age of the clade	Unimportant
Dating errors (radiometric dates; biostratigraphy)	Can be large, but typically small	Typically unimportant, less than 1% the age of the clade	Can be important due to lack of effective biostratigraphy
ΔT_{Gap}	Can be very large; the fossil and rock record is often very spotty at this timescale	Small to large (depends on the group)	Typically large
$\Delta T_{Div\text{-}1stApo}$			
Preservation ~constant	Similar to ΔT_{Gap}	Smaller or much smaller than ΔT_{Gap}	N/A—most groups have changed their preservation potential
Preservation drops at clade base	N/A—most groups this young didn't change their preservation potential	Can be much larger than ΔT_{Gap}	Typically much larger than ΔT_{Gap}

was driven by the glacial-interglacial cycles that repeatedly exposed and drowned the Sunda Shelf over the last million years. Instead, it appears that the ant lineages diverged from 1 to 20 million years—the COI gene would have had to have evolved ~10 times faster than the 1.5% rate to support the glacial-interglacial hypothesis.

5.2.1 Almost All Clades Are Embedded in More Inclusive Clades That Have a Fossil Record

Almost all clades, at least within animals and plants, lie within more inclusive clades where minimum and maximum age constraints are available (e.g., see Benton et al., 2015). Thus, at some level, temporal constraints can always be found for most groups, even if the dating precision might be low within the unfossiliferous ingroup.

6 SUMMARY

The quality of temporal calibration is highly variable, depending on the group and the fossil record available. Nonetheless, some generalizations can be made as a function of the age of the group, and its correlates, the group's geographic range and species richness (**Table 5**). If care is taken with the paleontological calibrations themselves, and with judicious analysis of data with multiple approaches, robust timetrees are well within our grasp for many taxa.

AUTHOR CONTRIBUTIONS

CM is sole author, and thus was responsible for all of its components.

FUNDING

This work was partially supported by the Philip Sandford Boone Chair in Paleontology at the University of California, Berkeley.

ACKNOWLEDGMENTS

Many colleagues over the years have contributed to my thinking on how to calibrate molecular phylogenies and quantify the incompleteness of the fossil record, including but not limited to: Steve Holland, Michael Foote, Steve Wang, Walter Fitch, and Simon Tavaré. This manuscript was greatly improved with input from the following reviewers: Philip Donoghue, Michel Laurin, Juliana Sterli, Rachel Warnock, and Walter Joyce, as well as Rauri Bowie and Carl Rothfels. Thanks to Jean-Bernard Caron for input on the dating of the Burgess Shale. The following also provided valuable input on previous incarnations of this manuscript, either directly or indirectly: David Hillis, Joe Felsenstein, Wayne Maddison, Nick Matzke, Josh Schraiber, and Tracy Heath. This work was supported in part by the Philip Sandford Boone Chair in Paleontology.

REFERENCES

Ali, M. S. M. (1983). The paleogeographic distribution of Clypeaster (Echinoidea) during the Cenozoic Era. *Neues Jahrb. für Geol. und Paläontologie Monatshefte* 8, 449–464.

Alroy, J. (1999). The fossil record of North American mammals: evidence for a Paleocene evolutionary radiation. *Syst. Biol.* 48, 107–118. doi: 10.1080/106351599260472

Antoine, P. O., Orliac, M. J., Atici, G., Ulusoy, I., Sen, E., Çubukçu, H. E., et al. (2012). A rhinocerotid skull cooked-to-death in a 9.2 ma-old ignimbrite flow of Turkey. *PLoS One* 7, 1–12. doi: 10.1371/journal.pone.0049997

Bapst, D. W. (2013). A stochastic rate-calibrated method for time-scaling phylogenies of fossil taxa. *Methods Ecol. Evol.* 4, 724–733. doi: 10.1111/2041-210X.12081

Barba-montoya, J., Reis, M., Schneider, H., Donoghue, P. C. J., and Yang, Z. (2018). Constraining uncertainty in the timescale of angiosperm evolution and the veracity of a cretaceous terrestrial revolution. *New Phytol.* 218, 819–834. doi: 10.1111/nph.15011

Behrensmeyer, A. K., Kidwell, S. M., and Gastaldo, R. A. (2000). Taphonomy and paleobiology. *Paleobiology* 26 (S4), 103–147. doi: 10.1017/S0094837300026907

Benton, M. J., and Donoghue, P. C. J. (2007). Paleontological evidence to date the tree of life. *Mol. Biol. Evol.* 24, 26–53. doi: 10.1093/molbev/msl150

Benton, M. J., Donoghue, P. C. J., Asher, R. J., Friedman, M., Near, T. J., and Vinther, J. (2015). Constraints on the timescale of animal evolutionary history. *Palaeontol. Electron.* 18, 1–106. doi: 10.26879/424

Bottjer, D. J., and Jablonski, D. (1988). Paleoenvironmental patterns in the evolution of post-paleozoic benthic marine invertebrates. *Palaios* 3, 540–560. doi: 10.2307/3514444

Burgess, S. D., Bowring, S., and Shen, S. (2014). High-precision timeline for Earth's most severe extinction. *Proc. Natl. Acad. Sci.* 111, 3316–3321. doi: 10.1073/pnas.1317692111

Clarke, J. T., Warnock, R. C. M., and Donoghue, P. C. J. (2011). Establishing a time-scale for plant evolution. *New Phytol.* 192, 266–301. doi: 10.1111/j.1469-8137.2011.03794.x

Cohen, A. S., Soreghan, M. J., and Scholz, C. A. (1993). Estimating the age of formation of lakes: an example from Lake Tanganyika, East African Rift system. *Geology* 21, 511–514. doi: 10.1130/0091-7613(1993)021<0511:ETAOFO>2.3.CO;2

Cohen, K. M., Finney, S. C., Gibbard, P. L., and Fan, J.-X. (2013). The ICS international chronostratigraphic chart. *Episodes* 36, 199–204. doi: 10.18814/epiiugs/2013/v36i3/002

Coiro, M., Doyle, J. A., and Hilton, J. (2019). How deep is the conflict between molecular and fossil evidence on the age of angiosperms? *New Phytol.* 223, 83–99. doi: 10.1111/nph.15708

Conran, J. G., Bannister, J. M., and Lee, D. E. (2009). Earliest orchid macrofossils: early miocene dendrobium and earina(orchidaceae: epidendroideae) from New Zealand. *Am. J. Bot.* 96, 466–474. doi: 10.3732/ajb.0800269

De Baets, K., Antonelli, A., and Donoghue, P. C. J. (2016). Tectonic blocks and molecular clocks. *Philos. Trans. R. Soc. B. Biol. Sci.* 371, 1–12. doi: 10.1098/rstb.2016.0098

Didier, G., Fau, M., and Laurin, M. (2017). Likelihood of tree topologies with fossils and diversification rate estimation. *Syst. Biol.* 66, 964–987. doi: 10.1093/sysbio/syx045

Donoghue, P. C. J., and Benton, M. J. (2007). Rocks and clocks: calibrating the tree of life using fossils and molecules. *Trends Ecol. Evol.* 22, 424–431. doi: 10.1016/j.tree.2007.05.005

Donoghue, P. C. J., and Yang, Z. (2016). The evolution of methods for establishing evolutionary timescales. *Philos. Trans. R. Soc. B Biol. Sci.* 371. 20160020. doi: 10.1098/rstb.2016.0020

Edwards, R. L., Gallup, C. D., and Cheng, H. (2003). Uranium-series dating of marine and lacustrine carbonates. *Rev. Mineral. Geochemistry* 52, 363–405. doi: 10.2113/0520363

Edwards, S. V., and Beerli, P. (2000). Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution (N. Y).* 54, 1839–1854. doi: 10.1111/j.0014-3820.2000.tb01231.x

Ezard, T. H. G., Aze, T., Pearson, P. N., and Purvis, A. (2011). Interplay between changing climate and species' ecology drives macroevolutionary dynamics. *Science* 332, 349–351. doi: 10.1126/science.1203060

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

Foote, M. (2007). Symmetric waxing and waning of marine invertebrate genera. *Paleobiology* 33, 517–529. doi: 10.1666/06084.1

Foote, M., Crampton, J. S., Beu, A. G., Marshall, B. A., Cooper, R. A., Maxwell, P. A., et al. (2007). Rise and fall of species occupancy in Cenozoic fossil mollusks. *Science* 318, 1131–1134. doi: 10.1126/science.1146303

Foote, M., Hunter, J., Janis, C. M., and Sepkoski, J. J. Jr. (1999). Evolutionary and preservational constraints on the origins of major biologic groups. *Science* 283, 1310–1314. doi: 10.1126/science.283.5406.1310

Friedman, M., and Brazeau, M. D. (2011). Sequences, stratigraphy and scenarios: what can we say about the fossil record of the earliest tetrapods? *Proc. R. Soc. B Biol. Sci.* 278, 432–439. doi: 10.1098/rspb.2010.1321

Friedman, M., Dornburg, A., Wainwright, P. C., Near, T. J., Eytan, R. I., Martin, C. H., et al. (2013). Molecular and fossil evidence place the origin of cichlid fishes long after Gondwanan rifting. *Proc. R. Soc. B Biol. Sci.* 280, 20131733–20131733. doi: 10.1098/rspb.2013.1733

Gavryushkina, A., Heath, T. A., Ksepka, D. T., Stadler, T., Welch, D., and Drummond, A. J. (2017). Bayesian total-evidence dating reveals the recent crown radiation of penguins. *Syst. Biol.* 66, 57–73. doi: 10.1093/sysbio/syw060

Gavryushkina, A., Welch, D., Stadler, T., and Drummond, A. J. (2014). Bayesian inference of sampled ancestor trees for epidemiology and fossil calibration. *PLoS Comput. Biol.* 10. e1003919. doi: 10.1371/journal.pcbi.1003919

Gustafsson, A. L. S., Verola, C. F., and Antonelli, A. (2010). Reassessing the temporal evolution of orchids with new fossils and a Bayesian relaxed clock, with implications for the diversification of the rare South American genus Hoffmannseggella (Orchidaceae: Epidendroideae). *BMC Evol. Biol.* 10, 177. doi: 10.1186/1471-2148-10-177

Haile-Selassie, Y. (2001). Late Miocene hominids from the. *Nature* 412, 178–181. doi: 10.1038/35084063

Heath, T. A., Huelsenbeck, J. P., and Stadler, T. (2014). The fossilized birth-death process for coherent calibration of divergence-time estimates. *Proc. Natl. Acad. Sci.* 111, E2957–E2966. doi: 10.1073/pnas.1319091111

Hedman, M. M. (2010). Constraints on clade ages from fossil outgroups. *Paleobiology* 36, 16–31. doi: 10.1666/0094-8373-36.1.16

Ho, S. Y. W., and Phillips, M. J. (2009). Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Syst. Biol.* 58, 367–380. 20150194. doi: 10.1093/sysbio/syp035

Ho, S. Y. W., Tong, K. J., Foster, C. S. P., Ritchie, A. M., Lo, N., and Crisp, M. D. (2015). Biogeographic calibrations for the molecular clock. *Biol. Lett.* 11. doi: 10.1098/rsbl.2015.0194

Holland, S. M. (1995). The stratigraphic distribution of fossils. *Paleobiology* 21, 92–109. doi: 10.1017/S0094837300013099

Holland, S. M. (2016). The non-uniformity of fossil preservation. *Philos. Trans. R. Soc. B Biol. Sci.* 371, 0–2. doi: 10.1098/rstb.2015.0130

Inoue, J., Donoghue, P. C. J., and Yang, Z. (2010). The impact of the representation of fossil calibrations on bayesian estimation of species divergence times. *Syst. Biol.* 59, 74–89. doi: 10.1093/sysbio/syp078

Iturralde-Vinent, M. A., and MacPhee, R. D. E. (1996). Age and paleogeographical origin of Dominican amber. *Science* 273, 1850–1852. doi: 10.1126/science.273.5283.1850

Jaanusson, V. (1976). "Faunal dynamics in the middle Ordovician (Viruan) of Balto-Scandia", *The Ordovician System: Proceedings of a Paleontological Association Symposium*, *Birminghan*. Ed. M. G. Bassett (Cardiff: University of Wales Press and National Museum of Wales), 301–326.

Jin, Y. G., Wang, Y., Wang, W., Shang, Q. H., Cao, C. Q., and Erwin, D. H. (2000). Pattern of marine mass extinction near the Permian-Triassic boundary in South China. *Science* 289, 432–436. doi: 10.1126/science.289.5478.432

Joyce, W. G., Parham, J. F., Lyson, T. R., Warnock, R. C. M., and Donoghue, P. C. J. (2013). A divergence dating analysis of turtles using fossil calibrations: an example of best practices. *J. Paleontol.* 87, 612–634. doi: 10.1666/12-149

Landis, M. J. (2017). Biogeographic dating of speciation times using paleogeographically informed processes. *Syst. Biol.* 66, 128–144. doi: 10.1093/sysbio/syw040

Lindberg, F. A. (1983). "Atlantic Coastal Plain Region," in *Correlation of stratigraphic units of North America (COSUNA) Project*. Eds. O. E. Child, G. Steels, and A. Salvador. Tulsa Oklahoma: American Association of Petroleum Geologists (AAPG), 1.

Liow, L. H., and Stenseth, N. C. (2007). The rise and fall of species: implications for macroevolutionary and macroecological studies. *Proc. R. Soc. B Biol. Sci.* 274, 2745–2752. doi: 10.1098/rspb.2007.1006

Lloyd, G. T., Bapst, D. W., Friedman, M., and Davis, K. E. (2016). Probabilistic divergence time estimation without branch lengths: dating the origins of dinosaurs, avian flight and crown birds. *Biol. Lett.* 12, 20160609. doi: 10.1098/rsbl.2016.0609

Loeza-Quintana, T., and Adamowicz, S. J. (2018). Iterative calibration: a novel approach for calibrating the molecular clock using complex geological events. *J. Mol. Evol.* 86, 118–137. doi: 10.1007/s00239-018-9831-2

Luo, Z. X. (2007). Transformation and diversification in early mammal evolution. *Nature* 450, 1011–1019. doi: 10.1038/nature06277

Magallon, S. A. (2004). Dating lineages: molecular and paleontological approaches to the temporal framework of clades. *Int. J. Plant Sci.* 165, S7–S21. doi: 10.1086/383336

Marjanović, D., and Laurin, M. (2008). Assessing confidence intervals for stratigraphic ranges of higher taxa: the case of Lissamphibia. *Acta Palaeontol. Pol.* 53, 413–432. doi: 10.4202/app.2008.0305

Marshall, C. R. (1990a). Confidence intervals on stratigraphic ranges. *Paleobiology* 16, 1–10. doi: 10.1017/S0094837300009672

Marshall, C. R. (1990b). The fossil record and estimating divergence times between lineages: maximum divergene times and the importance of reliable phylogenies. *J. Mol. Evol.* 30, 400–408. doi: 10.1007/BF02101112

Marshall, C. R. (1995a). Distinguishing between sudden and gradual extinctions in the fossil record: predicting the position of the Cretaceous-Tertiary iridium anomaly using the ammonite fossil record on Seymour Island, Antarctica. *Geology* 23, 731–734. doi: 10.1130/0091-7613(1995)023<0731:DBSAGE>2.3.CO;2

Marshall, C. R. (1995b). "Stratigraphy, the true order of species originations and extinctions, and testing ancestor-descendant hypotheses among Caribbean Neogene bryozoans," in *New approaches to speciation in the fossil record*. Eds. D. H. Erwin and R. L. Anstey. (New York: Columbia University Press), 208–235.

Marshall, C. R. (1997). Confidence intervals on stratigraphic ranges with nonrandom distributions of fossil horizons. *Paleobiology* 23, 165–173. doi: 10.1017/S0094837300016766

Marshall, C. R. (1999). Fossil gap analysis supports early tertiary origin of trophically diverse avian orders: Comment. *Geology* 27, 95–96. doi: 10.1130/0091-7613(1999)027<0095:FGASET>2.3.CO;2

Marshall, C. R. (2006). Explaining the Cambrian "explosion" of animals. *Annu. Rev. Earth Planet. Sci.* 34, 355–384. doi: 10.1146/annurev.earth.33.031504.103001

Marshall, C. R. (2008). A simple method for bracketing absolute divergence times on molecular phylogenies using multiple fossil calibration points. *Am. Nat.* 171, 726–742. doi: 10.1086/587523

Marshall, C. R. (2010). "Using confidence intervals to quantify the uncertainty in the end-points of stratigraphic ranges," in *Quantitative methods in paleobiology*. Eds. J. Alroy, and G. Hunt (The Paleontological Society), 291–316. doi: 10.1017/S1089332600001911

Marshall, C. R. (2017). Five palaeobiological laws needed to understand the evolution of the living biota. *Nat. Ecol. Evol.* 1, 0165. doi: 10.1038/s41559-017-0165

Marshall, C. R., Finnegan, S., Clites, E. C., Holroyd, P. A., Bonuso, N., Cortez, C., et al. (2018). Quantifying the dark data in museum fossil collections as palaeontology undergoes a second digital revolution. *Biol. Lett.* 14, 2–5. doi: 10.1098/rsbl.2018.0431

Marshall, C. R., and Swift, H. (1992). DNA-DNA hybridization phylogeny of sand dollars and highly reproducible extent of hybridization values. *J. Mol. Evol* 34, 31–44. doi: 10.1007/BF00163850

Marshall, C. R., and Valentine, J. W. (2010). The importance of preadapted genomes in the origin of the animal bodyplans and the cambrian explosion. *Evolution (N. Y).* 64, 1189–1201. doi: 10.1111/j.1558-5646.2009.00908.x

Marshall, C. R., and Ward, P. D. (1996). Sudden and gradual molluscan extinctions in the latest Cretaceous of western European Tethys. *Science* 274, 1360–1363. doi: 10.1126/science.274.5291.1360

Marshall, C., and Schultze, H.-P. (1992). Relative importance of molecular, neontological, and paleontological data in understanding the biology of the vertebrate invasion of land. *J. Mol. Evol.* 35, 93–101. doi: 10.1007/BF00183220

Martin, R. D. (1993). Primate origins: plugging the gaps. *Nature* 363, 223–234. doi: 10.1038/363223a0

Martínez, S., and Mooi, R. (2005). Extinct and extant sand dollars (Clypeasteroida: Echinoidea) from Uruguay. *Rev. Biol. Trop.* 53, 1–7.

McArthur, J. M., Howarth, R. J., and Shields, G. A. (2012). "Strontium isotope stratigraphy," in *the Geologic time scale 2012*. Eds. Gradstein, F. M., Ogg, J. G., Schmitz, M., and Ogg, G. (Boston: Elsevier), 127–144. doi: 10.1016/B978-0-444-59425-9.00007-X

Min, Z., and Schultze, H.-P. (2001). "Interrelationships of basal osteichthyans," in *Major events in early vertebrate evolution*. Ed. P. E. Ahlberg (Taylor & Francis London), 289–314.

Mooi, R., Martínez, S., and Parma, S. G. (2000). Phylogenetic systematics of tertiary monophorasterid sand dollars (Clypeasteroida: Echinoidea) from South America. *J. Paleontol.* 74, 263–281. doi: 10.1017/S0022336000031486

Mooi, R., and Peterson, D. (2000). A new species of Leodia (Clypeasteroida: Echinoidea) from the Neogene of Venezuela and its importance in the phylogeny of mellitid sand dollars. *J. Paleontol.* 74, 1083–1092. doi: 10.1017/S0022336000017637

Near, T. J., Meylan, P. A., and Shaffer, H. B. (2005). Assessing concordance of fossil calibration points in molecular clock studies: an example using turtles. *Am. Nat.* 165, 137–146. doi: 10.2307/3473141

Near, T. J., and Sanderson, M. J. (2004). Assessing the quality of molecular divergence time estimates by fossil calibrations and fossil-based model selection. *Philos. Trans. R. Soc. B Biol. Sci.* 359, 1477–1483. doi: 10.1098/rstb.2004.1523

Nowak, M. D., Smith, A. B., Simpson, C., and Zwickl, D. J. (2013). A simple method for estimating informative node age priors for the fossil calibration of molecular divergence time analyses. *PLoS One* 8, e66245. doi: 10.1371/journal.pone.0066245

O'Reilly, J. E., dos Reis, M., and Donoghue, P. C. J. (2015). Dating tips for divergence-time estimation. *Trends Genet.* 31, 637–650. doi: 10.1016/j.tig.2015.08.001

Ou, C.-Y., Ciesielski, C. A., Myers, G., Bandea, C. I., Luo, C.-C., Korber, B. T. M., et al. (1992). Molecular epidemiology of HIV transmission in a dental practice. *Science* 256, 1165–1171. doi: 10.1126/science.256.5060.1165

Parham, J. F., Donoghue, P. C. J., Bell, C. J., Calway, T. D., Head, J. J., Holroyd, P. A., et al. (2012). Best practices for justifying fossil calibrations. *Syst. Biol.* 61, 346–359. doi: 10.1093/sysbio/syr107

Patzkowsky, M. E., and Holland, S. M. (2012). *Stratigraphic paleobiology: understanding the distribution of fossil taxa in time and space*. Chicago: University of Chicago Press. doi: 10.7208/chicago/9780226649399.001.0001

Peng, S., Babcock, L. E., and Cooper, R. A. (2012). *The Cambrian Period*. Eds. F. M. Gradstein, J. G. Ogg, M. Schmitz, and G. Ogg (Elsevier). doi: 10.1016/B978-0-444-59425-9.00019-6

Pereira, A. G., Sterli, J., Moreira, F. R. R., and Schrago, C. G. (2017). Multilocus phylogeny and statistical biogeography clarify the evolutionary history of major lineages of turtles. *Mol. Phylogenet. Evol.* 113, 59–66. doi: 10.1016/j.ympev.2017.05.008

Pyron, R. A. (2011). Divergence time estimation using fossils as terminal taxa and the origins of lissamphibia. *Syst. Biol.* 60, 466–481. doi: 10.1093/sysbio/syr047

Quek, S. P., Davies, S. J., Ashton, P. S., Itino, T., and Pierce, N. E. (2007). The geography of diversification in mutualistic ants: a gene's-eye view into the Neogene history of Sundaland rain forests. *Mol. Ecol.* 16, 2045–2062. doi: 10.1111/j.1365-294X.2007.03294.x

Quental, T. B., and Marshall, C. R. (2010). Diversity dynamics: molecular phylogenies need the fossil record. *Trends Ecol. Evol.* 25, 435–441. doi: 10.1016/j.tree.2010.05.002

Ramírez, S. R., Gravendeel, B., Singer, R. B., Marshall, C. R., and Pierce, N. E. (2007). Dating the origin of the orchidaceae from a fossil orchid with its pollinator. *Nature* 448, 1042–1045. doi: 10.1038/nature06039

Raup, D. M. (1985). Mathematical models of cladogenesis. *Paleobiology* 11, 42–52. doi: 10.1017/S0094837300011386

Reis, M. D., Gunnell, G. F., Barba-Montoya, J., Wilkins, A., Yang, Z., and Yoder, A. D. (2018). Using phylogenomic data to explore the effects of relaxed clocks and calibration strategies on divergence time estimation: primates as a test case, *Syst. Biol.* 67, 594–615. doi: 10.1093/sysbio/syy001

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, 1291–1300. doi: 10.1890/08-0316.1

Roberts, D. L., and Solow, A. R. (2003). Flightless birds: when did the dodo become extinct? *Nature* 426, 245–245. doi: 10.1038/426245a

Romano, S. L., and Palumbi, S. R. (1996). Evolution of scieractinian corals inferred from molecular systematics. *Science* 271, 640–642. doi: 10.1126/science.271.5249.640

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. *Syst. Biol.* 61, 973–999. doi: 10.1093/sysbio/sys058

Sadler, P. M. (1981). Sediment accumulation rates and the completeness of stratigraphic sections. *J. Geol.* 89, 569–584. doi: 10.1086/628623

Sauquet, H., Ho, S. Y. W., Gandolfo, M. A., Jordan, G. J., Wilf, P., Cantrill, D. J., et al. (2012). Testing the impact of calibration on molecular divergence times using a fossil-rich group: the case of nothofagus (Fagales). *Syst. Biol.* 61, 289–313. doi: 10.1093/sysbio/syr116

Sawada, Y., Saneyoshi, M., Nakayama, K., Sakai, T., Itaya, T., Hyodo, M., et al. (2006). "The ages and geological backgrounds of Miocene hominoids Nacholapithecus, Samburupithecus, and Orrorin from Kenya," in *Human origins and environmental backgrounds* (USA: Springer), 71–96. doi: 10.1007/0-387-29798-7_6

Shaffer, H. B., McCartney-Melstad, E., Near, T. J., Mount, G. G., and Spinks, P. Q. (2017). Phylogenomic analyses of 539 highly informative loci dates a fully resolved time tree for the major clades of living turtles (Testudines). *Mol. Phylogenet. Evol.* 115, 7–15. doi: 10.1016/j.ympev.2017.07.006

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

Smith, A. B. (1994). Systematics and the fossil record: documenting evoutionary patterns. *Blackwell Science*. Oxford, UK: Blackwell.

Solow, A. R., and Roberts, D. L. (2003). A nonparametric test for extinction based on a sighting record. *Ecology* 84, 1329–1332. doi: 10.1890/0012-9658(2003)084[1329:ANTFEB]2.0.CO;2

Souto, C., Mooi, R., Martins, L., Menegola, C., and Marshall, C. R. (2019). Homoplasy and extinction: the phylogeny of the cassidulid echinoids (Echinodermata). *Zool. J. Linn. Soc* XX, 1–39. doi: 10.1093/zoolinnean/zlz060

Sprain, C. J., Renne, P. R., Vanderkluysen, L., Pande, K., Self, S., and Mittal, T. (2019). The eruptive tempo of deccan volcanism in relation to the cretaceous-paleogene boundary. *Science* 363, 866–870. doi: 10.1126/science.aav1446

Springer, M. S. (1995). Molecular clocks and the incompleteness of the fossil record. *J. Mol. Evol.* 41, 531–538. doi: 10.1007/BF00175810

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

Stanley, G. D. J. (2003). The evolution of modern corals and their early history. *Earth-Science Rev.* 60, 195–225. doi: 10.1016/S0012-8252(02)00104-6

Steiper, M. E., and Young, N. M. (2008). Timing primate evolution: Lessons from the discordance between molecular and paleontological estimates. *Evol. Anthropol.* 17, 179–188. doi: 10.1002/evan.20177

Sterli, J., Pol, D., and Laurin, M. (2013). Cladistics palaeontological dating and the timing of turtle diversification. *Cladistics* 29, 233–246. doi: 10.1111/j.1096-0031.2012.00425.x

Stirling, C. H., Esat, T., Lambeck, K., McCulloch, M., Blake, S., Lee, D.-C., et al. (2001). Orbital forcing of the marine isotope stage 9 interglacial. *Science* 291, 290–293. doi: 10.1126/science.291.5502.290

Strauss, D., and Sadler, P. M. (1989). Classical confidence intervals and Bayesian probability estimates for ends of local taxon ranges. *Math. Geol.* 21, 411–427. doi: 10.1007/BF00897326

Tavaré, S., Marshall, C. R., Will, O., Soligo, C., and Martin, R. D. (2002). Using the fossil record to estimate the age of the last common ancestor of extant primates. *Nature* 416, 726–729. doi: 10.1038/416726a

Taylor, M. E. (1987). "Biostratigraphy and paleobiogeography," in *Fossil invertebrates*". Eds. R. S. Boardman, A. H. Cheetham, and A. J. Rowell (Palo Alto California: Blackwell Scientific Publications), 52–66.

Taylor, R. E., and Bar-Yosef, O. (2014). *Radiocarbon dating: an archaeological perspective*. New York: Routledge.

Valentine, J. W. (1989). How good was the fossil record? Clues from the Californian Pleistocene. *Paleobiology* 15, 83–94. doi: 10.1017/S0094837300009295

Van Damme, D., and Pickford, M. (2003). The late cenozoic thiaridae (Mollusca, Gastropoda, Cerithioidea) of the albertine rift valley (Uganda-Congo) and their bearing on the origin and evolution of the Tanganyikan thalassoid malacofauna. *Hydrobiologia* 498, 1–83. doi: 10.1023/A:1026298512117

Wagner, P. J. (1998). A likelihood approach for evaluating estimates of phylogenetic relationships among fossil taxa. *Paleobiology* 24, 430–449. doi: 10.1017/S0094837300020091

Wagner, P. J. (2019). On the probabilities of branch durations and stratigraphic gaps in phylogenies of fossil taxa when rates of diversification and sampling vary over time. *Paleobiology* 45, 30–55. doi: 10.1017/pab.2018.35

Wagner, P. J., and Marcot, J. D. (2013). Modelling distributions of fossil sampling rates over time, space and taxa: assessment and implications for macroevolutionary studies. *Methods Ecol. Evol.* 4, 703–713. doi: 10.1111/2041-210X.12088

Wang, S. C., Chudzicki, D. J., and Everson, P. J. (2009). Optimal estimators of the position of a mass extinction when recovery potential is uniform. *Paleobiology* 35, 447–459. doi: 10.1666/0094-8373-35.3.447

Wang, S. C., Everson, P. J., Zhou, H. J., Park, D., and Chudzicki, D. J. (2016). Adaptive credible intervals on stratigraphic ranges when recovery potential is unknown. *Paleobiology* 42, 240–256. doi: 10.1017/pab.2015.37

Wang, S. C., and Marshall, C. R. (2016). Estimating times of extinction in the fossil record. *Biol. Lett.* 12, 20150989. doi: 10.1098/rsbl.2015.0989

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. *Proc. R. Soc. B. Biol. Sci.* 282, 20141013. doi: 10.1098/rspb.2014.1013

Warnock, R. C. M., Yang, Z., and Donoghue, P. C. J. (2017). Testing the molecular clock using mechanistic models of fossil preservation and molecular evolution. *Proc. R. Soc. B Biol. Sci.* 284, 20170227. doi: 10.1098/rspb.2017.0227

West, K., and Michel, E. (2000). Ancient lakes: biodiversity, ecology and evolution. *Adv. Ecol. Res.* 31, 331–354. doi: 10.1016/S0065-2504(00)31015-7

White, T. D., Asfaw, B., Beyene, Y., Haile-Selassie, Y., Lovejoy, C. O., Suwa, G., et al. (2009). Ardipithecus ramidus and the paleobiology of early hominids. *Science* 326, 75–87. doi: 10.1126/science.1175802

Wilkinson, R. D., Steiper, M. E., Soligo, C., Martin, R. D., Yang, Z., and Tavaré, S. (2011). Dating primate divergences through an integrated analysis of palaeontological and molecular data. *Syst. Biol.* 60, 16–31. doi: 10.1093/sysbio/syq054

Wilson, A. B., Glaubrecht, M., and Meyer, A. (2004). Ancient lakes as evolutionary reservoirs: evidence from the thalassoid gastropods of Lake Tanganyika. *Proc. R. Soc. B Biol. Sci.* 271, 529–536. doi: 10.1098/rspb.2003.2624

Yang, Z., and Rannala, B. (2006). Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Mol. Biol. Evol.* 23, 212–226. doi: 10.1093/molbev/msj024

Zhang, C., Stadler, T., Klopfstein, S., Heath, T. A., and Ronquist, F. (2016). Total-evidence dating under the fossilized birth-death process. *Syst. Biol.* 65, 228–249. doi: 10.1093/sysbio/syv080

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Marshall. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

APPENDIX A: RELATIONSHIP BETWEEN THE NUMBER OF FOSSIL LOCALITIES AND AGE OF THE *FAD* OF A TAXON, AND THE MEAN AND VARIANCE OF THE EQUIVALENT LOGNORMAL PRIOR

Following Section 3.1.2.1, the desired lognormal prior should have a mode, $e^{\left(\mu-\sigma^2\right)}$, which extends an average gap size (FAD/n) below the oldest fossil (*FAD*). It should also have a 95% tail, equivalent to the $p = 0.95$ quantile, $e^{\left(\mu+\sqrt{2}\sigma erf^{-1}\left(2p-1\right)\right)}$, extending $FAD(1 - p)^{-1/n}$ beyond the age of the *FAD*. Thus, we have two unknowns (μ and σ^2) and two equations, one for the mode and one for the 95% quantile:

$$e^{\left(\mu-\sigma^2\right)} = FAD/n \tag{A1}$$

$$e^{\left(\mu+\sqrt{2}\sigma erf^{-1}\left(2p-1\right)\right)} = FAD\left(1-p\right)^{-1/n} \tag{A2}$$

Taking natural logarithms, (A1) and (A2) become, respectively:

$$\mu-\sigma^2 = \ln\left(FAD\right)-\ln\left(n\right) \tag{A3}$$

$$\mu+\sqrt{2}\sigma erf^{-1}\left(2p-1\right) = \ln\left(FAD\right)-\left(\ln\left[1-p\right]\right)/n \tag{A4}$$

Rearranging equation (A3) gives the mean, μ:

$$\mu = \ln\left(FAD\right)-\ln\left(n\right)+\sigma^2 \tag{A5}$$

which is equation (6) in the text. We now need an expression for the variance, σ^2. This can be found by substituting equation (A5) into equation (A4). With some re-arranging we have:

$$\sigma^2+\sigma\sqrt{2}erf^{-1}\left(2p-1\right)+\left(\ln\left[1-p\right]/n-\ln\left[n\right]\right)=0 \tag{A6}$$

We can solve for σ using the quadratic formula:

$$\sigma = \left(-b \pm\sqrt{b^2-4ac}\right)/2a$$

where $a = 1$, $b=\sqrt{2}erf^{-1}\left(2p-1\right)$, and $c = (\ln)[1 - p]/ n - \ln[n])$. Given that $erf^{-1}(2p - 1) = 1.163087$ and $\ln (1- p) = -2.9957$ for $p = 0.95$, we have:

$$\sigma = -\sqrt{2}\left(1.163087\right)/2 \pm 0.5\sqrt{\left(\sqrt{2}\left(1.163087\right)\right)^2-4(-2.9957\ /n-\ln\left[n\right])}$$

After simplification and squaring, we arrived at equation (7) in the text:

$$\sigma^2 = \left(-0.8224+0.5\sqrt{2.7055-4(-2.9957\ /n-\ln\left[n\right])}\right)^2 \tag{A8}$$

The decision to add rather than subtract $\sqrt{b^2-4ac}$ was based on numerical tests to determine which of the two options returned the correct mode and 95th percentile.

frontiers in Genetics

SYSTEMATIC REVIEW
published: 29 November 2019
doi: 10.3389/fgene.2019.01241

Evolutionary Models for the Diversification of Placental Mammals Across the KPg Boundary

Mark S. Springer[1], Nicole M. Foley[2], Peggy L. Brady[1], John Gatesy[3] and William J. Murphy[2]*

[1] Department of Evolution, Ecology, and Evolutionary Biology, University of California, Riverside, Riverside, CA, United States, [2] Department of Veterinary Integrative Biosciences, Texas A&M University, College Station, TX, United States, [3] Division of Vertebrate Zoology, American Museum of Natural History, New York, NY, United States

OPEN ACCESS

Edited by:
Michel Laurin, UMR7207 Centre de recherche sur la paléobiodiversité et les paléoenvironnements (CR2P), France

Reviewed by:
Emmanuel Gheerbrant, UMR7207 Centre de recherche sur la paléobiodiversité et les paléoenvironnements (CR2P), France
Carlos G. Schrago, Federal University of Rio de Janeiro, Brazil
J. David Archibald, San Diego State University, United States

***Correspondence:**
Mark S. Springer
springer@ucr.edu

Specialty section:
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Genetics

Received: *28 June 2019*
Accepted: *08 November 2019*
Published: *29 November 2019*

Citation:
Springer MS, Foley NM, Brady PL, Gatesy J and Murphy WJ (2019) Evolutionary Models for the Diversification of Placental Mammals Across the KPg Boundary. Front. Genet. 10:1241. doi: 10.3389/fgene.2019.01241

Deciphering the timing of the placental mammal radiation is a longstanding problem in evolutionary biology, but consensus on the tempo and mode of placental diversification remains elusive. Nevertheless, an accurate timetree is essential for understanding the role of important events in Earth history (e.g., Cretaceous Terrestrial Revolution, KPg mass extinction) in promoting the taxonomic and ecomorphological diversification of Placentalia. Archibald and Deutschman described three competing models for the diversification of placental mammals, which are the *Explosive*, *Long Fuse*, and *Short Fuse Models*. More recently, the *Soft Explosive Model* and *Trans-KPg Model* have emerged as additional hypotheses for the placental radiation. Here, we review molecular and paleontological evidence for each of these five models including the identification of general problems that can negatively impact divergence time estimates. The *Long Fuse Model* has received more support from relaxed clock studies than any of the other models, but this model is not supported by morphological cladistic studies that position Cretaceous eutherians outside of crown Placentalia. At the same time, morphological cladistics has a poor track record of reconstructing higher-level relationships among the orders of placental mammals including the results of new pseudoextinction analyses that we performed on the largest available morphological data set for mammals (4,541 characters). We also examine the strengths and weaknesses of different timetree methods (node dating, tip dating, and fossilized birth-death dating) that may now be applied to estimate the timing of the placental radiation. While new methods such as tip dating are promising, they also have problems that must be addressed if these methods are to effectively discriminate among competing hypotheses for placental diversification. Finally, we discuss the complexities of timetree estimation when the signal of speciation times is impacted by incomplete lineage sorting (ILS) and hybridization. Not accounting for ILS results in dates that are older than speciation events. Hybridization, in turn, can result in dates than are younger or older than speciation dates. Disregarding this potential variation in "gene" history across the genome can distort phylogenetic branch lengths and divergence estimates when multiple unlinked genomic loci are combined together in a timetree analysis.

Keywords: KPg boundary, placental radiation, relaxed clocks, timetrees, tip dating

INTRODUCTION

Placentalia is the crown clade of eutherian mammals and includes 18–19 different orders with living representatives plus other major groups that are entirely extinct (e.g., Meridiungulata, Creodonta, Dinocerata, Mesonychia, Embrithopoda, Desmostylia, and Leptictida). Resolving the timing of the placental mammal radiation, both between orders and within orders, is a longstanding problem in evolutionary biology (Szalay, 1968; Gingerich, 1977; Young, 1981; Springer et al., 2003; Meredith et al., 2011; dos Reis et al., 2012; Halliday et al., 2019). Elucidation of the timing of this radiation has important implications for understanding the role of the KPg mass extinction in promoting the radiation of placental mammals.

The traditional view based on paleontology is that placental mammals began to diversity near the end of the Cretaceous, but with the bulk of the interordinal radiation occurring after the KPg mass extinction ~66 Ma (e.g., Szalay, 1968; Gingerich, 1977; Young, 1981; Carroll, 1988). Nevertheless, some paleontologists have allowed for the possibility that incipient cladogenesis among extant placental lineages may have occurred as far back as 85–80 Ma (Szalay, 1968; Young, 1981; Carroll, 1988) or even as far back as the Early Cretaceous. For example, McKenna and Bell (1997) included 22 genera from the Late Cretaceous and one genus from the Early Cretaceous in the crown-group Placentalia.

Early studies based on molecular data employed strict or local molecular clocks (Dickerson, 1971; Li et al., 1990), sometimes with culling of genes for which a constant rate of evolution was rejected by likelihood ratio tests and/or linearized tree tests (Hedges et al., 1996; Kumar and Hedges, 1998), or with adjustments for rate variation based on a reference taxon (Springer, 1997; Springer et al., 1997). Quartet dating (Rambaut and Bromham, 1998) allowed for limited rate variation under a 2-rate model and was also applied to early divergences in the placental radiation (Eizirik et al., 2001; Madsen et al., 2001; Murphy et al., 2001a; Scally et al., 2001). The general consensus of these studies was that most interordinal cladogenesis occurred prior to the KPg boundary. Moreover, many of these studies pushed back the estimate for the most recent common ancestor of Placentalia to ~100 Ma or more (Li et al., 1990; Hedges et al., 1996; Kumar and Hedges, 1998; Eizirik et al., 2001; Madsen et al., 2001; Murphy et al., 2001b). Some of these studies also suggested that plate tectonic events could have been important drivers of the early placental radiation (Hedges et al., 1996; Springer et al., 1997; Eizirik et al., 2001; Murphy et al., 2001a; Murphy et al., 2001b; Scally et al., 2001; Wildman et al., 2007).

Archibald and Deutschman (2001) proposed three competing models for the placental radiation based on published paleontological and molecular dating studies. These are the *Explosive, Long Fuse*, and *Short Fuse Models* of diversification (**Figure 1**). The *Explosive Model* corresponds to the widely held view among paleontologists that most or all of the placental radiation occurred after the KPg mass extinction. This model also suggests a fundamental role for the mass extinction in promoting the interordinal radiation of placental mammals. The *Short Fuse* and *Long Fuse Models*, in turn, emerged from molecular-based studies and posit deeper temporal roots for the placental radiation. In the *Long Fuse Model*, interordinal cladogenesis is primarily concentrated before the KPg boundary whereas intraordinal cladogenesis occurred after the end Cretaceous mass extinction. The *Short Fuse Model* posits even more ancient interordinal cladogenesis, in some cases as far back as the Jurassic, along with the beginnings of intraordinal cladogenesis in numerous orders in the Cretaceous. These three models have provided a useful framework for subsequent studies of the placental radiation.

Several developments near the turn of the millennium helped to shape the last ~20 years of timetree studies on placental mammals. First, Thorne et al. (1998) developed a Bayesian Markov chain Monte Carlo (MCMC) method that allows each branch to have its own rate of evolution. Second, the traditional phylogeny for placental orders based on morphology (e.g., Szalay, 1977; Novacek, 1992) was overhauled by molecular studies that employed multigene data sets and improved models of sequence evolution. The results of these studies clustered placental orders into four major clades (Afrotheria, Xenarthra, Euarchontoglires, Laurasiatheria) with the additional clustering of Euarchontoglires and Laurasiatheria into Boreoeutheria (Springer and de Jong, 2001). Of these five clades only Xenarthra was recovered by previous morphological analyses. This overhaul began with the recognition of Afrotheria (Springer et al., 1997; Stanhope et al., 1998a; Stanhope et al., 1998b) and came to full fruition in multigene studies that provided robust support for the four major clades of placental mammals (Madsen et al., 2001; Murphy et al., 2001a; Murphy et al., 2001b; Scally et al., 2001; Waddell et al., 2001; Springer et al., 2004). All four major clades, as well as Boreoeutheria, have been corroborated by retroposon insertions (Nishihara et al., 2006; Möller-Krull et al., 2007; Nishihara et al., 2009). This thoroughly revised phylogeny for placental mammals also overturned previous molecular hypotheses based on mitogenomes (D'Erchia et al., 1996; Reyes et al., 2000; Arnason et al., 2002) and early analyses of nuclear genes with limited taxon sampling (e.g., Graur et al., 1991; Li et al., 1992; Graur et al., 1992) that positioned rodents or erinaceids (e.g., hedgehogs, moon rats) as the earliest branches of the placental tree. The root of the placental tree remains contentious, but is now centered on three competing hypotheses: Afrotheria versus Boreoeutheria + Xenarthra; Afrotheria + Xenarthra versus Boreoeutheria; and Xenarthra versus Afrotheria + Boreoeutheria (Scally et al., 2001; Murphy et al., 2007; Morgan et al., 2013; Romiguier et al., 2013; Tarver et al., 2016). Resolution of this uncertainty is important for dating the placental tree.

Most timetree studies based on well-corroborated molecular topologies have recovered the majority of interordinal divergences in the Cretaceous and are generally compatible with the *Long Fuse Model* (Springer et al., 2003; Delsuc et al., 2004; Springer et al., 2005; Murphy et al., 2007; Meredith et al., 2011; dos Reis et al., 2012; Tarver et al., 2016; Foley et al., 2016). An exception is the supertree analysis by Bininda-Emonds et al. (2007), which recovered even older divergence times that are generally compatible with the *Short Fuse Model*. By contrast, the authors of recent morphological cladistic studies have argued that their results provide renewed support for a strict version of the *Explosive Model* of diversification by positioning all

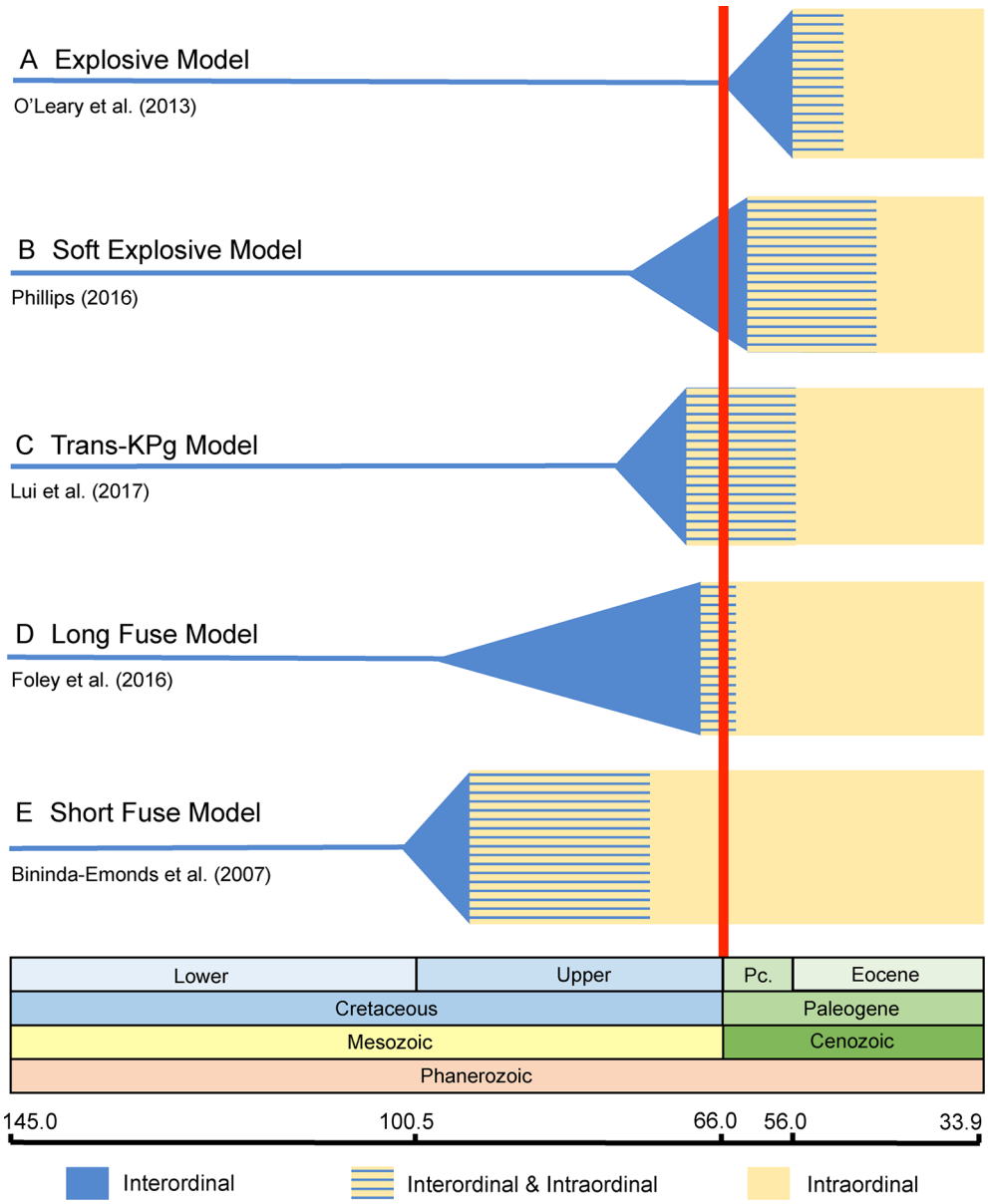


FIGURE 1 | Graphical summary of the five competing models of diversification for placental mammals. Approximate dates that were used to illustrate each model are derived from representative studies as indicated in the figure. **(A)** *Explosive Model*. **(B)** *Soft Explosive Model*. **(C)** *Trans-KPg Model*. **(D)** *Long Fuse Model*. **(E)** *Short Fuse Model*. For the *Short Fuse Model*, some molecular estimates for the base of Placentalia are older than the date obtained by Bininda-Emonds et al. (2007), e.g., Kumar and Hedges (1998) obtained a date of ~129 Ma.

Cretaceous taxa outside of Placentalia (Wible et al., 2007; Wible et al., 2009; O'Leary et al., 2013). For example, O'Leary et al. (2013) claimed that all interordinal cladogenesis occurred after the KPg mass extinction based on their combined analyses of morphology and molecules.

Recently, Phillips (2016) and Liu et al. (2017a) proposed new models of placental diversification that are intermediate between the *Long Fuse* and *Explosive Models*. Specifically, Phillips (2016) proposed the *Soft Explosive Model* and Liu et al. (2017a) proposed the *Trans-KPg Model* (**Figure 1**). These models are slight variations of the same theme, that interordinal diversification extended across the KPg and well into the Cenozoic, when most intraordinal diversification occurred, and bring the total number of models from three to five.

A common denominator of relaxed molecular clock analyses of the placental radiation is that they have generally relied on node-dating approaches that calibrate a rooted tree by constraining the age of one or more internal nodes (Springer et al., 2003; Delsuc et al., 2004; Springer et al., 2005; dos Reis et al., 2012; Emerling et al., 2015; Phillips, 2016; Foley et al., 2016; Tarver et al., 2016; Liu et al., 2017a). Node dating has recently come under scrutiny, perhaps most importantly because the maximum age and prior

probability distribution for a calibrated node are subjective (Heath et al., 2014; Lee and Palci, 2015). Given this and other potential problems with node dating, alternative approaches for timetree inference have gained more traction. One popular method is tip dating (Pyron, 2011; Ronquist et al., 2012). This approach was originally developed for dating evolutionary trees of RNA viruses with samples that were taken at different years (Rambaut, 2000). Dated tips provide a unique source of information for estimating rates of evolution and time-scaling a tree (Lee and Palci, 2015). Tip dating of sequential samples of RNA viruses was co-opted for use with evolutionary trees that include fossil organisms. To achieve this goal, the RNA clock for sequentially sampled viruses has been replaced with a morphological clock for phenotypic characters that are scored for extinct and extant taxa (Lee and Palci, 2015). Tip dating can also take advantage of molecular matrices for extant taxa, in which case the term "total evidence dating" is sometimes used because the data sets contain both molecular and morphological characters (Lee and Palci, 2015). For convenience we use the term tip dating for the remainder of this paper. In tip dating, the molecular and morphological data matrices are simultaneously used to estimate the phylogenetic placement of fossils and calibrate the tree (Arcila et al., 2015). An additional advantage of tip dating is that all extinct species for a given clade can be included in analysis, rather than just the oldest fossil as in node dating.

In addition to tip dating, Heath et al. (2014) suggested a new method for timetree estimation that uses a single model for the speciation-extinction-fossilization process. This model is known as the fossilized birth–death model and has only four parameters (speciation rate, extinction rate, fossil recovery rate, proportion of sampled extant species) that require prior assumptions. Fossilized birth–death dating can be implemented with tip dating (Gavryushkina et al., 2017), but in its original incarnation (Heath et al., 2014) fossilized birth-death dating was performed with molecular data only. A more recent implementation of fossilized birth–death dating requires fossil ages and a set of trees, but does not require molecular data (Didier and Laurin, 2018).

An additional issue that affects the estimation of species divergence times with molecular data is that coalescence times for individual genes are expected to exceed speciation times. For segments of the genome that disagree with the species tree because of incomplete lineage sorting (ILS), coalescence times will always exceed speciation times (Angelis and dos Reis, 2015; dos Reis et al., 2016). ILS, also known as deep coalescence when viewed from the perspective of looking back in time, occurs when alleles fail to coalesce in the most recent common ancestor of two taxa and instead coalesce deeper in the gene tree. A consequence of ILS is that divergence times on gene trees will overestimate speciation times. However, even gene segments that agree with the species tree are expected to have coalescence times that exceed speciation times. The opposite pattern may occur when two taxa hybridize with each other. Specifically, gene flow between two taxa, either involving portions of the nuclear genome and/or the mitogenome, will result in a divergence time estimate for these taxa that is younger than the actual speciation time for the same taxa when the introgressed DNA regions are employed in timetree analyses. Recent studies suggest that extensive introgression has occurred in several mammalian clades (e.g., Li et al., 2016; Árnason et al., 2018; Palkopoulou et al., 2018; Li et al., 2019), so this issue deserves consideration in future timetree studies given that all current molecular clock estimation models assume no gene flow among species lineages.

Here, (1) we review the supporting arguments and shortcomings of each of the five models of placental diversification, including the identification of general problems that can negatively impact divergence time estimates; (2) examine the pros and cons of different timetree methods (node dating, tip dating, fossilized birth–death dating) that may now be applied to estimate the timing of the placental radiation; and (3) discuss the complexities of timetree estimation when the genetic signal for speciation times is complicated by the coalescence process and hybridization (Hallström and Janke, 2008).

REVIEW AND COMPARISON OF MODELS

Explosive Model

The *Explosive Model* posits that the vast majority of placental cladogenesis, both interordinal and intraordinal, occurred near or after the KPg boundary (66 Ma) (**Figure 1A**) (Archibald and Deutschman, 2001). During the first ~10 million years of the Cenozoic, diversification of terrestrial placental taxa occurred rapidly in response to available niche space vacated by non-avian dinosaurs (Carroll, 1997; O'Leary et al., 2013). Support for *Explosive Model* is derived from direct reading of the fossil record and also from trees derived from the analysis of morphological data that exclude all or most Mesozoic eutherians from crown Placentalia (Gingerich, 1977; Archibald and Deutschman, 2001; Gingerich et al., 2001; Wible et al., 2009; Goswami et al., 2011; O'Leary et al., 2013; Davies et al., 2017). Instead, most Mesozoic eutherians are positioned as stem placental lineages (Archibald and Deutschman, 2001); throughout the remainder of our discussion, we refer to extinct eutherians that are outside of Placentalia as "stem placentals." Recent versions of the *Explosive Model*, which are based on cladistic analyses of large morphological and combined data sets (Wible et al., 2007; Wible et al., 2009; O'Leary et al., 2013; Halliday, et al., 2016; Halliday et al., 2017; Halliday et al., 2019), suggest an extreme version of the *Explosive Model* that is consistent with just a single placental ancestor crossing the KPg boundary.

A literal reading of the fossil record indicates that there is a striking increase in the abundance of extinct eutherian species on the Paleocene side of the KPg boundary. This increase (e.g. from 11 extinct eutherian species in the Late Cretaceous to 139 in the early Tertiary) is viewed as supporting evidence for the *Explosive Model* (Archibald and Deutschman, 2001). Several studies have investigated if this apparent increase is an artifact related to limited sampling in the Late Cretaceous (Alroy, 1999; Benton et al., 2000; Archibald and Deutschman, 2001; Davies et al., 2017). The resulting quantitative analyses suggest that the explosive increase in morphological and taxonomic diversity after the KPg boundary is biologically significant and is not due to a poor fossil record in the Cretaceous (Alroy, 1999; Davies et al., 2017).

Reconstructions of ancestral areas for placental mammals further suggest that the interordinal radiation of Boreoeutheria occurred in Eurasia and North America (Springer et al., 2011), areas that contain some of the best-known Late Cretaceous fossil localities. These results suggest that the current distribution of sampling localities should be sufficient to uncover Late Cretaceous crown boreoeutherian fossils if they are present (Phillips, 2016). A caveat is that there is no fossil record of Cretaceous eutherians in Africa so potential placental fossils on this continent remain unsampled (Phillips, 2016). Other landmasses with a poor or missing fossil record of eutherians from all or most of the Cretaceous include Antarctica, Madagascar, and India. Also, an important criticism of the *Explosive Model* of placental diversification is that it relies on the accurate phylogenetic placement of extinct eutherians from the Cretaceous as stem placentals. However, the placement of some extinct taxa is subject to significant uncertainty for a variety of reasons (see below). An additional criticism of extreme versions of the *Explosive Model* (e.g., O'Leary et al., 2013) is that the nucleotide substitution rates for basal branches of Placentalia would have been extremely high, more representative of DNA viruses than those typically observed in mammals, to fit the *Explosive Model* (Springer et al., 2013). While the *Explosive Model* is the hypothesis that is best supported by traditional interpretations of the fossil record, it has not yet been supported by any rigorous molecular analysis.

Soft Explosive Model

The *Soft Explosive Model* allows for cladogenesis among the major superordinal groups (Xenarthra, Afrotheria, Laurasiatheria, and Euarchontoglires) in the Cretaceous, but places the remainder of placental interordinal diversification near or after the KPg boundary (Phillips, 2016) (**Figure 1B**). Like the *Explosive Model*, this hypothesis suggests that the rapid interordinal diversification seen after the KPg boundary occurred in response to ecospace filling in the absence of non-avian dinosaurs (Phillips, 2016; Phillips and Fruciano, 2018). The *Soft Explosive Model* does not preclude a few crown placentals from the late Cretaceous, but suggests that the vast majority of Late Cretaceous eutherians are stem placentals rather than members of Placentalia. However, as discussed below there are significant problems with the placement of extinct mammalian orders based on parsimony or likelihood analyses of morphological characters. Phillips (2016) suggested that Cretaceous dates for most interordinal splits, as are commonly recovered in studies that support the *Long Fuse Model*, are the result of rate transference errors that can be avoided by removing fossil calibrations for taxa that are large and/or long-lived. These taxa generally have slower rates of molecular evolution relative to small-bodied, short-lived mammals with shorter generation times that might better approximate most early placental taxa (Bromham, 2011). However, recent analyses have shown that not calibrating large or long-lived taxa can result in zombie lineages, where taxa have a fossil record that is older than the divergence time estimated from molecular data (Springer et al., 2017). While ghost lineages are the expected result of an incomplete fossil record (Springer and Lilje, 1988; Strauss and Sadler, 1989; Springer, 1990; Marshall, 1997), zombie lineages are logically impossible if extinct taxa have been correctly identified because divergence times cannot be younger than minimum ages implied by the fossil record (Springer et al., 2017). Indeed, omitting or using too few fossil calibrations for large or long-lived taxa biases analyses to underestimate the ages of these lineages, and can also drag the ages of deeper nodes towards the present (Springer et al., 2017). This debate has continued in the literature with the focus once more returning to the issue of fossil calibrations (Phillips and Fruciano, 2018), which due to their somewhat subjective nature are a long recognized and ongoing source of conflict in node-dating analyses (Yang and Rannala, 2006; Donoghue and Benton, 2007; Inoue et al., 2009; Pyron, 2009; Parham et al., 2012).

Trans-KPg Model

The *Trans-KPg Model* is similar to the *Short Fuse Model* in suggesting that much of the interordinal diversification (cladogenesis) of placental mammals occurred after the KPg mass extinction. In contrast to the latter model, however, the *Trans-KPg Model* suggests that interordinal diversification was part of a continuous radiation in the Late Cretaceous and early Cenozoic that was uninterrupted by the KPg mass extinction (Liu et al., 2017a) (**Figure 1C**). The steady rate of interordinal diversification of placental mammals through time is proposed to coincide with a parallel radiation of herbivorous multituberculates in response to the gradual increase in ecological opportunity afforded by the rise of the angiosperms (Liu et al., 2017a). Similar to the *Soft Explosive Model*, timetrees that support the *Trans-KPg Model* are compromised by extensive zombie lineages (Gatesy and Springer, 2017; also see below), in addition to homology errors in the underlying data set (Gatesy and Springer, 2017). One reanalysis of Liu et al.'s (2017a) data set that purportedly corrected these homology errors yielded divergence time estimates that presumably contain the same host of zombie lineages as the first attempt because the authors claimed that the new divergence times were strongly correlated (0.9997) with the original divergence times (Liu et al., 2017b). A different reanalysis based on a revised suite of fossil calibrations supported the *Soft Explosive Model* (Phillips and Fruciano, 2018), further highlighting the sensitivity of timetrees to different node-based fossil calibration schemes.

Long Fuse Model

The *Long Fuse Model* posits that all or most interordinal cladogenesis occurred in the Cretaceous whereas the majority of intraordinal diversification took place after the KPg boundary (Archibald and Deutschman, 2001) (**Figure 1D**). Under this scenario, the initial diversification of placental mammals began in the Cretaceous, possibly in response to the Cretaceous Terrestrial Revolution and the associated diversification of flowering plants and insects (Meredith et al., 2011). Like the *Explosive Model*, the *Long Fuse Model* suggests an important role for the KPg boundary event, but restricts its impact to intraordinal splitting and ecological/phenotypic diversification, which exploded after the KPg mass extinction event in response to newly available niche

space (Meredith et al., 2011). This hypothesis is most strongly favored by analyses of molecular datasets comprising multiple gene fragments for small and large numbers of taxa (Kumar and Hedges, 1998; Eizirik et al., 2001; Murphy et al., 2001a; Murphy et al., 2001b; Springer et al., 2003; Springer et al., 2005; Murphy et al., 2007; Meredith et al., 2011; Lartillot and Delsuc, 2012; Emerling et al., 2015; Hedges et al., 2015; Foley et al., 2016; Springer et al., 2017) and genome wide data (Wildman et al., 2007; dos Reis et al., 2012; dos Reis et al., 2014; Tarver et al., 2016; Wu et al., 2017).

The *Long Fuse Model* predicts the occurrence of placental fossils deep in the Cretaceous. Possible eutherian forms are recognized in the fossil record as far back as the Jurassic with the discovery of *Juramaia sinensis* in China (Luo et al., 2011), although the phylogenetic placement of this taxon is contentious and some analyses have recovered *Juramaia* as a stem therian (e.g., Krause et al., 2014). Indeed, relationships among various eutherian forms that appear in the fossil record prior to the KPg boundary are controversial, with much debate centering over the correct assignment of extinct taxa to the stem of Placentalia or to the crown clade. This problem is exacerbated by the fragmentary skeletal remains that have been recovered for many of these taxa.

Fossils attributed to the Late Cretaceous families Zalambdalestidae and Zhelestidae were originally considered placentals (Archibald, 1996; Archibald et al., 2001). Specifically, cladistic analyses suggested that zalambdalestids represent a paraphyletic stem group to Glires (lagomorphs and rodents) whereas zhelestids form a clade with Ungulata (Archibald et al., 2001). Subsequent analyses with expanded taxon sampling have excluded zalambdalestids and zhelestids from crown Placentalia, instead recovering these fossils as stem placentals (Wible et al., 2009; Archibald and Averianov, 2012; O'Leary et al., 2013; Zhou et al., 2013). These contrasting results also highlight the importance of missing data. Another candidate crown placental from the Cretaceous is *Protungulatum coombsi*, which is known from at least 300,000 years before the KPg boundary in the Late Cretaceous Hell Creek Formation of Montana (Archibald et al., 2011). O'Leary et al. (2013) analyses of the morphological data set for mammals with the largest number (4541) of characters, as well as a combined analyses of this matrix with DNA data, reconstructed the position of *Protungulatum* as a crown laurasiatherian, thereby providing some paleontological support for a Cretaceous origin of Placentalia. However, like many fossil eutherians the position of *Protungulatum* is controversial. More recently, Halliday et al. (2017; 2019) recovered a stem placental position for *Protungulatum*. Another intriguing candidate for membership in Placentalia is *Gypsonictops*, which has now been reported from the Turonian (93.9–89.8 Ma) (Cohen and Cifelli, 2015; Cohen, 2017; Halliday et al., 2019). Halliday et al. (2019) recovered *Gypsonictops* (family Gypsonictopidae) and *Leptictis* (family Leptictidae) as sister taxa just outside of Placentalia. Numerous authors have also recognized an association of these families together in Leptictida (Gunnell et al., 2007; Wible et al., 2007; Wible et al., 2009). O'Leary et al. (2013) included *Leptictis* in their cladistic analysis of 4541 characters and recovered this taxon inside of Placentalia. However, Leptictidae is only known from the Cenozoic and its inclusion in Placentalia does not mandate a Cretaceous age for Placentalia. Still, taken together, the results of O'Leary et al. (2013) and Halliday et al. (2019) hint at the possible inclusion of Leptictida in crown Placentalia. More specifically, if Leptictidae and Gypsonictopidae are sister taxa, and if this clade is positioned in crown Placentalia rather than the stem group, then the main paleontological objection to the *Long Fuse Model* would be largely blunted.

Short Fuse Model

The *Short Fuse Model* posits interordinal and some intraordinal diversification of placental mammals well back in the Late (Upper) Cretaceous (Archibald and Deutschman, 2001) (**Figure 1E**). The initiation of interordinal cladogenesis may even extend as far back as the Upper Jurassic (Archibald and Deutschman, 2001). According to this model, the mass extinction event at the KPg boundary did not play a significant role in the interordinal diversification of present-day mammals nor the ecomorphological divergence of many ordinal level crown clades. Unlike the *Explosive* and *Long Fuse Models*, both of which are widely advocated in the literature, support for the *Short Fuse Model* is restricted to a relatively small number of studies. These include early molecular clock analyses (e.g., Kumar and Hedges, 1998), a supertree analysis (Bininda-Emonds et al., 2007), and more recently morphological clock studies (Puttick et al., 2016; Caldas and Schrago, 2019). The most explicit support for the *Short Fuse Model* comes from Bininda-Emonds et al. (2007), who used a matrix representation with parsimony approach to build a supertree representing ~99% of mammalian species-level diversity. However, the molecular dating analysis employed local molecular clocks and a pure birth model to interpolate some divergence times. Bininda-Emonds et al. (2007) concluded that the KPg extinction had no effect on the diversification of extant lineages, and instead suggested that increased diversification in the Eocene may have been triggered by the Early Eocene Climatic Optimum (Bininda-Emonds et al., 2007). The conclusion that extant lineages experienced accelerated rates of diversification in the Eocene was not supported by a subsequent study that employed relaxed clock methods (Meredith et al., 2011). Puttick et al. (2016) performed tip dating with the morphological data set (4,541 characters) of O'Leary et al. (2013) and recovered interordinal and intraordinal divergence times for the placental radiation that are even older than those of Bininda-Emonds et al. (2007) (see *Challenges for Tip Dating*).

By contrast with these timetree studies, Tavaré et al. (2002) and Wilkinson et al. (2011) used modeling approaches to address the question of whether or not divergence times within crown Primates (Euprimates) extend as far back as the KPg boundary. If intraordinal divergence times in Primates extend into the Mesozoic, then interordinal divergences for deeper nodes must be at least this old. These modeling approaches incorporated parameters for fossil preservation rates, the mean longevity of fossil primate species, and the number of extant primate species. Tavaré et al. (2002) concluded that crown Primates last shared a common ancestor ~81.5 Ma. Wilkinson et al. (2011) obtained posterior estimates of divergence times for several nodes within Primates based on their modeling approach and then used these estimates as priors in an MCMC analysis with DNA sequences.

Similar to Tavaré et al. (2002); Wilkinson et al. (2011) concluded that Primates last shared a common ancestor ~84.5 Ma. Thus, both of these studies are consistent with the predictions of the *Short Fuse Model*. However, Phillips (2016) criticized several assumptions of these models including logistic species accumulation and long times to speciation (2–3 myr), both of which favor a long period of missing history early in primate evolution.

NODE DATING AND BEYOND

Several problems are potentially of concern for node- tip-, and fossilized birth–death dating methods that can be applied to the placental radiation. Other shortcomings are restricted to a subset of these methods. In this section we first address common problems and then examine unique problems that are associated with specific methods.

Homology

An important issue for all molecular timetree methods is the underlying quality of the DNA or protein alignments. In the Sanger sequencing era, it was straightforward to inspect individual alignments for misaligned regions or problematic sequences from smaller sets of orthologous genes. Similarly, gene trees were routinely inspected for red flags such as unexpected relationships that may indicate contamination or paralogy. However, it is no longer practical to inspect/edit thousands of alignments that are tens or even hundreds of kilobases in length and contain hundreds of taxa. Nevertheless, this does not excuse researchers from assessing the quality of their alignments and gene trees. Indeed, numerous phylogenomic data sets (Struck et al., 2011; Chiari et al., 2012; Song et al., 2012; Kumar et al., 2013; Jarvis et al., 2014; Feijoo and Parada, 2017; Chen et al., 2017; Liu et al., 2017a) contain alignments with homology problems that impact the results and main conclusions of these studies (Struck, 2013; Springer and Gatesy, 2016; Brown and Thomson, 2017; Gatesy and Springer, 2017; Springer and Gatesy, 2018a; Springer and Gatesy, 2018b). These problems could have been avoided with appropriate screening procedures to flag problematic alignments and gene trees. **Figure 2** shows an example of yet another phylogenomic data set (Chen et al., 2017) with large-scale homology problems that impact the major conclusions of this study. Even without inspecting all of the constituent alignments and gene trees, it is possible to ascertain if there are systematic problems *via* targeted or even random sampling of the individual alignments and trees. One approach for targeted inspection is to view alignments that correspond to the gene trees with the highest Robinson-Foulds (RF) distances (Robinson and Foulds, 1981). RF Distances Filter (Simmons et al., 2016) is especially useful for this purpose and outputs normalized RF distances between gene trees (or between gene trees and a species tree) that range from 0 for identical trees to 1 for trees with no internal branches in common. Problematic sequence alignments as shown in **Figure 2** are not difficult to recognize, especially for a trained systematist who is acquainted with the taxonomy of their group. Reciprocal BLAST searches and re-alignments, sometimes in conjunction with new phylogenetic analyses, can be used to verify if problematic regions of an alignment correspond to orthologous regions of the same gene or not (Springer and Gatesy, 2018a; Springer and Gatesy, 2018b). Similarly, a targeted approach may be used to inspect all alignments with long branches that exceed a specified threshold (Mason et al., 2016). Springer and Gatesy (2018b) used both of these approaches (highest RF distances, long branches) to identify alignments with orthology problems for several phylogenomic data sets including Kumar et al.'s (2013) data set for Euarchontoglires and Jarvis et al.'s (2014) data set for birds. We agree with Bromham (2019, p. 3) that the "safest approach is to only analyze those alignments for which you are certain of homology for all columns and rows, resisting the temptation to analyze unverified alignments for the sake of expedience." Homology errors in alignments will be propagated in all subsequent steps (e.g., phylogeny reconstruction, estimation of divergence dates) and should be avoided. Ongoing efforts to develop new methods to screen genomic alignments (e.g., Ali et al., 2019) for such errors should reduce this source of error moving forward. It is also important for authors to make all gene alignments available so that *ad hoc* criteria used to exclude genes (or regions thereof) can be evaluated by other researchers.

Zombie and Ghost Lineages

An additional red flag for timetree analyses is the occurrence of zombie lineages, where estimated divergence times are younger than minimum ages implied by fossils (Springer and Gatesy, 2016; Springer and Gatesy, 2018c). Zombie lineages are evident in several recent studies that have addressed the timing of the placental radiation (Phillips, 2016; Sato et al., 2016; Liu et al., 2017a). The most extreme example is Liu et al. (2017a) where the estimated divergence date for sperm whale [a toothed whale (Odontoceti)] to minke whale [a baleen whale (Mysticeti)] is only 2.9 Ma. This estimated date is more than an order of magnitude younger than the age of the oldest mysticete fossil (*Mystacodon*, 36.4 Ma) (Gatesy and Springer, 2017; Lambert et al., 2017a; de Muizon et al., 2019) and is also younger than numerous extinct mysticete and physeteroid (sperm whale) genera (**Figure 3**). By contrast, McGowen et al.'s (2009) timetree for Cetacea accommodates all of these fossils without any zombie lineages (**Figure 3**). At the opposite end of the spectrum, Barido-Sottani et al.'s (2019) fossilized birth–death analysis of Cetacea resulted in excessively long ghost lineages when fossil ages were estimated from an uncertain age range using midpoints or randomly sampled from these same age ranges. Specifically, the most recent common ancestor of crown Cetacea was estimated at > 60 Ma with midpoint ages and > 50 Ma with random draws from uncertain age ranges. The former date is more than 23 million years older than the earliest known crown cetaceans (Lambert et al., 2017a; de Muizon et al., 2019) and ten million years older than *Ambulocetus* (= walking whale), which is an early transitional form (stem Cetacea) that retained short limbs and large feet for swimming (Thewissen et al., 1996; Madar et al., 2002). Zombie lineages and ghost lineages should both be carefully compared to the fossil record in timetree analyses. Long ghost lineages are sometimes required because of a poor fossil record, as

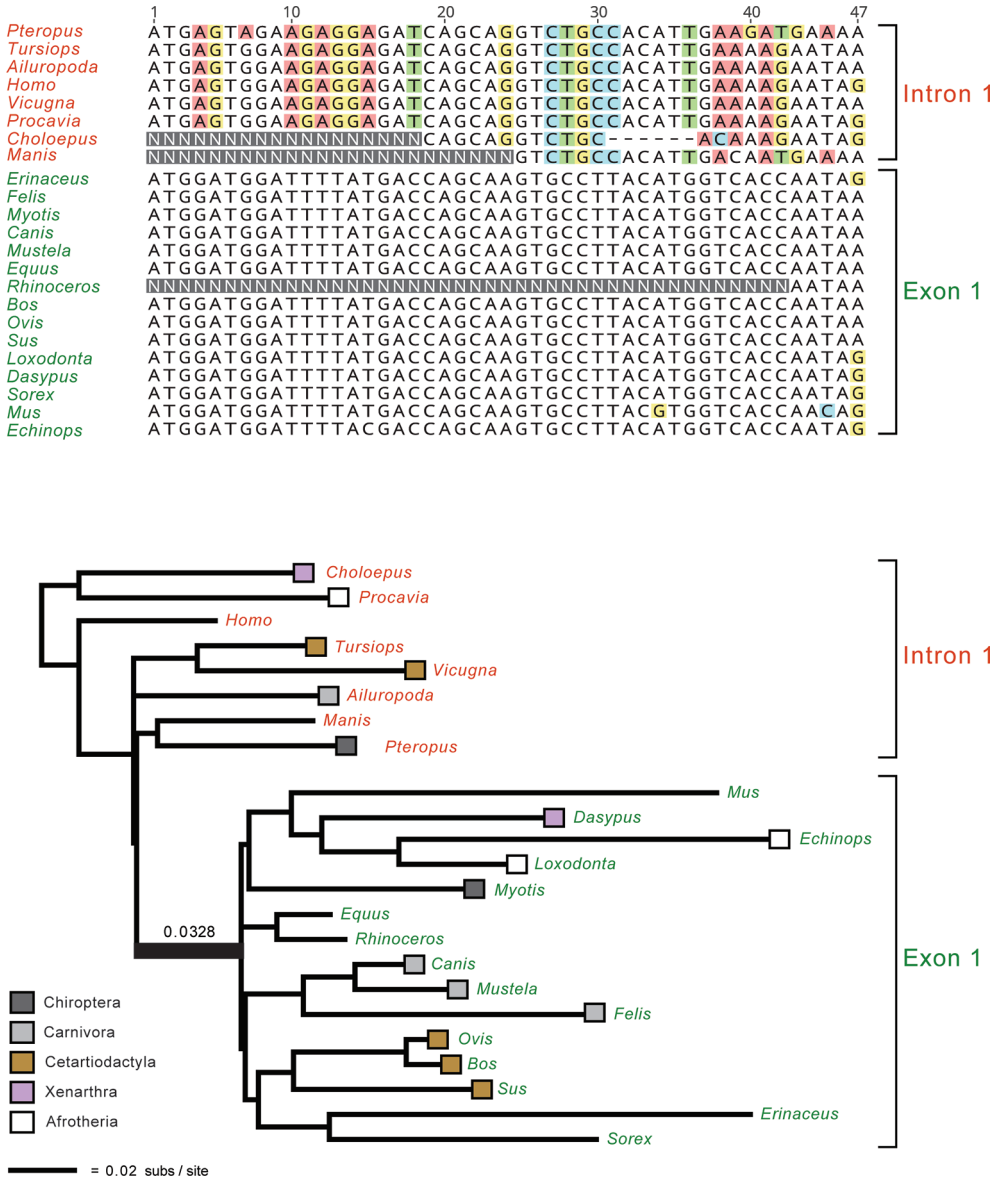


FIGURE 2 | Example of a homology problem from Chen et al.'s (2017) phylogenomic data set for Laurasiatheria and outgroups. Partial *ETV1* gene alignment (top) and gene tree for the full *ETV1* alignment are shown. Protein-coding sequences for 15 taxa (green lettering) are for exon 1 and begin on the start codon ATG, but the first eight taxa in the alignment (red lettering) instead are represented by sequence from intron 1 of *ETV1*. Faulty annotation and subsequent misalignment of protein-coding sequence to non-coding sequence results in 20 'pseudo-synapomorphies' for a clade that contradicts five well-established mammalian clades. The long internal branch that subtends this clade, 0.0328 substitutions per site, is indicated. Nucleotides that differ from the majority nucleotide at each position in the alignment are highlighted in colored boxes.

is the case for craseonycterid and myzopodid bats (Teeling et al., 2005), but the long ghost lineages for Cetacea implied by Barido-Sottani et al.'s (2019) analyses are less reasonable given the much more complete fossil record for cetaceans than for craseonycterid or myzopodid bats.

Phylogenetic Placement of Fossils

All timetree methods are critically dependent on the accurate phylogenetic placement of extinct taxa, whether through *a priori* decisions based on previous analyses and observations (node-dating, fossilized birth–death dating) or through the simultaneous estimation of phylogenetic relationships and divergence times (tip dating). This task is especially difficult for placental orders because of widespread ecomorphological convergence and correlated character evolution (Springer et al., 2007; Springer et al., 2013; Springer et al., 2017). For example, highly specialized myrmecophagy has evolved independently in Old World pangolins (Pholidota), African aardvarks (Tubulidentata), and New World anteaters (Xenarthra). Some

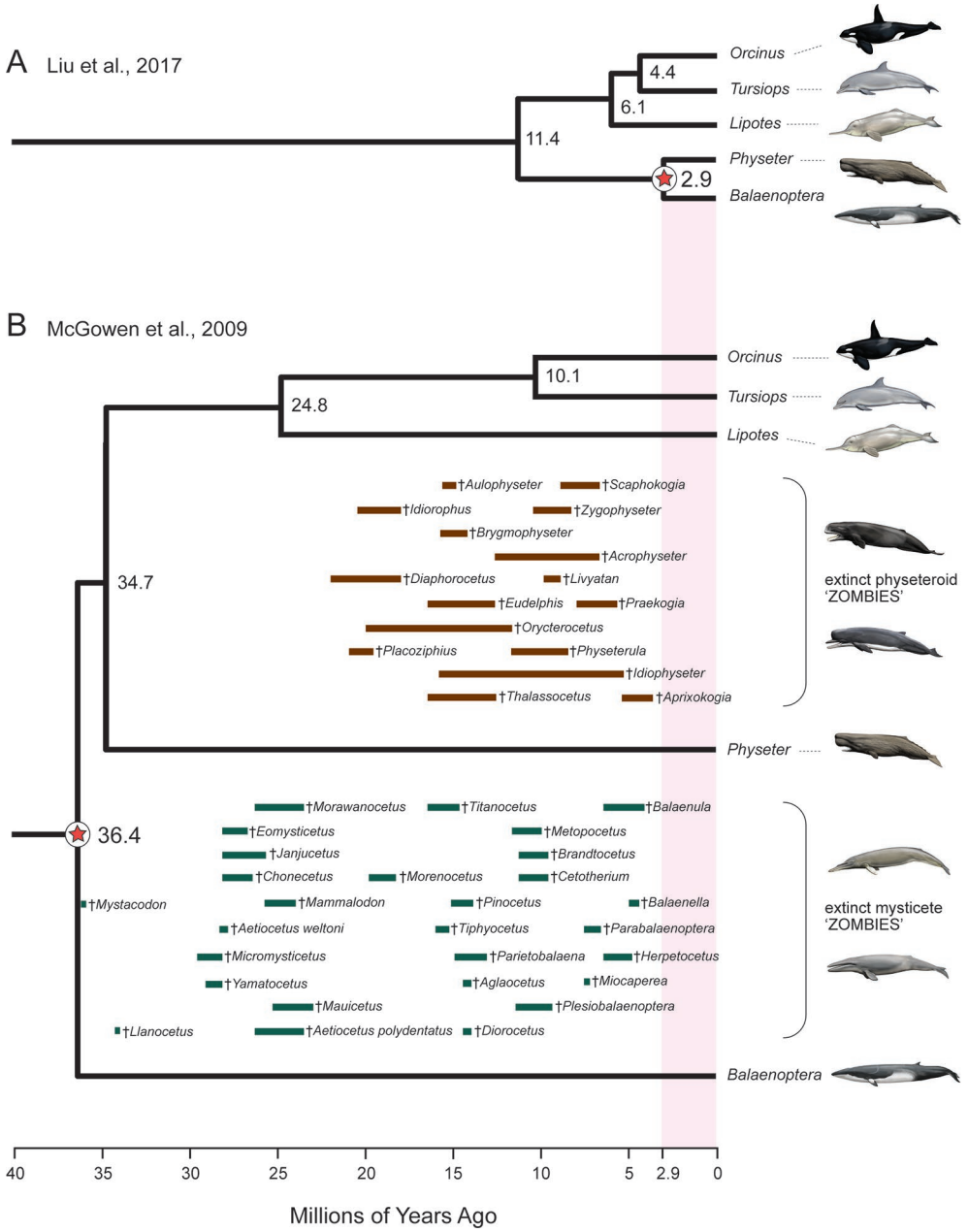


FIGURE 3 | Example of 'zombie' whale lineages implied by the timetree for mammals of Liu et al. (2017a). Due to inadequate density of fossil calibrations in this molecular clock study, the slowly evolving cetacean clade shows extremely shallow divergences **(A)** relative to previous molecular clock analyses such as McGowen et al. (2009) **(B)**. Numerous extinct sperm whales (Physeteroidea) and baleen whales (Mysticeti) are found in strata that are much older than the divergence time estimate between *Physeter* (giant sperm whale) and *Balaenoptera* (rorqual baleen whale) in **(A)** but not in **(B)**. Geological range estimates for extinct mysticetes (green bars) and physeteroids (brown bars) are from Marx and Fordyce (2015) and Lambert et al. (2017a; 2017b). Paintings are by C. Buell.

or all of these taxa routinely cluster together in morphological cladistic analyses (Novacek, 1986; O'Leary et al., 2013). Darwin (1859) was aware of the general problem of ecomorphological convergence and noted that adaptation to similar conditions will conceal, rather than reveal, genealogical relationships. Total evidence phylogenetic analyses that combine morphological and molecular data matrices together can mitigate this problem for extant taxa, but there is no guarantee that extinct taxa will be accurately placed based on morphological data alone, especially if extinct taxa are from orders (e.g., Creodonta, Mesonychia) that are only distantly related to living forms. One approach to assess the severity of this problem is through pseudoextinction analyses that render all representatives of a living order extinct by retaining osteological characters but recoding molecular and soft morphological characters as missing. The logic behind this approach is that only hard parts are typically fossilized in extinct

taxa (Springer et al., 2007). Springer et al. (2007) showed that the majority of placental orders moved to different phylogenetic positions when they were treated as pseudoextinct and also that some of these orders became polyphyletic. One caveat is that Springer et al. (2007) examined a relatively small osteological data set of 185 characters from Asher et al. (2003) and raised the possibility that larger morphological data sets would overcome the problems that beset smaller data sets if these problems were statistical in nature (e.g., see Sterli et al., 2013) and resulted from small sample size. O'Leary et al.'s (2013) massive morphological data set (4,541 phenomic characters) provided an opportunity to re-evaluate the effects of pseudoextinction without the potential problem of small sample size. **Figure 4** shows the results of a pseudoextinction analysis with maximum parsimony for the extant orders of placental mammals and marsupial outgroups. As was the case in Springer et al.'s (2007; 2008) pseudoextinction analyses, the majority of placental orders moved to a different interordinal location when pseudoextinct (i.e., treated as fossils and just coded for hard parts). In addition, three of these orders (Afrosoricida, Cetartiodactyla, Eulipotyphla) become para- or polyphyletic (**Figure 4**). Distantly related insectivores (Afrosoricida, Eulipotyphla) group with each other, and all three orders with highly specialized myrmecophages (Xenarthra, Pholidota, Tubulidentata) cluster with one of the other myrmecophagous orders when treated as pseudoextinct. These results suggest that an entirely extinct clade of myrmecophagous placental mammals might join with one of the other myrmecophagous groups even if the true phylogenetic position of this extinct group is elsewhere in the overall tree. For example, the phylogenetic position of *Eurotamandua*, an enigmatic myrmecophage from the middle Eocene of Europe, is likely to be conflated with other myrmecophages such as pangolins or anteaters even if myrmecophagy originated independently in this taxon. Indeed, previous assessments of the phylogenetic affinities of this taxon based on putative synapomorphies and cladistic analyses suggest that *Eurotamandua* is closely related to Vermilingua (anteaters) (Storch, 1981), to Pholidota (pangolins) (McKenna and Bell, 1997), to Palaenodonta (an extinct relative of Pholidota) (Rose, 1999), or to Xenarthra (Halliday et al., 2019). There are also cases of extinct taxa whose phylogenetic position shifts to a seemingly less accurate position when morphological data for all taxa (extinct and extant) are analyzed in combination with molecular data for extant taxa in a total evidence analyses. One example is the extinct taxon *Rodhocetus*, which belongs to the stem cetacean family Rodhocetidae. The position of this taxon based on morphology only is with other cetaceans (Gatesy et al., 2013; O'Leary et al., 2013). However, *Rodhocetus* is outside of a clade that contains other cetartiodactyls plus perissodactyls in O'Leary et al. (2013) total evidence analysis. This result shows that molecular data do not always improve the phylogenetic placement of extinct taxa, especially for incompletely preserved fossils. *Rodhocetus* is only scored for 386 of 4,541 characters in O'Leary et al.'s phenomic character matrix.

The inclusion of extinct and extant taxa in the same analysis has the potential to break up long branches and improve phylogenetic accuracy, but diachronous terminals (i.e., terminals of different ages) may also create problems for morphological cladistic analysis. Namely, diachronous terminals create opportunities for long-branch misplacement because root to tip distances are longer for extant taxa than for fossils (Wang et al., 2005; Springer et al., 2017). This problem mimics lineage-specific rate variation in analyses of molecular data with extant taxa. The phylogenetic placement of fossils can also be negatively impacted by the inevitable bias of the fossil record to preserve hard (biomineralized) morphological structures. This bias can systematically distort phylogeny. Specifically, Sansom and Wills (2013) showed that fossils are more likely to move stemward than crownward when they are only known for biomineralized characters. The causes of stemward slippage are not entirely clear, although Sansom and Wills (2013) suggest that fundamental taphonomic biases associated with the preservation of hard versus soft part characters cause fossils to be interpreted as erroneously primitive. The result of this "stemward slippage" is that divergence dates will be underestimated (Sansom and Wills, 2013). Finally, a recent study on morphological evolution in placental mammals concluded that it may be very difficult to distinguish early members of the major placental groups from stem eutherians on the basis of skeletal and dental characters because Cretaceous forms were not ecologically diverse and may appear very similar to each other (Halliday et al., 2019). In a similar vein, previous authors hypothesized that placentals from the Cretaceous were small and may have diversified phylogenetically before they diverged morphologically and acquired the diagnostic features of crown placental orders (Easteal, 1999; Madsen et al., 2001; Springer and Murphy, 2007). For these reasons, it is difficult to have confidence in the phylogenetic placement of fossils that are only distantly related to extant forms. In addition, these problems are more likely to impact deeper nodes because the placement of extinct taxa becomes more uncertain with increasing phylogenetic depth.

In spite of potential difficulties with convergent evolution and diachronous terminals, fossils remain fundamentally important for understanding the timing of the placental radiation. Similarly, fossils are critical for deciphering sequences of character evolution because they record unique combinations of morphological characters that are unknown in living mammals (Lee and Palci, 2015). On the other hand, the misplacement of these fossils in a phylogenetic analysis may distort the resulting estimates of both divergence times and ancestral character states. To the extent that we are confident in the phylogenetic placement of fossils we may also be more confident in both timetree analyses and ancestral character state transformations. Here, the placement of fossils may be more reliable when they belong to groups that also have extensive living representatives such as Tenrecidae (Asher and Hofreiter, 2006), Primates (Pattinson et al., 2015), and Rodentia (Asher et al., 2019). However, the placement of extinct taxa without living representatives remains more elusive. For example, the entirely extinct Plesiadapiformes are generally recognized as a paraphyletic taxon at the base of Euprimates (Silcox et al., 2017), but some cladistic analyses with more comprehensive taxon sampling place the earliest known plesiadapiform genus, *Purgatorius*, outside of Placentalia (Halliday et al., 2017; Halliday et al., 2019). Similarly, Gheerbrant et al. (2018)

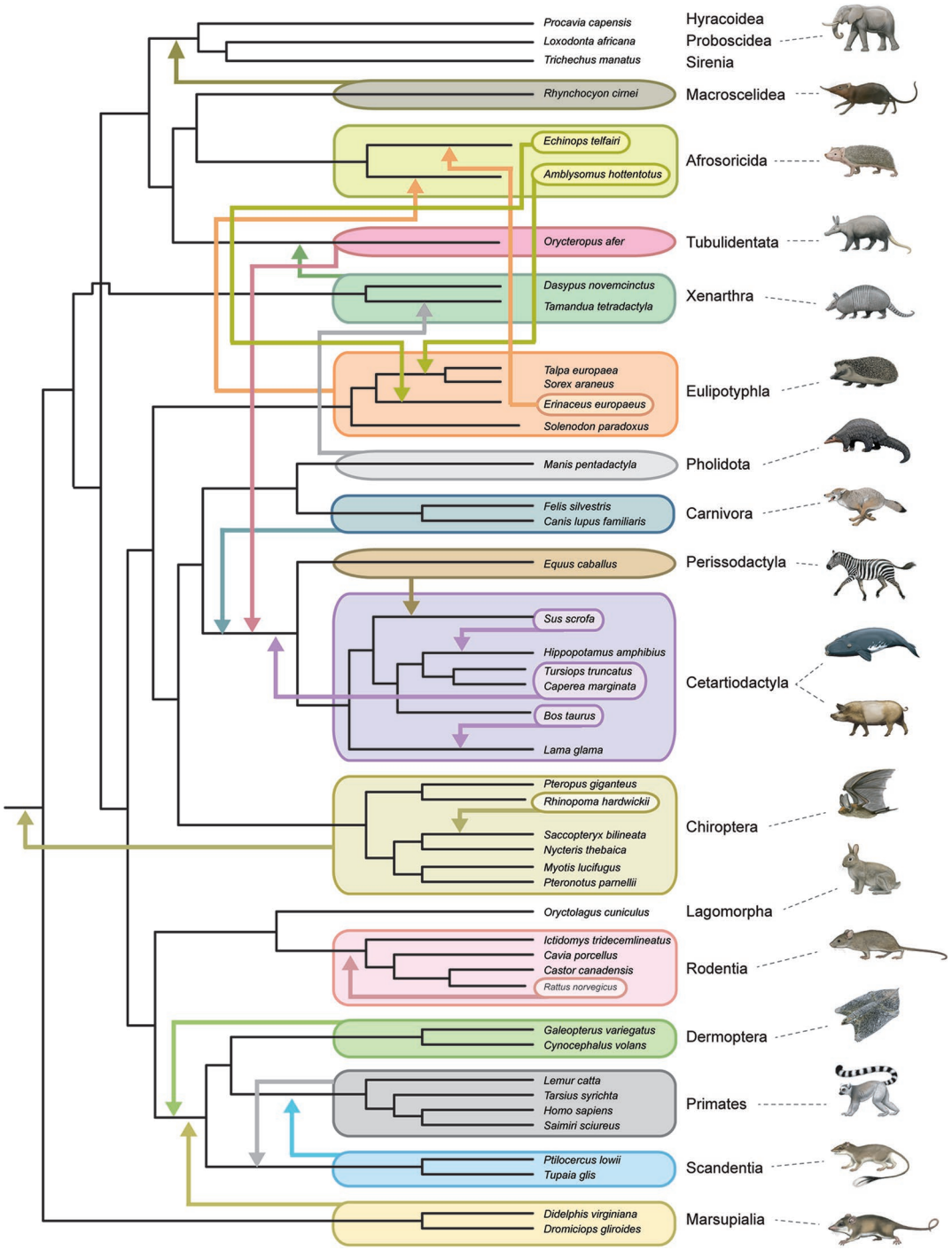


FIGURE 4 | Summary of pseudoextinction results for the reanalysis of morphological data from Morphobank Project 773 (O'Leary et al., 2013). Analyses were performed with a molecular scaffold that was based on robustly supported clades (>95% bootstrap support) from Meredith et al.'s (2011) phylogenetic analysis of 26 nuclear loci. The molecular scaffold included several polytomies that are not yet confidently resolved by molecular data: trichotomy at root of Placentalia (Afrotheria, Boreoeutheria, Xenarthra), paenungulate trichotomy (Hyracoidea, Proboscidea, Sirenia), Euarchontoglires trichotomy (Primatomorpha, Glires, Scandentia), and Laurasiatheria polytomy (Carnivora+Pholidota, Chiroptera, Cetartiodactyla, Perissodactyla). Pseudoextinct taxa were made pseudoextinct by recoding soft tissue characters as missing and deleting the pseudoextinct taxon from the molecular scaffold. Maximum parsimony analyses of 19 ordinal level taxa were individually executed with PAUP 4.0a165 (Swofford, 2002) and compared to the master scaffold. Parsimony analyses for each pseudoextinct clade were performed with 1000 random input orders of taxa and tree-bisection and reconnection branch swapping. Mammalian orders that showed shifts in phylogenetic position in these analyses are indicated by arrows that show the movements of entire clades as well as the repositioning of subtaxa within or among orders. Only four orders (Lagomorpha, Hyracoidea, Proboscidea, Sirenia) did not show changes to phylogenetic relationships in these analyses. Monotreme outgroups were included in the original analysis but were pruned from the tree shown here. Paintings are by C. Buell.

recovered the extinct order Embrithopoda as a clade of stem tethytheres, but other analyses have positioned this order elsewhere within Paenungulata or even deeper in Afrotheria (Tabuce et al., 2007; Cooper et al., 2014; Erdal et al., 2016). An even more difficult fossil group is Anagalida, which minimally includes the families Anagalidae and Pseudictopidae. Representatives of these families have been recovered as stem Glires, the sister taxon to Macroscelidea, or even as stem placentals in different phylogenetic analyses (Meng et al., 2003; Meng, 2004; Asher et al., 2019).

There are no easy solutions for elucidating ecomorphological convergence among extant and extinct placental mammals. One positive result for a longstanding phylogenetic problem concerns the phylogenetic placement of two recently extinct orders of South American ungulates, Notoungulata and Litopterna. Morphological studies have placed one or both of these orders in a variety of different locations on the placental tree. O'Leary et al. (2013) included a representative of each of these orders in their phylogenetic analysis of the mammalian radiation. They recovered a stem euungulate (Cetartiodactyla + Perissodactyla) position for *Protolipterna*, an early representative of the order Litopterna, and a nested position within Paenungulata (Proboscidea + Sirenia + Hyracoidea) for *Thomashuxleya*, a representative of the order Notoungulata. More recently, amino acid sequences for ancient collagen molecules from extinct members of these orders have been determined using mass spectrometry (Buckley, 2015; Welker et al., 2015). Phylogenetic analyses based on these sequences show that the representative litoptern (*Macrauchenia*) and notoungulate (*Toxodon*) are sister taxa to each other and that this monophyletic group is the sister taxon to Perissodactyla (Buckley, 2015; Welker et al., 2015). This clade was named Panperissodactyla (Welker et al., 2015) and was partially corroborated by a phylogenetic analysis of mitogenomic sequences by Westbury et al. (2017) that demonstrated a sister-group relationship between Litopterna (*Macrauchenia*) and Perissodactyla (Notoungulata not included in the analysis). Molecular sequences are not immune to homoplasy, as for example the lysozyme protein in foregut fermenting ruminants, colobus monkeys, and the hoatzin (Kornegay et al., 1994) and a handful of hearing proteins such as prestin in echolocating bats and toothed whales (Liu et al., 2010; Davies et al., 2012). However, convergent changes in these genes are limited to replacement substitutions and do not extend broadly across the genome to other loci. Liu et al. (2010) found that echolocating dolphins cluster with echolocating horseshoe and Old World leaf-nosed bats based on amino acid sequences for prestin, but analyses based on nucleotide alignments, which index both replacement and silent substitutions, recovered the accepted species tree and were not misled by convergence. Further, we are unaware of any phylogenomic analyses that group ruminants with colobus monkeys or echolocating bats with toothed whales. By contrast, there are several groups of ecomorphologically similar mammals (e.g., ant and termite eaters) that group together based on O'Leary et al. (2013) massive data set that includes morphological characters from many different parts of the body (Springer et al., 2013). Finally, given that Panperissodactyla is supported by independently segregating molecular markers (mitogenomes and collagen protein sequences), it seems unlikely that this relationship is driven by convergent evolution.

Challenges for Node Dating

Since 2003, node dating with a relaxed molecular clock has been the main approach used to estimate divergence times in different taxa including the timing of the placental radiation. Node dating is based on calibrating internal nodes against the fossil record (Ronquist et al., 2012). It is easy to apply with limited information from the fossil record, but like other methods (i.e., tip dating, fossilized birth-death dating) is not guaranteed to yield accurate divergence dates given some of the problems noted below. Node dating is implemented in several popular programs (e.g., mcmctree, BEAST). This approach does not require a morphological data matrix and can be implemented with both soft and/or hard-bounded calibrations. One potential problem with node dating is the use of unrelated priors (treewide prior, node-specific calibration) for each calibrated node (Heath et al., 2014). However, this problem can be avoided by applying a birth–death process to the uncalibrated nodes conditioned on the calibrated nodes (Yang and Rannala, 2006). A more serious problem is that probability densities for maximum age bounds are usually based on subjective or arbitrary criteria and are rarely informed by biological processes and/or detailed knowledge of the fossil record (Benton and Donoghue, 2007; Ho and Phillips, 2009; Heath et al., 2014; Arcila et al., 2015; Lee and Palci, 2015). The fossilization process is modeled only indirectly in node dating and in isolation from other forms of data (Heath et al., 2014). Models for branch-rate variation (e.g., lognormal, exponential) and its deployment (e.g., independent, autocorrelated) are drawn from statistical distributions that are convenient and tractable, but not necessarily reflective of real biological processes. This same criticism applies to tip dating methods (below). Meredith et al. (2011) showed that autocorrelated and independent models for the deployment of rate variation both perform poorly unless there is a dense network of calibrated nodes to combat (1) zombie lineages in large-bodied mammals with slow rates of evolution, and (2) excessively old divergences in small-bodied mammals with fast rates of evolution. Trends toward increased body size in extant mammalian orders may bias estimates of interordinal divergence times if calibrations are applied to large-bodied clades (Phillips, 2016), but this problem can be partially mitigated with hard-bounded constraints that enforce maximum ages (Meredith et al., 2011) and/or the exclusion of large-bodied taxa from timetree analyses of placental mammals (Springer et al., 2003; Springer et al., 2017).

Challenges for Tip Dating

In tip dating, morphological characters are coded for extinct and extant taxa and included in a combined data matrix that also includes molecular data for extant taxa (and in some cases recently extinct taxa). Tip dating employs a single probabilistic model that encompasses all of the different data types (fossil ages, molecular data matrix, and morphological data matrix) and then jointly estimates all of the model parameters, including a dated phylogeny, in a single analysis. However, current implementations

of tip dating have limitations. First, the phylogenetic placement of extinct taxa based on morphological data may be highly inaccurate because of correlated homoplasy, which occurs when multiple characters are correlated with each other and with the same environmental variables (Springer et al., 2007; Springer et al., 2017). Such correlations may be driven by adaptation to similar niches or by developmental constraints. The inclusion of molecular data can help to tease apart homology from homoplasy for extant taxa, but most fossils can only be scored for morphological data with their attendant problems of correlated character evolution. Second, the delineation of morphological characters and character states is intrinsically more subjective than is the case for molecular data, where there are just four nucleotides for DNA and 20 amino acids for proteins. Third, the notion of morphological clocks is problematic. Puttick et al. (2016) analyzed O'Leary et al. (2013) phenomic character matrix for extinct and extant mammals with a morphological clock model and obtained divergence time estimates for the most recent common ancestor of Placentalia that range from Late Jurassic (146.2 Ma) to Early Cretaceous (132.2 Ma) in age, much older than node-dating estimates based on molecular data sets that are generally in the range of 100–90 million years (Meredith et al., 2011; dos Reis et al., 2012; Emerling et al., 2015; Foley et al., 2016; Tarver et al., 2016; Springer et al., 2017). Similarly, Puttick et al. (2016) estimated dates for the most recent common ancestors of other superordinal groups that are consistently older than dates based on relaxed molecular clocks. Afrotheria (138.5–123.6 Ma), Euarchontoglires (139.1–125.1 Ma), and Laurasiatheria (142.6–128.3 Ma) all have dates that are tens of millions of years older than relaxed clock studies. Puttick et al.'s (2016) analyses also recovered Cretaceous dates for several crown orders including Cetartiodactyla (98.8–85.6 Ma), Chiroptera (88–80 Ma), and Eulipotyphla (106–91.2 Ma). Puttick et al. (2016) concluded that current implementations of tip dating analyses are prone to estimate ancient divergence estimates when based solely on morphological data. These authors recommended that the results of such analyses be treated with caution. Caldas and Schrago (2019) compared the results of molecular and morphological clocks with internal node calibrations and found that the majority of estimated ages were older with the morphological clock than the molecular clock. However, Caldas and Schrago (2019) estimated interordinal ages based on the morphological clock are younger than Puttick et al. (2016) estimated ages based on the morphological clock with tip dating. Taken together, the results of these studies (Puttick et al., 2016; Caldas and Schrago, 2019) suggest that morphological clocks and tip dating both contribute to older ages than are typically recovered with molecular clocks and node dating for placental mammals.

Models for morphological character evolution, such as the Mk model (Lewis, 2001), have been borrowed from molecular evolution as if morphological characters evolve under the same model as molecular characters. Molecular models may be tractable, but are unlikely to reflect realistic morphological character evolution. For example, most molecular models assume uniform branch rates, so that the probabilities of change for all characters, whether fast or slow, increase or decrease in concert with each other on each branch (Goloboff et al., 2018). As discussed by Goloboff et al. (2018), this assumption seems especially implausible for morphology. Finally, the collection of morphological data matrices is time consuming and expensive relative to the amount of data returned, and is not practical for most taxa on the scale of O'Leary et al.'s (2013) data set with > 4,500 phenomic characters for 86 mammaliaform taxa. Nevertheless, the development of these data matrices is crucial for various aspects of timetree estimation, either indirectly for node dating approaches or directly for tip dating approaches.

Challenges for Fossilized Birth–Death Dating

The fossilized birth–death model serves as a single prior probability distribution for divergence time dating that is used to calibrate and estimate node ages. Arbitrary calibration densities are not required as is the case for node dating. Indeed, the only assumptions are: (i) constant speciation rate, (ii) constant extinction rate, (iii) fossils are recovered along branches of the species tree according to a Poisson process, and (iv) each extant species is sampled with probability p. The original implementation of fossilized birth–death dating is the DPPDiv program of Heath et al. (2014). One limitation of this version of fossilized birth–death dating is that it does not allow for the inclusion of morphological characters in the analysis and only considers the age of each fossil. DPPDiv therefore requires the assignment of fossils to specific calibration nodes in the phylogeny based on prior information as is also true for node dating. More recently, fossilized birth–death dating has been combined with tip dating in BEAST2 (Gavryushkina et al., 2017), but this requires molecular and/or morphological data matrices for fossil and extant taxa and is not currently an option for most mammalian clades. In addition, fossilized birth–death dating assumes constant speciation and extinction rates and may be ill suited to investigate the timing of the placental radiation that spans the KPg mass extinction in four of five evolutionary models (**Figure 1**). Constant speciation and extinction are also unlikely to hold across diverse taxa with widely varying life histories. For example, speciation rates in the order Rodentia (> 2,000 extant species) have historically been much higher than in the order Tubulidentata (one extant species). We expect that future versions of fossilized birth–death dating may allow for different speciation and extinction rates in different sectors of a phylogenetic tree. Third, it is unlikely that fossils are recovered along branches according to a Poisson process. Rather, fossil recovery rates are spatially and temporally non-uniform and vary across different continents, time periods, taxonomic groups, and depositional environments (Holland, 2016). An additional issue for fossilized birth–death dating pertains to the sampling of fossils with imprecise ages that are represented as age ranges in the literature or in the Paleobiology Database.

Heath et al. (2014) original description of the fossilized birth–death method for timetree estimation provided an illustration of their approach with an empirical data set for Ursidae (bears). In this example, Heath et al. (2014) employed a molecular data set that included complete mitogenomes and a single nuclear gene, and randomly sampled each extinct ursid and fossil outgroup

with an imprecise age from a uniform distribution of its given range. This procedure resulted in divergence dates for Ursidae that are similar to the mitogenomic node-based estimates of Krause et al. (2008). These results suggest that fossilized birth–death dating can recover dates that are generally in line with dates that are estimated with other accepted methods. At the same time, Heath et al. (2014) suggested that it would be preferable to treat fossil ages as random variables by placing prior densities on fossil occurrence times conditional on their estimated age ranges. More recently, Barido-Sottani et al. (2019) found that fixing fossil ages to the midpoint or a random point drawn from within the stratigraphic age range resulted in biased estimates of divergence times. Specifically, estimated ages for Cetacea were much older than other studies (e.g., Steeman et al., 2009; McGowen et al., 2009; Marx and Fordyce, 2015) and imply huge gaps in the fossil record. By contrast, continual MCMC sampling of fossil ages from a prior distribution resulted in divergence time estimates that are in much better agreement with previous studies.

Timetree Analyses With ILS and Hybridization

Timetree analyses are aimed at estimating speciation times (incipient cladogenesis *sensu* de Queiroz, 1998), but timetree estimation is complicated by the observation that individual chromosomes and chromosomal segments may have different evolutionary histories. These differences can be the result of several processes, including coalescence (with or without ILS), recent and ancient gene flow, homoplasy, demography, natural selection, and sex-specific biases in gene flow. Each of these processes may differentially shape genetic variation within distinct, and non-random regions of a genome (Li et al., 2016; Li et al., 2019). Phylogenetic analysis of different genomic segments that are influenced by these different processes will yield trees with unique branching patterns and branch lengths, which when converted into time can produce a range of divergence estimates (Nachman and Payseur, 2012; Leaché et al., 2014; Fontaine et al., 2015). For example, coalescence will result in dates that are older than speciation times (**Figure 5**). In the absence of gene flow, coalescence times are expected to be older than speciation times even without ILS (**Figure 5A**). Under complete neutrality, the expected coalescence time for sequences that are sampled from individuals belonging to two different, completely isolated species is T + 2N, where T is the species divergence time and N is the population size of the ancestral species (Angelis and dos Reis, 2015). From this equation it is clear that expected coalescence times become increasingly older than actual speciation times with increasing ancestral population size. In some cases coalescence will not occur in the immediate common ancestor of two sister species and will occur even deeper in the tree, in which case it is referred to as deep coalescence (**Figures 5B, C**). In contrast, recent hybridization between sister taxa or the ancestors of extant sister taxa will result in divergence estimates between the hybridizing taxa that are younger than the initial time since divergence between the two lineages (**Figure 6A**). Hybridization between closely related non-sister taxa, including lineages that subsequently went extinct, can result in divergence time estimates between non-hybridizing taxa that are older than speciation times for some clades and younger for others (**Figures 6B–E**) (e.g., Figueiró et al., 2017; Barlow et al., 2018). Importantly, the direction of introgression has a critical role in altering divergence times for different clades (**Figure 6B** versus **Figure 6C**). Finally, hybridization may result in a new species that coexists with its parental lineages (**Figure 6F**). Hybrid speciation is rare in mammals, but hybrid origins have been suggested for the Neotropical bat *Artibeus schwartzi* (Larsen et al., 2010) and the Clymene dolphin (*Stenella clymene*) (Amaral et al., 2014).

Disregarding this potential variation in "gene" history across the genome can distort phylogenetic branch lengths and divergence estimates when multiple unlinked genomic loci are concatenated

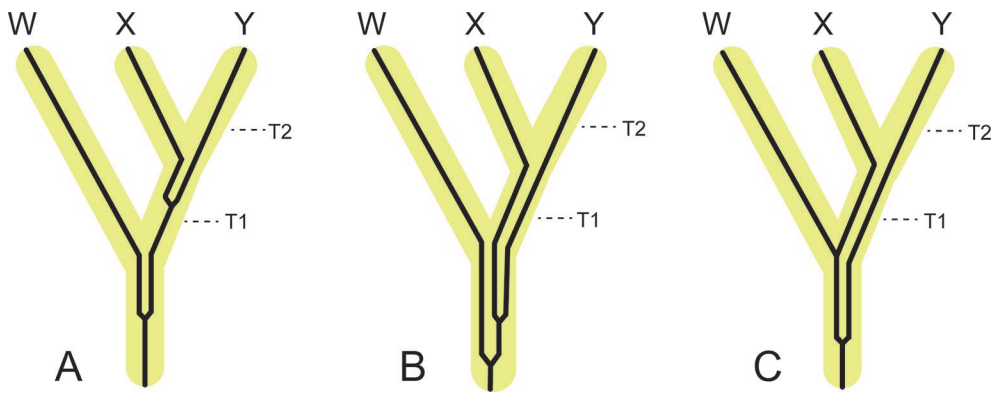


FIGURE 5 | Examples of the effects of coalescence for individual genes on divergence time estimation relative to speciation times (i.e., incipient cladogenesis with cessation of gene flow), T1 and T2. Gene tree lineages are thin and black and are contained within thick and yellow species tree lineages. **(A)** Coalescence of a gene in the most recent common ancestor of X and Y and in the most recent common ancestor of X+Y and W. The topology of the species tree and the topology of the gene tree are congruent. Coalescence times for this gene exceed species divergence times, but by less than one internal branch. **(B)** Deep coalescence of a gene in the common ancestor of W, X, and Y in which the gene tree topology agrees with the species tree topology. **(C)** Deep coalescence of a gene in the common ancestor of W, X, and Y in which the gene tree topology conflicts with the species tree topology. All coalescences of genes in this figure are deeper than speciation times, so molecular clock estimates from these gene trees would be older than the true speciation times (T1 and T2).

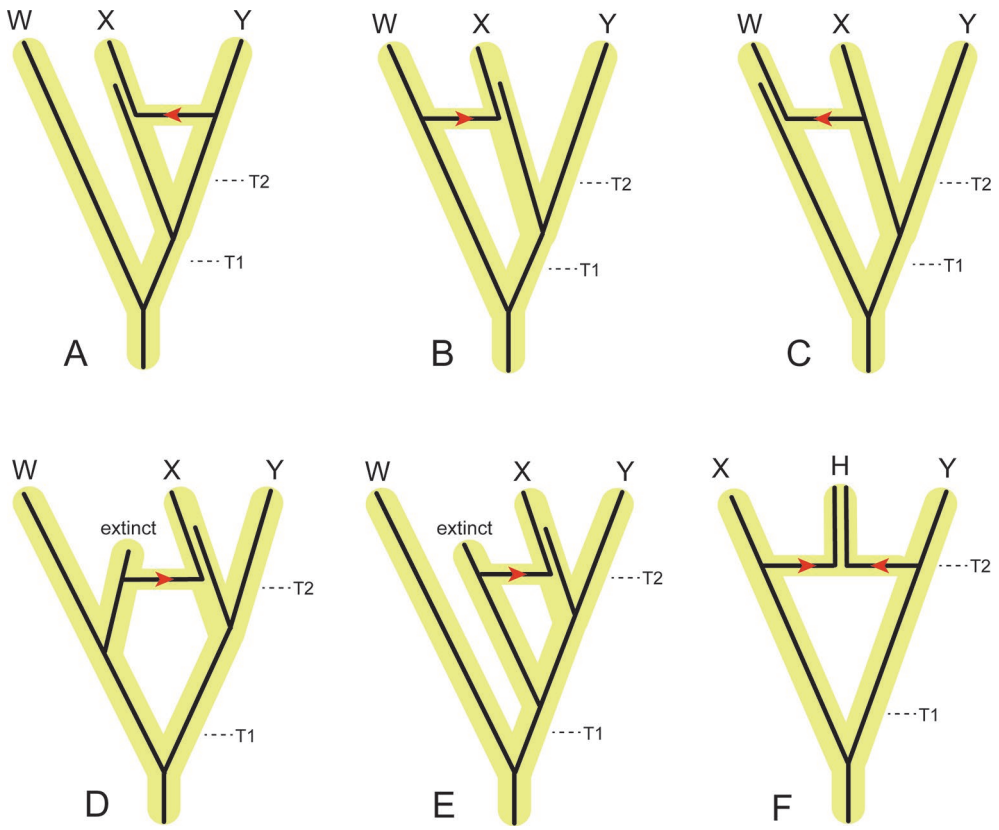


FIGURE 6 | Examples of the effects of introgression/hybridization on divergence time estimation relative to speciation (incipient cladogenesis) times for individual gene segments that each have a single evolutionary history. **(A)** Introgression of a gene from the ancestor of Y to the ancestor of X. This gene flow pathway will decrease the estimated divergence time between X and Y relative to the actual speciation time (T2), but have no effect on the estimated divergence time between W and X+Y. **(B)** Introgression of a gene from the ancestor of W to the ancestor of X. This gene flow pathway will increase the amount of divergence between X and Y relative to the speciation time T2, and decrease the divergence between W and X+Y relative to the speciation time T1. If this gene flow pattern is pervasive through the genome, then the democratic vote (i.e., count of different genes supporting each topology) of traditional concatenation and coalescence methods will flip the topology to [(W,X),Y]. **(C)** Introgression of a gene from the ancestor of X to the ancestor of W. This gene flow pathway will have no effect on the estimated divergence between X and Y relative to the speciation time T2, but will decrease the estimated divergence between W and X+Y relative to the speciation time T1. If this gene flow pattern is pervasive through the genome, then the democratic vote will flip the topology to [(W,X),Y]. **(D)** Introgression of a gene from an extinct relative of W to the ancestor of X. Introgressed genes of this type will increase the estimated divergence between X and Y relative to the speciation time T2, and decrease the estimated divergence between W and X+Y relative to the speciation time T1. If this gene flow pattern is pervasive through the genome, then the democratic vote will flip the topology to [(W,X),Y]. **(E)** Introgression of a gene from an extinct relative of X+Y to the ancestor of X. Introgressed genes of this type will increase the estimated divergence between X and Y relative to the speciation time T2, but have no effect on the estimated divergence between W and X + Y relative to the speciation time T1. **(F)** Hybridization between the ancestors of X and Y results in a new species, (H), that coexists with the parental lineages that terminate in species X and Y. If the majority of the genome of H is derived from the ancestral lineage to X, then the democratic vote across the genome will favor the topology ((H,X),Y). Conversely, if the majority of the genome of H is derived from the ancestral lineage to Y, then the democratic vote across the genome will favor the topology [(H,Y),X].

or combined in a coalescence analysis (Schierup and Hein, 2000; Posada and Crandall, 2002; Lemey and Posada, 2009; Li et al., 2019). Timetree methods such as *BEAST (Heled and Drummond, 2010) allow individual loci to have unique histories under the multispecies coalescent with ILS, but more complicated models that deal with ILS and introgression are still in their infancy, especially in cases where hybridization effectively overwhelms the phylogenetic signal of speciation across the majority of the genome.

Hybridization

ILS has been widely recognized as a source of gene tree variation, and coalescent methods to accommodate this variation (together with other sources of variation) are becoming more widespread (e.g., Hobolth et al., 2007; Degnan and Rosenberg, 2009; Hobolth et al., 2011; Hibbins and Hahn, 2019; Bravo et al., 2019). However, most coalescence approaches assume all discordance among loci results from ILS (Liu et al., 2009), thus disregarding the potential contributions from post-speciation gene flow to phylogenomic discordance. The expansion of whole genome data has led to the recognition that interspecific hybridization is a widespread phenomenon across the tree of life and must be accounted for in phylogenomics and timetree estimation (e.g., Ai et al., 2015; Fontaine et al., 2015; Lamichhaney et al., 2015; Li et al., 2016; Martin and Jiggins, 2017; Árnason et al., 2018; Barlow et al., 2018; Palkopoulou et al., 2018; Li et al., 2019). A few cases in

the literature (mosquitoes, butterflies, and cats) have shown that extensive hybridization can effectively replace the phylogenetic signal of original branching events across the majority of the genome, and in these instances the original signal for the deepest divergence point between taxa is only present in a minority of the genome (Fontaine et al., 2015; Edelman et al., 2019; Li et al., 2019). In such situations, traditional concatenation and coalescence approaches that use all of the data ("democratic vote" methods, Degnan and Rosenberg 2009) may fail to construct an accurate phylogeny of the original branching events, instead producing trees that reflect the most recent and extensive bouts of gene flow. These studies illustrate the importance of teasing apart segments of the genome that have different histories. We recommend that researchers examine X and Z chromosomal regions, especially low-recombination regions of these chromosomes, to determine if they record different histories than other regions of the genome. An additional point is that hybridization, if not accounted for, has the potential to result in zombie lineages where estimated divergence times are younger than minimum ages for speciation that are implied by the fossil record.

Recombination

A second important observation from recent phylogenomic studies is that in lineages with an extensive history of hybridization and introgression, signatures of ancient (original) branching events are more rapidly depleted from chromosomal regions with high rates of meiotic recombination and more localized effects of linked selection (Brandvain et al., 2014; Schumer et al., 2018; Li et al., 2019; Martin et al., 2019). Conversely, regions of low recombination, particularly the X and Z sex chromosomes, are enriched for genomic segments that support the original species tree prior to reticulation (Fontaine et al., 2015; Edelman et al., 2019; Li et al., 2019; Martin et al., 2019). This enrichment on the sex chromosomes may be due to a higher density of reproductive isolating loci, reduced effective population size and hence reduced ILS, or some combination of these processes (Pease and Hahn, 2013; Presgraves, 2018). Perhaps paradoxically, Wang and Hahn (2018) showed that speciation genes, if they participate in Dobzhansky-Muller incompatibilities with other loci *via* epistatic interactions, are more likely to have gene trees that are discordant with the species tree.

Li et al. (2019) demonstrated a strong topological bias in high-recombination regions that are enriched for signatures of gene flow, supporting observations from previous simulation studies (Posada, 2000; Schierup and Hein, 2000; Leaché et al., 2014). The degree of branch length (and timetree) distortion is dependent on the temporal context and intensity of gene flow throughout the evolutionary history of the group (Li et al., 2019). Li et al. (2019) concluded that some star-like phylogenies could be artifacts of combining sequences derived from regions of the genome with elevated recombination rates and histories of gene flow, rather than accurate depictions of the diversification process. Collectively, these studies indicate that recombination rate is an important parameter to consider in phylogenetic inference and divergence time estimation (Edelman et al., 2019; Li et al., 2019; Martin et al., 2019). At the same time, recombination rate is a difficult parameter to include in most studies due to the rarity of recombination maps for most taxa. Yet this is likely to change in the near future as new linkage disequilibrium-based approaches allow for estimation of genome-wide recombination rates in a broader array of non-model organisms from population genomic sequence data (Stevison et al., 2016). Future studies should investigate the influence of local recombination rates and properly parsed out coalescence genes (e.g., Hobolth et al., 2007) on tree shape and divergence time estimation. One area of interest is to determine whether any of the previously supported models for mammalian evolution based on molecular studies are biased because of the distorting effects of combining loci from regions of the genome with highly variable or elevated recombination rates.

CONCLUSIONS

The reconstruction of a reliable timetree for placental mammals is fundamentally important for understanding the potential role of the KPg extinction and other events in Earth history in promoting mammalian diversification. However, an agreed upon timetree remains elusive. Indeed, the number of models for placental diversification has increased, rather than decreased, over the last two decades. The list of competing models now includes the *Explosive*, *Soft Explosive*, *Trans-KPg*, *Long Fuse*, and *Short Fuse Models*. Unprecedentedly large phylogenomic and multigene data sets for placental mammals have become available in the last decade (Meredith et al., 2011; dos Reis et al., 2012; Song et al., 2012; dos Reis et al., 2014; Emerling et al., 2015; Foley et al., 2016; Tarver et al., 2016; Chen et al., 2017; Liu et al., 2017a; Wu et al., 2017). Similarly, O'Leary et al. (2013) and Halliday et al. (2019) have assembled the largest morphological data sets for Eutheria in this same time span. Molecular and morphological data can now be analyzed, either separately or in combination with each other, with increasingly complex approaches to timetree reconstruction (Kumar, 2005; Pyron, 2011; Ronquist et al., 2012; Heath et al., 2014; Kumar and Hedges, 2016) that have the potential to discriminate among competing models for placental diversification. However, some of the largest phylogenomic data sets have pervasive homology problems, often due to limitations of fragmented draft genome assemblies and their gene annotations, that limit their usefulness for phylogeny reconstruction and timetree estimation (Springer and Gatesy, 2016; Gatesy and Springer, 2017; Springer and Gatesy, 2018a; Springer and Gatesy, 2018b). Thus improving the contiguity and annotation of genome assemblies across the mammalian tree will reduce the probability of these artifacts. Similarly, new methods for timetree estimation have potential shortcomings that must be addressed if we are to reconstruct an accurate timetree for placental mammals. For example, tip dating methods employ morphological clock models that are conveniently borrowed from molecular evolutionary genetics, but these models may not be appropriate for morphological data. On the paleontological front, new fossil discoveries have the potential to provide decisive evidence for or against some of the models for placental diversification, but this requires that the fossils can be unambiguously placed in the eutherian tree. Variation in "gene" histories that results from the coalescent process (including

ILS) and hybridization can distort phylogenetic branch lengths and divergence estimates when multiple unlinked genomic loci are combined together in a timetree analysis. The acquisition of high quality genomes for more and more mammalian taxa, combined with methods for detecting recombination and introgression, should help to facilitate the identification of genomic regions with different histories. The partitioning of the genome into these regions with different histories will be an important step in estimating species divergence times in the radiation of placental mammals. Finally, the acquisition of large-scale genomic data sets provides an opportunity for culling loci that exhibit a poor fit to models of sequence evolution. For example, an important conclusion from Liu et al. (2017a) is that suboptimal molecular clock loci and methods are a major cause of discordance among different studies that have investigated the timing of the placental radiation. A caveat here is that Liu et al.'s (2017a) results are also tainted by extensive homology problems and zombie lineages (Gatesy and Springer, 2017). Nevertheless, the important point here is that different models that are employed in timetree estimation, whether they be models of sequence evolution or models of rate variation across branches of a phylogenetic tree, should be adequate to describe the relevant process instead of just better fitting than other models. Timetree estimation is highly interdisciplinary, and we remain optimistic that improved estimates of the timing of the placental radiation will result from new fossil discoveries, additional high quality genomes, and improved models and methods for the analysis of these data.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. These data can be found here: Project DOI: 10.7934/P773, http://dx.doi.org/10.7934/P773.

AUTHOR CONTRIBUTIONS

MS, NF, JG, and WM conceived the study. PB performed pseudoextinction analyses. JG and MS collected data from genomic and fossil databases. MS, NF, JG, and WM wrote the manuscript. PB, NF, and JG constructed figures. PB provided comments on the draft manuscript. All authors read and approved the final draft for submission.

FUNDING

This work was supported by National Science Foundation grant DEB-1457735 (JG and MS) and DEB-1753760 (WM).

ACKNOWLEDGMENTS

We thank Michel Laurin and four referees for helpful comments on an earlier draft of this manuscript. Animal paintings are by Carl Buell.

REFERENCES

Árnason, U., Lammers, F., Kumar, V., Nilsson, M. A., and Janke, A. (2018). Whole-genome sequencing of the blue whale and other rorquals finds signatures for introgressive gene flow. *Sci. Adv.* 4, eaap9873. doi: 10.1126/sciadv.aap9873

Ai, H., Fang, X., Yang, B., Huang, Z., Chen, H., Mao, L., et al. (2015). Adaptation and possible ancient interspecies introgression in pigs identified by whole-genome sequencing. *Nat. Genet.* 47, 217–225. doi: 10.1038/ng.3199

Ali, R. H., Bogusz, M., and Whelan, S. (2019). Identifying clusters of high confidence homologies in multiple sequence alignments. *Mol. Biol. Evol.* 36, 2340–2351. doi: 10.1093/molbev/msz142

Alroy, J. (1999). The fossil record of North American mammals: evidence for a Paleocene evolutionary radiation. *Syst. Biol.* 48, 107–118. doi: 10.1080/106351599260472

Amaral, A. R., Lovewell, G., Coelho, M. M., Amato, G., and Rosenbaum, H. C. (2014). Hybrid speciation in a mammal mammal: the clymene dolphin (*Stenella clymene*). *PLoS One* 9, e83645. doi: 10.1371/journal.pone.0083645

Angelis, K., and dos Reis, M. (2015). The impact of ancestral population size and incomplete lineage sorting on Bayesian estimation of species divergence times. *Curr. Zool.* 61, 874–885. doi: 10.1093/czoolo/61.5.874

Archibald, J. D., and Averianov, A. (2012). Phylogenetic analysis, taxonomic revision, and dental ontogeny of the Cretaceous Zhelestidae (Mammalia: Eutheria). *Zool. J. Linn. Soc.* 164, 361–426. doi: 10.1111/j.1096-3642.2011.00771.x

Archibald, J. D., and Deutschman, D. H. (2001). Quantitative analysis of the timing of the origin and diversification of extant placental orders. *J. Mamm. Evol.* 8, 107–124. doi: 10.1023/A:1011317930838

Archibald, J. D., Averianov, A. O., and Ekdale, E. G. (2001). Late Cretaceous relatives of rabbits, rodents, and other extant eutherian mammals. *Nature* 414, 62–65. doi: 10.1038/35102048

Archibald, J. D., Zhang, Y., Harper, T., and Cifelli, R. L. (2011). *Protungulatum*, confirmed Cretaceous occurrence of an otherwise Paleocene eutherian (placental?) mammal. *J. Mamm. Evol.* 18, 153–161. doi: 10.1007/s10914-011-9162-1

Archibald, J. D. (1996). Fossil evidence for a late Cretaceous origin of "hoofed" mammals. *Science* 272, 1150–1153. doi: 10.1126/science.272.5265.1150

Arcila, D., Pyron, R. A., Tyler, J. C., Ortí, G., and Betancur-R, R. (2015). An evaluation of fossil tip-dating versus node-age calibrations in tetraodontiform fishes (Teleostei: Percomorphaceae). *Mol. Phylogenet. Evol.* 82, 131–145. doi: 10.1016/j.ympev.2014.10.011

Arnason, U., Adegoke, J. A., Bodin, K., Born, E. W., Esa, Y. B., Gullberg, A., et al. (2002). Mammalian mitogenomic relationships and the root of the eutherian tree. *Proc. Natl. Acad. Sci. U.S.A.* 99, 8151–8156. doi: 10.1073/pnas.102164299

Asher, R. J., and Hofreiter, M. (2006). Tenrec phylogeny and the noninvasive extraction of nuclear DNA. *Syst. Biol.* 55, 181–194. doi: 10.1080/10635150500433649

Asher, R. J., Novacek, M. J., and Geisler, J. H. (2003). Relationships of endemic African mammals and their fossil relatives based on morphological and molecular evidence. *J. Mamm. Evol.* 10, 131–194. doi: 10.1023/A:1025504124129

Asher, R. J., Smith, M. R., Rankin, A., and Emry, R. J. (2019). Congruence, fossils and the evolutionary tree of rodents and lagomorphs. *R. Soc. Open Sci.* 6, 190387. doi: 10.1098/rsos.190387

Barido-Sottani, J., Aguirre-Fernández, G., Hopkins, M. J., Stadler, T., and Warnock, R. (2019). Ignoring stratigraphic age uncertainty leads to erroneous estimates of species divergence times under the fossilized birth – death process. *Proc. R. Soc. B* 286, 20190685. doi: 10.1098/rspb.2019.0685

Barlow, A., Cahill, J. A., Hartmann, S., Theunert, C., Xenikoudakis, G., Fortes, G. G., et al. (2018). Partial genomic survival of cave bears in living brown bears. *Nat. Ecol. Evol.* 2, 1563–1570. doi: 10.1038/s41559-018-0654-8

Benton, M. J., and Donoghue, P. C. (2007). Paleontological evidence to date the tree of life. *Mol. Biol. Evol.* 24, 26–53. doi: 10.1093/molbev/msl150

Benton, M. J., Wills, M., and Hitchin, R. (2000). Quality of the fossil record through time. *Nature* 403, 534–537. doi: 10.1038/35000558

Bininda-Emonds, O. R., Cardillo, M., Jones, K. E., MacPhee, R. D., Beck, R. M., Grenyer, R., et al. (2007). The delayed rise of present-day mammals. *Nature* 446, 507–512. doi: 10.1038/nature05634

Brandvain, Y., Kenney, A. M., Flagel, L., Coop, G., and Sweigart, A. L. (2014). Speciation and introgression between *Mimulus nasutus* and *Mimulus guttatus*. *PLoS Genet.* 10, e1004410. doi: 10.1371/journal.pgen.1004410

Bravo, G. A., Antonelli, A., Bacon, C. D., Bartoszek, K., Blom, M. P. K., Huynh, S., et al. (2019). Embracing heterogeneity: coalescing the Tree of Life and the future of phylogenomics. *PeerJ* 7, e6399. doi: 10.7717/peerj.6399

Bromham, L. (2011). The genome as a life-history character: Why rate of molecular evolution varies between mammal species. *Phil. Trans. R. Soc. Lond. B* 366, 2503–2513. doi: 10.1098/rstb.2011.0014

Bromham, L. (2019). Six impossible things before breakfast: assumptions, models, and belief in molecular dating. *Trends Ecol. Evol.* 34, 474–486. doi: 10.1016/j.tree.2019.01.017

Brown, J. M., and Thomson, R. C. (2017). Bayes factors unmask highly variable information content, bias, and extreme influence in phylogenomic analyses. *Syst. Biol.* 66, 517–530. doi: 10.1093/sysbio/syw101

Buckley, M. (2015). Ancient collagen reveals evolutionary history of the endemic South American 'ungulates'. *Proc. R. Soc. B* 282, 20142671. doi: 10.1098/rspb.2014.2671

Caldas, I. V., and Schrago, C. G. (2019). Data partitioning and correction for ascertainment bias reduce the uncertainty of placental mammal divergence times inferred from the morphological clock. *Ecol. Evol.* 9, 2255–2262. doi: 10.1002/ece3.4921

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

Carroll, R. L. (1997). *Patterns and Processes of Vertebrate Evolution* (Cambridge, United Kingdom: Cambridge University Press).

Chen, M.-Y., Liang, D., and Zhang, P. (2017). Phylogenomic resolution of the phylogeny of laurasiatherian mammals: Exploring phylogenetic signals within coding and noncoding sequences. *Genome Biol. Evol.* 9, 1998–2012. doi: 10.1093/gbe/evx147

Chiari, Y., Cahais, V., Galtier, N., and Delsuc, F. (2012). Phylogenomic analyses support the position of turtles as the sister group of birds and crocodiles (Archosauria). *BMC Biol.* 10, 65. doi: 10.1186/1741-7007-10-65

Cohen, J. E., and Cifelli, R. L. (2015). "The first eutherian mammals from the Early Late Cretaceous of North America," in *75th Society of Vertebrate Paleontology Annual Meeting* (Dallas, TX, USA: Society of Vertebrate Paleontology), 108.

Cohen, J. E. (2017). Radiation of Tribosphenic Mammals During the Early Late Cretaceous (Turonian) of North America. Unpubl. Ph.D. dissertation. Univ. Oklahoma.

Cooper, L. N., Seiffert, E. R., Clementz, M., Madar, S. I., Bajpai, S., Hussain, S. T., et al. (2014). Anthracobunids from the Middle Eocene of India and Pakistan are stem perissodactyls. *PLoS One* 9, e109232. doi: 10.1371/journal.pone.0109232

D'Erchia, A. M., Gissi, C., Pesole, G., Saccone, C., and Arnason, U. (1996). The guinea pig is not a rodent. *Nature* 381, 597–600. doi: 10.1038/381597a0

Darwin, C. (1859). *The Origin of Species by Means of Natural Selection* (London: John Murray).

Davies, K. T. J., Cotton, J. A., Kirwan, J. D., Teelling, E. C., and Rossiter, S. J. (2012). Parallel signatures of sequence evolution among hearing genes in echolocating mammals: an emerging model of genetic convergence. *Heredity* 108, 480–489. doi: 10.1038/hdy.2011.119

Davies, T. W., Bell, M. A., Goswami, A., and Halliday, T. J. (2017). Completeness of the eutherian mammal fossil record and implications for reconstructing mammal evolution through the Cretaceous/Paleogene mass extinction. *Paleobiology* 43, 521–536. doi: 10.1017/pab.2017.20

de Muizon, C., Bianucci, G., Martínez-Cáceres, M., and Lambert, O. (2019). *Mystacodon selenensis*, the earliest known toothed mysticete (Cetacea, Mammalia) from the late Eocene of Peru: anatomy, phylogeny, and feeding adaptations. *Geodiversitas* 41, 401–499. doi: 10.5252/geodiversitas2019v41a11

de Queiroz, K. (1998). "The general lineage concept of species, species criteria, and the process of speciation: A conceptual unification and terminological recommendations," in *Endless Forms: Species and Speciation*. Eds. D. J. Howard, and S. H. Berlocher (New York: Oxford University Press), 57–75.

Degnan, J. H., and Rosenberg, N. A. (2009). Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends Ecol. Evol.* 24, 332–340. doi: 10.1016/j.tree.2009.01.009

Delsuc, F., Vizcaino, S. F., and Douzery, E. J. P. (2004). Influence of Tertiary paleoenvironmental changes on the diversification of South American mammals: a relaxed molecular clock study with xenarthrans. *BMC Evol. Biol.* 4, 11. doi: 10.1186/1471-2148-4-11

Dickerson, R. E. (1971). The structure of cytochrome c and the rates of molecular evolution. *J. Mol. Evol.* 1, 26–45. doi: 10.1007/BF01659392

Didier, G., and Laurin, M. (2018). Exact distribution of divergence times from fossil ages and topologies. *bioRxiv*, 490003. doi: 10.1101/490003

Donoghue, P. C., and Benton, M. J. (2007). Rocks and clocks: Calibrating the tree of life using fossils and molecules. *Trends Ecol. Evol.* 22, 424–431. doi: 10.1016/j.tree.2007.05.005

dos Reis, M., Inoue, J., Hasegawa, M., Asher, R. J., Donoghue, P. C., and Yang, Z. (2012). Phylogenomic datasets provide both precision and accuracy in estimating the timescale of placental mammal phylogeny. *Proc. R. Soc. Lond. B* 279, 3491–3500. doi: 10.1098/rspb.2012.0683

dos Reis, M., Donoghue, P. C., and Yang, Z. (2014). Neither phylogenomic nor palaeontological data support a Palaeogene origin of placental mammals. *Biol. Lett.* 10, 20131003. doi: 10.1098/rsbl.2013.1003

dos Reis, M., Donoghue, P. C., and Yang, Z. (2016). Bayesian molecular clock dating of species divergences in the genomics era. *Nat. Rev. Genet.* 17, 71–80. doi: 10.1038/nrg.2015.8

Easteal, S. (1999). Molecular evidence for the early divergence of placental mammals. *BioEssays* 21, 1052–1058. doi: 10.1002/(SICI)1521-1878(199912)22:1<1052::AID-BIES9>3.0.CO;2-6

Edelman, N. B., Frandsen, P. B., Miyagi, M., Clavijo, B., Davey, J., Dikow, R., et al. (2019). Genomic architecture and introgression shape a butterfly radiation. *Science* 366, 594–599. doi: 10.1126/science.aaw2090

Eizirik, E., Murphy, W. J., and O'Brien, S. J. (2001). Molecular dating and biogeography of the early placental mammal radiation. *J. Hered.* 92, 212–219. doi: 10.1093/jhered/92.2.212

Emerling, C. A., Huynh, H. T., Nguyen, M. A., Meredith, R. W., and Springer, M. S. (2015). Spectral shifts of mammalian ultraviolet-sensitive pigments (short wavelength-sensitive opsin 1) are associated with eye length and photic niche evolution. *Proc. R. Soc. Lond. B* 282, 20151817. doi: 10.1098/rspb.2015.1817

Erdal, O., Antoine, P.-O., and Sen, S. (2016). New material of *Palaeoamasia kansui* (Embrithopoda, Mammalia) from the Eocene of Turkey and a phylogenetic analysis of Embrithopoda at the species level. *Palaeontology* 59, 631–655. doi: 10.1111/pala.12247

Feijoo, M., and Parada, A. (2017). Macrosystematics of eutherian mammals combining HTS data to expand taxon coverage. *Mol Phylogenet Evol.* 113, 76–83. doi: 10.1016/j.ympev.2017.05.004

Figueiró, H. V., Li, G., Trindade, F. J., Assis, J., Pais, F., Fernandes, G., et al. (2017). Genome-wide signatures of complex introgression and adaptive evolution in the big cats. *Sci. Adv.* 3, e1700299. doi: 10.1126/sciadv.1700299

Foley, N. M., Springer, M. S., and Teeling, E. C. (2016). Mammal madness: Is the mammal tree of life not yet resolved? *Phil. Trans. R. Soc. Lond. B* 371, 20150140. doi: 10.1098/rstb.2015.0140

Fontaine, M. C., Pease, J. B., Steele, A., Waterhouse, R. M., Neafsey, D. E., Sharakhov, I. V., et al. (2015). Extensive introgression in a malaria vector species complex revealed by phylogenomics. *Science* 347, 1258524. doi: 10.1126/science.1258524

Gatesy, J., and Springer, M. S. (2017). Phylogenomic red flags: Homology errors and zombie lineages in the evolutionary diversification of placental mammals. *Proc. Natl. Acad. Sci. U.S.A.* 114, E9431–E9432. doi: 10.1073/pnas.1715318114

Gatesy, J., Geisler, J. H., Chang, J., Buell, C., Berta, A., Meredith, R. W., et al. (2013). A phylogenetic blueprint for a modern whale. *Mol. Phylogenet. Evol.* 66, 479–506. doi: 10.1016/j.ympev.2012.10.012

Gavryushkina, A., Heath, T. A., Ksepka, D. T., Stadler, T., Welch, D., and Drummond, A. J. (2017). Bayesian total-evidence dating reveals the recent crown radiation of penguins. *Syst. Biol.* 66, 57–73. doi: 10.1093/sysbio/syw060

Gheerbrant, E., Schmitt, A., and Kocsis, L. (2018). Early African fossils elucidate the origin of embrithopod mammals. *Curr. Biol.* 28, 2167–2173. doi: 10.1016/j.cub.2018.05.032

Gingerich, P. D., ul Haq, ,. M., Zalmout, I. S., Khan, I. H., and Malkani, M. S. (2001). Origin of whales from early artiodactyls: Hands and feet of Eocene Protocetidae from Pakistan. *Science* 293, 2239–2242. doi: 10.1126/science.1063902

Gingerich, P. (1977). "Patterns of evolution in the mammalian fossil record," in *Patterns of Evolution as Illustrated by the Fossil Record*, vol. pp . Ed. A. Hallam (Amsterdam, Netherlands: Elsevier Scientific, Amsterdam), 469–500. doi: 10.1016/S0920-5446(08)70335-2

Goloboff, P. A., Torres, A., and Arias, J. S. (2018). Weighted parsimony outperforms other methods of phylogenetic inference under models appropriate for morphology. *Cladistics* 34, 407–437. doi: 10.1111/cla.12205

Goswami, A., Prasad, G. V., Upchurch, P., Boyer, D. M., Seiffert, E. R., Verma, O., et al. (2011). A radiation of arboreal basal eutherian mammals beginning in the Late Cretaceous of India. *Proc. Natl. Acad. Sci. U.S.A.* 108, 16333–16338. doi: 10.1073/pnas.1108723108

Graur, D., Hide, W. A., and Li, W.-H. (1991). Is the guinea-pig a rodent? *Nature* 351, 649–652. doi: 10.1038/351649a0

Graur, D., Zarkikh, A., Hide, W. A., and Li, W.-H. (1992). The biochemical phylogeny of guinea pigs and gundies and the paraphyly of the order Rodentia. *Comp. Biochem. Physiol. B* 101, 495–498. doi: 10.1016/0305-0491(92)90327-N

Gunnell, G. F., Bown, T. M., and Bloch, J. I. (2007). *Leptictida. In: Evolution of Tertiary Mammals of North America* Vol. 2. Eds. C. M. Janis, G. F. Gunnell, and M. D. Uhen (Cambridge: Cambridge University Press), 82–88. doi: 10.1017/CBO9780511541438.007

Halliday, T. J. D., Upchurch, P., and Goswami, A. (2016). Eutherians experienced elevated evolutionary rates in the immediate aftermath of the Cretaceous-Palaeogene mass extinction. *Proc. R. Soc. Lond. B* 283, 20153026. doi: 10.1098/rspb.2015.3026

Halliday, T. J. D., Upchurch, P., and Goswami, A. (2017). Resolving the relationships of Paleocene placental mammals. *Biol. Rev.* 92, 521–550. doi: 10.1111/brv.12242

Halliday, T. J., dos Reis, M., Tamuri, A. U., Ferguson-Gow, H., Yang, Z., and Goswami, A. (2019). Rapid morphological evolution in placental mammals post-dates the origin of the crown group. *Proc. R. Soc. Lond. B* 286, 20182418. doi: 10.1098/rspb.2018.2418

Hallström, B. M., and Janke, A. (2008). Resolution among major placental mammal interordinal relationships with genome data imply that speciation influenced their earliest radiations. *BMC Evol. Biol.* 8, 162. doi: 10.1186/1471-2148-8-162

Heath, T. A., Huelsenbeck, J. P., and Stadler, T. (2014). The fossilized birth-death process for coherent calibration of divergence-times estimates. *Proc. Natl. Acad. Sci. U.S.A.* 111, E2957–E2966. doi: 10.1073/pnas.1319091111

Hedges, S. B., Parker, P. H., Sibley, C. G., and Kumar, S. (1996). Continental breakup and the ordinal diversification of birds and mammals. *Nature* 381, 226–229. doi: 10.1038/381226a0

Hedges, S. B., Marin, J., Suleski, M., Paymer, M., and Kumar, S. (2015). Tree of life reveals clock-like speciation and diversification. *Mol. Biol. Evol.* 32, 835–845. doi: 10.1093/molbev/msv037

Heled, J., and Drummond, A. J. (2010). Bayesian inference of species trees from multilocus data. *Mol. Biol. Evol.* 27, 570–580. doi: 10.1093/molbev/msp274

Hibbins, M. S., and Hahn, M. W. (2019). The timing and direction of introgression under the multispecies network coalescent. *Genetics* 211, 1059–1073. doi: 10.1534/genetics.118.301831

Ho, S., and Phillips, M. J. (2009). Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Syst. Biol.* 58, 367–380. doi: 10.1093/sysbio/syp035

Hobolth, A., Christensen, O. F., Mailund, T., and Schierup, M. H. (2007). Genomic relationships and speciation times of human, chimpanzee, and gorilla inferred from a coalescent hidden Markov model. *PLoS Genet.* 3, e7. doi: 10.1371/journal.pgen.0030007

Hobolth, A., Dutheil, J. Y., Hawks, J., Schierup, M. H., and Mailund, T. (2011). Incomplete lineage sorting patterns among human, chimpanzee, and orangutan suggest recent orangutan speciation and widespread selection. *Genome Res.* 21, 349–356. doi: 10.1101/gr.114751.110

Holland, S. M. (2016). The non-uniformity of fossil preservation. *Philos. Trans. R. Soc. B* 371, 20150130. doi: 10.1098/rstb.2015.0130

Inoue, J., Donoghue, P. C., and Yang, Z. (2009). The impact of the representation of fossil calibrations on Bayesian estimation of species divergence times. *Syst. Biol.* 59, 74–89. doi: 10.1093/sysbio/syp078

Jarvis, E. D., Mirarab, S., Aberer, A. J., Li, B., Houde, P., Li, C., et al. (2014). Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science* 346, 1320–1331. doi: 10.1126/science.1253451

Kornegay, J. R., Schilling, J. W., and Wilson, A. C. (1994). Molecular adaptation of a leaf-eating bird: stomach lysozyme of the hoatzin. *Mol. Biol. Evol.* 11, 921–928. doi: 10.1093/oxfordjournals.molbev.a040173

Krause, J., Unger, T., Nocon, A., Malaspinas, A. S., Kolokotronis, S.-O., Stiller, M., et al. (2008). Mitochondrial genomes reveal an explosive radiation of extinct and extant bears near the Miocene-Pliocene boundary. *BMC Evol. Biol.* 8, 220. doi: 10.1186/1471-2148-8-220

Krause, D. W., Hoffmann, S., Wible, J. R., Kirk, E. C., Schultz, J. A., von Koenigswald, W., et al. (2014). First cranial remains of a gondwanatherian mammal reveal remarkable mosaicism. *Nature* 515, 512–517. doi: 10.1038/nature13922

Kumar, S., and Hedges, S. B. (1998). A molecular timescale for vertebrate evolution. *Nature* 392, 917–920. doi: 10.1038/31927

Kumar, S., and Hedges, S. B. (2016). Advances in time estimation methods for molecular data. *Mol. Biol. Evol.* 33, 863–869. doi: 10.1093/molbev/msw026

Kumar, V., Hallström, B. M., and Janke, A. (2013). Coalescent-based genome analyses resolve the early branches of the Euarchontoglires. *PLoS One* 8, e60019. doi: 10.1371/journal.pone.0060019

Kumar, S. (2005). Molecular clocks: Four decades of evolution. *Nat. Rev. Gen.* 6, 654–662. doi: 10.1038/nrg1659

Lambert, O., Martínez-Cáceres, M., Bianucci, G., Di Celma, C., Salas-Gismondi, R., Steurbaut, W., et al. (2017a). Earliest mysticete from the Late Eocene of Peru sheds new light on the origin of baleen whales. *Curr. Biol.* 27, 1535–1541. doi: 10.1016/j.cub.2017.04.026

Lambert, O., Bianucci, G., and de Muizon, C. (2017b). Macroraptorial sperm whales (Cetacea, Odontoceti, Physeteroidea) from the Miocene of Peru. *Zool. J. Linn. Soc.* 179, 404–474. doi: 10.1111/zoj.12456

Lamichhaney, S., Berglund, J., Almén, M. S., Maqbool, K., Grabherr, M., Martinez-Barrio, A., et al. (2015). Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature* 518, 371–375. doi: 10.1038/nature14181

Larsen, P. A., Marchán Rivadeneira, M. R., and Baker, R. J. (2010). Natural hybridization generates mammalian lineage with species characteristics. *Proc. Natl. Acad. Sci. U.S.A.* 107, 11447–11452. doi: 10.1073/pnas.1000133107

Lartillot, N., and Delsuc, F. (2012). Joint reconstruction of divergence times and life-history evolution in placental mammals using a phylogenetic covariance model. *Evolution* 66, 1773–1787. doi: 10.1111/j.1558-5646.2011.01558.x

Leaché, A. D., Harris, R. B., Rannala, B., and Yang, Z. (2014). The influence of gene flow on species tree estimation: a simulation study. *Syst Biol.* 63, 17–30. doi: 10.1093/sysbio/syt049

Lee, M. S. Y., and Palci, A. (2015). Morphological phylogenetics in the genomic age. *Curr. Biol.* 25, R922–R929. doi: 10.1016/j.cub.2015.07.009

Lemey, P., and Posada, D. (2009). "Introduction to recombination detection," in *The Phylogenetic Handbook*, 2nd ed. Eds. P. Lemey, M. Salemi, and Vandamme (Cambridge: Cambridge University Press), 493–518. doi: 10.1017/CBO9780511819049.017

Lewis, P. O. (2001). A likelihood approach to estimating phylogeny from discrete morphological character data. *Syst. Biol.* 50, 913–925. doi: 10.1080/106351501753462876

Li, W.-H., Gouy, M., Sharp, P., O'hUigin, C., and Yang, Y.-W. (1990). Molecular phylogeny of Rodentia, Lagomorpha, Primates, Artiodactyla, and Carnivora and molecular clocks. *Proc. Natl. Acad. Sci. U. S. A.* 87, 6703–6707. doi: 10.1073/pnas.87.17.6703

Li, W.-H., Hide, W. A., Zharkikh, A., Ma, D. P., and Graur, D. (1992). The molecular taxonomy and evolution of the guinea pig. *J. Hered.* 83, 174–181. doi: 10.1093/oxfordjournals.jhered.a111188

Li, G., Davis, B. W., Eizirik, E., and Murphy., W. J. (2016). Phylogenomic evidence for ancient hybridization in the genomes of living cats (Felidae). *Genome Res.* 26, 1–11. doi: 10.1101/gr.186668.114

Li, G., Figueiró, H. V., Eizirik, E., and Murphy, W. J. (2019). Recombination-aware phylogenomics reveals the structured genomic landscape of hybridizing cat species. *Mol. Biol. Evol.* 36, 2111–2126. doi: 10.1093/molbev/msz139

Liu, L., Yu, L., Kubatko, L. S., Pearl, D. K., and Edwards, S. V. (2009). Coalescent methods for estimating phylogenetic trees. *Mol. Phylogenet. Evol.* 53, 320–328. doi: 10.1016/j.ympev.2009.05.033

Liu, Y., Cotton, J. A., Shen, B., Han, X., Rossiter, S. J., and Zhang, S. (2010). Convergent sequence evolution between echolocating bats and dolphins. *Curr. Biol.* 20, R53–R54. doi: 10.1016/j.cub.2009.11.058

Liu, L., Zhang, J., Rheindt, F. E., Lei, F., Qu, Y., Wang, Y., et al. (2017a). Genomic evidence reveals a radiation of placental mammals uninterrupted by the KPg boundary. *Proc. Natl. Acad. Sci. U.S.A.* 114, E7282–E7290. doi: 10.1073/pnas.1616744114

Liu, L., Zhang, J., Rheindt, F. E., Lei, F., Qu, Y., Wang, Y., et al. (2017b). Reply to Gatesy and Springer: Claims of homology errors and zombie lineages do not compromise the dating of placental diversification. *Proc. Natl. Acad. Sci. U.S.A.* 114, E9433–E9434. doi: 10.1073/pnas.1715371114

Luo, Z. X., Yuan, C. X., Meng, Q. J., and Ji, Q. (2011). A Jurassic eutherian mammal and divergence of marsupials and placentals. *Nature* 476, 442–445. doi: 10.1038/nature10291

Möller-Krull, M., Delsuc, F., Churakov, G., Marker, C., Superina, M., Brosius, J., et al. (2007). Retroposed elements and their flanking regions resolve the evolutionary history of xenarthran mammals (armadillos, anteaters, sloths). *Mol. Biol. Evol.* 24, 2573–2582. doi: 10.1093/molbev/msm201

Madar, S. I., Thewissen, J. G. M., and Hussain, S. T. (2002). Additional holotype remains of *Ambulocetus natans* (Cetacea, Ambulocetidae) and their implications for locomotion in early whales. *J. Vertebr. Paleontol.* 22, 405–422. doi: 10.1671/0272-4634(2002)022[0405:AHROAN]2.0.CO;2

Madsen, O., Scally, M., Douady, C. J., Kao, D. J., DeBry, R. W., Adkins, R., et al. (2001). Parallel adaptive radiations in two major clades of placental mammals. *Nature* 409, 610. doi: 10.1038/35054544

Marshall, C. R. (1997). Confidence intervals on stratigraphic ranges with nonrandom distributions of fossil horizons. *Paleobiology* 23, 165–173. doi: 10.1017/S0094837300016766

Martin, S. H., and Jiggins, C. D. (2017). Interpreting the genomic landscape of introgression. *Curr. Opin. Genet. Develop.* 47, 69–74. doi: 10.1016/j.gde.2017.08.007

Martin, S. H., Davey, J. W., Salazar, C., and Jiggins, C. D. (2019). Recombination rate variation shapes barriers to introgression across butterfly genomes. *PLoS Biol.* 17, e2006288. doi: 10.1371/journal.pbio.2006288

Marx, F. G., and Fordyce, R. E. (2015). Baleen boom and bust: a synthesis of mysticete phylogeny, diversity and disparity. *R. Soc. open sci.* 2, 140434. doi: 10.1098/rsos.140434

Mason, V. C., Li, G., Minx, P., Schmitz, J., Churakov, G., Doronina, L., et al. (2016). Genomic analysis reveals remarkable hidden biodiversity within colugos, the sister group to primates. *Sci. Adv.* 2, e1600633. doi: 10.1126/sciadv.1600633

McGowen, M. R., Spaulding, M., and Gatesy, J. (2009). Divergence date estimation and a comprehensive molecular tree of extant cetaceans. *Mol.Phylogenet.Evol.* 53, 891–906. doi: 10.1016/j.ympev.2009.08.018

McKenna, M. C., and Bell, S. K. (1997). *Classification of Mammals Above the Species Level* (New York: Columbia Univ. Press).

Meng, J., Hu, Y., and Li, C. (2003). The osteology of *Rhombomylus* (Mammalia, Glires): implications for phylogeny and evolution of Glires. *Bull. Am. Mus. Nat. Hist.* 275, 1–247. doi: 10.1206/0003-0090(2003)275<0001:TOORMG>2.0.CO;2

Meng, J. (2004). Phylogeny and divergence of basal Glires. *Bull. Am. Mus. Nat. Hist.* 285, 93–109. doi: 10.1206/0003-0090(2004)285<0093:C>2.0.CO;2

Meredith, R. W., Janecka, J. E., Gatesy, J., Ryder, O. A., Fisher, C. A., Teeling, E. C., et al. (2011). Impacts of the Cretaceous Terrestrial Revolution and KPg extinction on mammal diversification. *Science* 334, 521–524. doi: 10.1126/science.1211028

Morgan, C. C., Foster, P. G., Webb, A. E., Pisani, D., McInerney, J. O., and O'Connell, M. J. (2013). Heterogeneous models place the root of the placental mammal phylogeny. *Mol. Biol. Evol.* 30, 2145–2156. doi: 10.1093/molbev/mst117

Murphy, W. J., Eizirik, E., Johnson, W. E., Zhang, Y. P., Ryder, O. A., and O'Brien, S. J. (2001a). Molecular phylogenetics and the origins of placental mammals. *Nature* 409, 614–618. doi: 10.1038/35054550

Murphy, W. J., Eizirik, E., O'Brien, S. J., Madsen, O., Scally, M., Douady, C. J., et al. (2001b). Resolution of the early placental mammal radiation using Bayesian phylogenetics. *Science* 294, 2348–2351. doi: 10.1126/science.1067179

Murphy, W. J., Pringle, T. H., Crider, T. A., Springer, M. S., and Miller, W. (2007). Using genomic data to unravel the root of the placental mammal phylogeny. *Genome Res.* 17, 413–421. doi: 10.1101/gr.5918807

Nachman, M. W., and Payseur, B. A. (2012). Recombination rate variation and speciation: theoretical predictions and empirical results from rabbits and mice. *Philos. Trans. R. Soc. Lond. B* 367, 409–421. doi: 10.1098/rstb.2011.0249

Nishihara, H., Hasegawa, M., and Okada, N. (2006). Pegasoferae, an unexpected mammalian clade revealed by tracking ancient retroposon insertions. *Proc. Natl. Acad. Sci. U.S.A.* 103, 9929–9934. doi: 10.1073/pnas.0603797103

Nishihara, H., Maruyama, S., and Okada, N. (2009). Retroposon analysis and recent geological data suggest near-simultaneous divergence of the three superorders of mammals. *Proc. Natl. Acad. Sci. U.S.A.* 106, 5235–5240. doi: 10.1073/pnas.0809297106

Novacek, M. J. (1986). The skull of leptictid insectivorans and the higher-level classification of eutherian mammals. *Bull. Am. Mus. Nat. Hist.* 183, 1–112.

Novacek, M. J. (1992). Mammalian phylogeny: shaking the tree. *Nature* 356, 121–125. doi: 10.1038/356121a0

O'Leary, M. A., Bloch, J. I., Flynn, J. J., Gaudin, T. J., Giallombardo, A., Giannini, N. P., et al. (2013). The placental mammal ancestor and the post-K-Pg radiation of placentals. *Science* 339, 662–667. doi: 10.1126/science.1229237

Palkopoulou, E., Lipson, M., Mallick, S., Nielsen, S., Rohland, N., Baleka, S., et al. (2018). A comprehensive genomic history of extinct and living elephants. *Proc. Natl. Acad. Sci. U.S.A.* 115, E2566–E2574. doi: 10.1073/pnas.1720554115

Parham, J. F., Donoghue, P. C., Bell, C. J., Calway, T. D., Head, J. J., Holroyd, P. A., et al. (2012). Best practices for justifying fossil calibrations. *Syst. Biol.* 61, 346–359. doi: 10.1093/sysbio/syr107

Pattinson, D. J., Thompson, R. S., Piotrowski, A. K., and Asher, R. J. (2015). Phylogeny, paleontology, and Primates: Do incomplete fossils bias the tree of life? *Syst. Biol.* 64, 169–186. doi: 10.1093/sysbio/syu077

Pease, J. B., and Hahn, M. W. (2013). More accurate phylogenies inferred from low-recombination regions in the presence of incomplete lineage sorting. *Evolution* 67, 2376–2384. doi: 10.1111/evo.12118

Phillips, M. J., and Fruciano, C. (2018). The soft explosive model of placental mammal evolution. *BMC Evol. Biol.* 18, 104. doi: 10.1186/s12862-018-1218-x

Phillips, M. J. (2016). Geomolecular dating and the origin of placental mammals. *Syst. Biol.* 65, 546–557. doi: 10.1093/sysbio/syv115

Posada, D., and Crandall, K. A. (2002). The effect of recombination on the accuracy of phylogeny estimation. *J. Mol. Evol.* 54, 396–402. doi: 10.1007/s00239-001-0034-9

Posada, D. (2000). How does recombination affect phylogeny estimation? *Trends Ecol. Evol.* 15, 489–490. doi: 10.1016/S0169-5347(00)02027-9

Presgraves, D. C. (2018). Evaluating genomic signatures of "the large X-effect" during complex speciation. *Mol. Ecol.* 27, 3822–3830. doi: 10.1111/mec.14777

Puttick, M. N., Thomas, G. H., and Benton, M. J. (2016). Dating Placentalia: Morphological clocks fail to close the molecular fossil gap. *Evolution* 70, 873–886. doi: 10.1111/evo.12907

Pyron, R. A. (2009). A likelihood method for assessing molecular divergence time estimates and the placement of fossil calibrations. *Syst. Biol.* 59, 185–194. doi: 10.1093/sysbio/syp090

Pyron, R. A. (2011). Divergence time estimation using fossils as terminal taxa and the origins of Lissamphibia. *Syst. Biol.* 60, 466–481. doi: 10.1093/sysbio/syr047

Rambaut, A., and Bromham, L. (1998). Estimating divergence dates from molecular sequences. *Mol. Biol. Evol.* 15, 442–448. doi: 10.1093/oxfordjournals.molbev.a025940

Rambaut, A. (2000). Estimating the rate of molecular evolution: incorporating non-contemporaneous sequences into maximum likelihood phylogenies. *Bioinformatics* 16, 395–399. doi: 10.1093/bioinformatics/16.4.395

Reyes, A., Gissi, C., Pesole, G., Catzeflis, F. M., and Saccone, C. (2000). Where do rodents fit in? Evidence from the complete mitochondrial genome of *Sciurus vulgaris*. *Mol. Biol. Evol.* 17, 979–983. doi: 10.1093/oxfordjournals.molbev.a026379

Robinson, D. F., and Foulds, L. R. (1981). Comparison of phylogenetic trees. *Math. Biosci.* 53, 131–147. doi: 10.1016/0025-5564(81)90043-2

Romiguier, J., Ranwez, V., Delsuc, F., Galtier, N., and Douzery, E. J. P. (2013). Less is more in mammalian phylogenomics: AT-rich genes minimize tree conflicts and unravel the root of placental mammals. *Mol. Biol. Evol.* 30, 2124–2144. doi: 10.1093/molbev/mst116

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. *Syst. Biol.* 61, 973–999. doi: 10.1093/sysbio/sys058

Rose, K. D. (1999). *Eurotamandua* and Palaenodonta: convergent or related? *PaläOntologische Zeitschrift*. 73, 395–401. doi: 10.1007/BF02988050

Sansom, R. S., and Wills, M. A. (2013). Fossilization causes organisms to appear erroneously primitive by distorting evolutionary trees. *Sci. Rep.* 3, 2545. doi: 10.1038/srep02545

Sato, J. J., Ohdachi, S. D., Echenique-Diaz, L. M., Borroto-Páez, R., Begué-Quiala, G., Delgado-Labañino, J. L., et al. (2016). Molecular phylogenetic analysis of nuclear genes suggests a Cenozoic over-water dispersal origin for the Cuban solenodon. *Sci. Rep.* 6, 31173. doi: 10.1038/srep31173

Scally, M., Madsen, O., Douady, C. J., de Jong, W. W., Stanhope, M. J., and Springer, M. S. (2001). Molecular evidence for the major clades of placental mammals. *J. Mamm. Evol.* 8, 239–277. doi: 10.1023/A:1014446915393

Schierup, M. H., and Hein, J. (2000). Consequences of recombination on traditional phylogenetic analysis. *Genetics* 156, 879–891.

Schumer, M., Xu, C., Powell, D., Durvasula, A., Skov, L., Holland, C., et al. (2018). Natural selection interacts with the local recombination rate to shape the evolution of hybrid genomes. *Science* 360, 656–660. doi: 10.1126/science.aar3684

Silcox, M. T., Bloch, J. I., Boyer, D. M., Chester, S. G. B., and López Torres, S. (2017). The evolutionary radiation of plesiadapiforms. *Evol. Anthropol.* 26, 74–94. doi: 10.1002/evan.21526

Simmons, M. P., Sloan, D. B., and Gatesy, J. (2016). The effects of subsampling gene trees on coalescent methods applied to ancient divergences. *Mol. Phylogenet. Evol.* 97, 76–89. doi: 10.1016/j.ympev.2015.12.013

Song, S., Liu, L., Edwards, S. V., and Wu, S. (2012). Resolving conflict in eutherian mammal phylogeny using phylogenomics and the multispecies coalescent model. *Proc. Natl. Acad. Sci. U.S.A.* 109, 14942–14947. doi: 10.1073/pnas.1211733109

Springer, M. S., and de Jong, W. W. (2001). Which mammalian supertree to bark up? *Science* 291, 1709–1711. doi: 10.1126/science.1059434

Springer, M. S., and Gatesy, J. (2016). The gene tree delusion. *Mol. Phylogenet. Evol.* 94, 1–33. doi: 10.1016/j.ympev.2015.07.018

Springer, M. S., and Gatesy, J. (2018a). Pinniped diphyly and bat triphyly: More homology errors drive conflicts in the mammalian tree. *J. Hered.* 109, 297–307. doi: 10.1093/jhered/esx089

Springer, M. S., and Gatesy, J. (2018b). On the importance of homology in the age of phylogenomics. *Syst. Biodivers.* 16, 210–228. doi: 10.1080/14772000.2017.1401016

Springer, M. S., and Gatesy, J. (2018c). Delimiting coalescence genes (c-genes) in phylogenomic data sets. *Genes* 9, 123. doi: 10.3390/genes9030123

Springer, M., and Lilje, A. (1988). Biostratigraphy and gap analysis: The expected sequence of biostratigraphic events. *J. Geol.* 96, 228–236. doi: 10.1086/629212

Springer, M. S., and Murphy, W. J. (2007). Mammalian evolution and biomedicine: new views from phylogeny. *Biol. Rev.* 82, 375–392. doi: 10.1111/j.1469-185X.2007.00016.x

Springer, M. S., Cleven, G. C., Madsen, O., de Jong, W. W., Waddell, V. G., Amrine, H. M., et al. (1997). Endemic African mammals shake the phylogenetic tree. *Nature* 388, 61–64. doi: 10.1038/40386

Springer, M. S., Murphy, W. J., Eizirik, E., and O'Brien, S. J. (2003). Placental mammal diversification and the Cretaceous-Tertiary boundary. *Proc. Natl. Acad. Sci. U.S.A.* 100, 1056–1061. doi: 10.1073/pnas.0334222100

Springer, M. S., Stanhope, M. J., Madsen, O., and de Jong, W. W. (2004). Molecules consolidate the placental mammal tree. *Trends Ecol. Evol.* 19, 430–438. doi: 10.1016/j.tree.2004.05.006

Springer, M. S., Murphy, W. J., Eizirik, E., and O'Brien, S. J. (2005). "Evidence for major placental clades," in *The Rise of Placental Mammals*. Eds. K. Rose, and J. Archibald (Baltimore, London: John Hopkins University Press), 45–64.

Springer, M. S., Burk-Herrick, A., Meredith, R., Eizirik, E., Teeling, E., O'Brien, S. J., et al. (2007). The adequacy of morphology for reconstructing the early history of placental mammals. *Syst. Biol.* 56, 673–684. doi: 10.1080/10635150701491149

Springer, M. S., Meredith, R. W., Eizirik, E., Teeling, E., and Murphy, W. J. (2008). Morphology and placental mammal phylogeny. *Syst. Biol.* 57, 499–503. doi: 10.1080/10635150802164504

Springer, M. S., Meredith, R. W., Janecka, J. E., and Murphy, W. J. (2011). The historical biogeography of Mammalia. *Phil. Trans. R. Soc. Lond. B* 366, 2478–2502. doi: 10.1098/rstb.2011.0023

Springer, M. S., Meredith, R. W., Teeling, E. C., and Murphy, W. J. (2013). Technical comment on "the placental mammal ancestor and the post-K-Pg radiation of placentals". *Science* 341, 613–613. doi: 10.1126/science.1238025

Springer, M. S., Emerling, C. A., Meredith, R. W., Janečka, J. E., Eizirik, E., and Murphy, W. J. (2017). Waking the undead: Implications of a soft explosive model for the timing of placental mammal diversification. *Mol. Phylogenet. Evol.* 106, 86–102. doi: 10.1016/j.ympev.2016.09.017

Springer, M. S. (1990). The effect of random range truncations on patterns of evolution in the fossil record. *Paleobiology* 16, 512–520. doi: 10.1017/S0094837300010228

Springer, M. S. (1997). Molecular clocks and the timing of the placental and marsupial radiations in relation to the Cretaceous-Tertiary boundary. *J. Mamm. Evol.* 4, 285–302. doi: 10.1023/A:1027378615412

Stanhope, M. J., Madsen, O., Waddell, V. G., Cleven, G. C., de Jong, W. W., and Springer, M. S. (1998a). Highly congruent molecular support for a diverse superordinal clade of endemic African mammals. *Mol. Phylogenet. Evol.* 9, 501–508. doi: 10.1006/mpev.1998.0517

Stanhope, M. J., Waddell, V. G., Madsen, O., de Jong, W., Hedges, S. B., Cleven, G. C., et al. (1998b). Molecular evidence for multiple origins of Insectivora and for a new order of endemic African insectivore mammals. *Proc. Natl. Acad. Sci. U.S.A.* 95, 9967–9972. doi: 10.1073/pnas.95.17.9967

Steeman, M. E., Hebsgaard, M. B., Fordyce, R. E., Ho, S. Y. W., Rabosky, D. L., Nielsen, R., et al. (2009). Radiation of extant cetaceans driven by restructuring of the oceans. *Syst. Biol.* 58, 573–585. doi: 10.1093/sysbio/syp060

Sterli, J., Pol, D., and Laurin, M. (2013). Incorporating phylogenetic uncertainty on phylogeny-based paleontological dating and the timing of turtle diversification. *Cladistics.* 29, 233–236. doi: 10.1111/j.1096-0031.2012.00425.x

Stevison, L. S., Woerner, A. E., Kidd, J. M., Kelley, J. L., Veeramah, K. R., McManus, K. F., et al. (2016). The time scale of recombination rate evolution in great apes. *Mol. Biol. Evol.* 33, 928–945. doi: 10.1093/molbev/msv331

Storch, G. (1981). *Eurotamandua joresi*, ein Myrmecophagide aus dem Eoziin der "Grube Messel" bei Darmstadt (Mammalia, Xenarthra). *Senckenbergiana Lethaea* 61, 247–289.

Strauss, D., and Sadler, P. M. (1989). Classical confidence intervals and Bayesian probability estimates for ends of local taxon ranges. *Math. Geol.* 21, 411–427. doi: 10.1007/BF00897326

Struck, T. H., Paul, C., Hill, N., Hartmann, S., Hösel, C., Kube, M., et al. (2011). Phylogenomic analyses unravel annelid evolution. *Nature* 471, 95–98. doi: 10.1038/nature09864

Struck, T. H. (2013). The impact of paralogy on phylogenomic studies - a case study on annelid relationships. *PLoS One* 8, e62892. doi: 10.1371/journal.pone.0062892

Swofford, D. L. (2002). PAUP* Phylogenetic Analysis Using Parsimony (* and Other Methods) Version 4. Sunderland, MA, Sinauer.

Szalay, F. S. (1968). The beginnings of Primates. *Evolution* 22, 19–36. doi: 10.1111/j.1558-5646.1968.tb03445.x

Szalay, F. S. (1977). "Phylogenetic relationships and a classification of the eutherian Mammalia," in *Major Patterns in Vertebrate Evolution*. Eds. M. K. Hecht, P. C. Goody, and B. M. Hecht (New York: Plenum Press), 315–374. doi: 10.1007/978-1-4684-8851-7_12

Tabuce, R., Marivaux, L., Adaci, M., Bensalah, M., Hartenberger, J.-L., Mahboudi, M., et al. (2007). Early Tertiary mammals from North Africa reinforce the molecular Afrotheria clade. *Proc. Roy. Soc. B* 274, 1159–1166. doi: 10.1098/rspb.2006.0229

Tarver, J. E., dos Reis, M., Mirarab, S., Moran, R. J., Parker, S., O'Reilly, J. E., et al. (2016). The interrelationships of placental mammals and the limits of phylogenetic inference. *Genome Biol. Evol.* 8, 330–344. doi: 10.1093/gbe/evv261

Tavaré, S., Marshall, C. R., Will, O., Soligo, C., and Martin, R. D. (2002). Using the fossil record to estimate the age of the last common ancestor of extant primates. *Nature* 416, 726–729. doi: 10.1038/416726a

Teeling, E. C., Springer, M. S., Madsen, O., Bates, P., O'Brien, S. J., and Murphy, W. J. (2005). A molecular phylogeny for bats illuminates biogeography and the fossil record. *Science* 307, 580–584. doi: 10.1126/science.1105113

Thewissen, J. G. M., Madar, S. I., and Hussain, S. T. (1996). *Ambulocetus natans*, an Eocene cetacean (Mammalia) from Pakistan. *Cour. Forsch. Senckenbg.* 191, 1–86.

Thorne, J. L., Kishino, H., and Painter, I. S. (1998). Estimating the rate of evolution of the rate of molecular evolution. *Mol. Biol. Evol.* 15, 1647–1657. doi: 10.1093/oxfordjournals.molbev.a025892

Waddell, P. J., Kishino, H., and Ota, R. (2001). A phylogenetic foundation for comparative mammalian genomics. *Genome Inform.* 12, 141–154. doi: 10.11234/gi1990.12.141

Wang, R. J., and Hahn, M. W. (2018). Speciation genes are more likely to have discordant gene trees. *Evol. Lett.* 2, 281–296. doi: 10.1002/evl3.77

Wang, X., McKenna, M. C., and Dashzeveg, D. (2005). *Amphicticeps* and *Amphicynodon* (Arctoidea, Carnivora) from Hsanda Gol Formation, Central Mongolia and phylogeny of basal arctoids with comments on zoogeography. *Am. Mus. Novitat.* 3483, 1–57. doi: 10.1206/0003-0082(2005)483[0001:AAAACF]2.0.CO;2

Welker, F., Collins, M. J., Thomas, J. A., Wadsley, M., Brace, S., Cappellini, E., et al. (2015). Ancient proteins resolve the evolutionary history of Darwin's South American ungulates. *Nature* 522, 81–84. doi: 10.1038/nature14249

Westbury, M., Baleka, S., Barlow, A., Hartmann, S., Paijmans, J. L. A., Kramarz, A., et al. (2017). A mitogenomic timetree for Darwin's enigmatic South American mammal *Macrauchenia patachonica*. *Nat. Comm.* 8, 15951. doi: 10.1038/ncomms15951

Wible, J. R., Rougier, G. W., Novacek, M. J., and Asher, R. J. (2007). Cretaceous eutherians and Laurasian origin for placental mammals near the K/T boundary. *Nature* 447, 1003–1006. doi: 10.1038/nature05854

Wible, J. R., Rougier, G. W., Novacek, M. J., and Asher, R. J. (2009). The eutherian mammal *Maelestes gobiensis* from the late Cretaceous of Mongolia and the phylogeny of Cretaceous Eutheria. *Bull. Am. Mus. Nat. Hist.* 3, 1–123. doi: 10.1206/623.1

Wildman, D. E., Uddin, M., Opazo, J. C., Liu, G., Lefort, V., Guindon, S., et al. (2007). Genomics, biogeography, and the diversification of placental mammals. *Proc. Natl. Acad. Sci. U.S.A.* 104, 14395–14400. doi: 10.1073/pnas.0704342104

Wilkinson, R. D., Steiper, M. E., Soligo, C., Martin, R. D., Yang, Z., and Tavaré, S. (2011). Dating primate divergences through an integrated analysis of palaeontological and molecular data. *Syst. Biol.* 60, 16–31. doi: 10.1093/sysbio/syq054

Wu, J., Yonezawa, T., and Kishino, H. (2017). Rates of molecular evolution suggest natural history of life history traits and a post-K-Pg nocturnal bottleneck of placentals. *Curr. Biol.* 27, 3025–3033. e3025. doi: 10.1016/j.cub.2017.08.043

Yang, Z., and Rannala, B. (2006). Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Mol. Biol. Evol.* 23, 212–226. doi: 10.1093/molbev/msj024

Young, J. Z. (1981). *The Life of Vertebrates*. 3rd edition (Oxford: Oxford University Press).

Zhou, C.-F., Wu, S., Martin, T., and Luo, Z.-X. (2013). A Jurassic mammaliaform and the earliest mammalian evolutionary adaptations. *Nature* 500, 163–167. doi: 10.1038/nature12429

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer EG and handling Editor declared their shared affiliation at the time of review.

Copyright © 2019 Springer, Foley, Brady, Gatesy and Murphy. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 17 January 2020
doi: 10.3389/fevo.2019.00457

A Total Evidence Phylogenetic Analysis of Pinniped Phylogeny and the Possibility of Parallel Evolution Within a Monophyletic Framework

Ryan S. Paterson[1]*, *Natalia Rybczynski*[1,2], *Naoki Kohno*[3,4] and *Hillary C. Maddin*[1]

[1] Department of Earth Sciences, Carleton University, Ottawa, ON, Canada, [2] Palaeobiology, Canadian Museum of Nature, Ottawa, ON, Canada, [3] Department of Geology and Paleontology, National Museum of Nature and Science, Tsukuba, Japan, [4] Graduate School of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Japan

In the present study, a series of phylogenetic analyses of morphological, molecular, and combined morphological-molecular datasets were conducted to investigate the relationships of 23 extant and 44 fossil caniforme genera, in order to test the phylogenetic position of putative stem pinniped *Puijila* within a comprehensive evolutionary framework. With *Canis* as an outgroup, a Bayesian Inference analysis employing tip-dating of a combined molecular-morphological (i.e., Total Evidence) dataset recovered a topology in which musteloids are the sister group to a monophyletic pinniped clade, to the exclusion of ursids, and recovered *Puijila* and *Potamotherium* along the stem of Pinnipedia. A similar topology was recovered in a parsimony analysis of the same dataset. These results suggest the pinniped stem may be expanded to include additional fossil arctoid taxa, including *Puijila, Potamotherium,* and *Kolponomos*. The tip-dating analysis suggested a divergence time between pinnipeds and musteloids of ∼45.16 million years ago (Ma), though a basal split between otarioids and phocids is not estimated to occur until ∼26.52 Ma. These results provide further support for prolonged freshwater and nearshore phases in the evolution of pinnipeds, prior to the evolution of the extreme level of aquatic adaptation displayed by extant taxa. Ancestral character state reconstruction was used to investigate character evolution, to determine the frequency of reversals and parallelisms characterizing the three extant clades within Pinnipedia. Although the phylogenetic analyses did not directly provide any evidence of parallel evolution within the pinniped extant families, it is apparent from the inspection of previously-proposed pinniped synapomorphies, within the context of a molecular-based phylogenetic framework, that many traits shared between extant pinnipeds have arisen independently in the three clades. Notably, those traits relating to homodonty and limb-bone specialization for aquatic locomotion appear to have multiple origins within the crown group, as suggested by the retention of the plesiomorphic conditions in early-diverging fossil members of the three extant families. Thus, while the present analysis identifies a new suite of morphological synapomorphies for Pinnipedia, the frequency of reversals and other homoplasies within the clade limit their diagnostic value.

Keywords: total evidence dating, pinniped, phylogenetic analysis, aquatic adaptation, parallel evolution

OPEN ACCESS

Edited by:
Michel Laurin, UMR7207 Centre de Recherche sur la Paléobiodiversité et les Paléoenvironnements (CR2P), France

Reviewed by:
Olaf R. P. Bininda-Emonds, University of Oldenburg, Germany
Morgan Churchill, University of Wisconsin–Oshkosh, United States

***Correspondence:**
Ryan S. Paterson
ryanpaterson@cmail.carleton.ca

Specialty section:
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Ecology and Evolution

Received: *13 May 2019*
Accepted: *13 November 2019*
Published: *17 January 2020*

Citation:
Paterson RS, Rybczynski N, Kohno N and Maddin HC (2020) A Total Evidence Phylogenetic Analysis of Pinniped Phylogeny and the Possibility of Parallel Evolution Within a Monophyletic Framework. Front. Ecol. Evol. 7:457. doi: 10.3389/fevo.2019.00457

INTRODUCTION

From the 1960's until the late 1980's, pinnipeds (seals, sea lions, and walruses) were widely believed to be diphyletic, based partly on the divergent locomotory styles and morphologies observed in extant taxa (Mivart, 1885; McLaren, 1960). A diphyletic view aligns phocids with musteloids (weasels, raccoons, and allies), and grouped otarioids (otariids: fur seals and sea lions; and odobenids: walrus and kin) with ursoids (bears) (**Figure 1A**; Ray, 1976; Tedford, 1976; Repenning and Tedford, 1977; de Muizon, 1982a,b; Wozencraft, 1989; Nojima, 1990). The paradigm shifted when Wyss (1987) provided evidence from inner ear morphology strongly suggesting a monophyletic relationship of phocids, otariids, and odobenids (**Figure 1B**). Wyss' careful observation of this anatomy and subsequent recognition of pinniped homologies corroborated earlier morphological (Weber, 1904; Gregory, 1910; Davies, 1958), cytogenetic (Fay et al., 1967), and biomolecular (Sarich, 1969; Arnason, 1974; Romero-Herrera et al., 1978; de Jong and Goodman, 1982; Miyamoto and Goodman, 1986; Sato et al., 2009) work identifying pinnipeds as monophyletic. From the publication of Wyss (1987) to the present, the vast majority of phylogenetic studies have confirmed Wyss' assertion, regardless if the nature of the evidence was molecular (Dragoo and Honeycutt, 1997; Flynn and Nedbal, 1998; Flynn et al., 2005; Fulton and Strobeck, 2006; Sato et al., 2006; Yonezawa et al., 2009), morphological (Berta and Wyss, 1994; Deméré and Berta, 2002; Kohno, 2006; Furbish, 2015), or a combination of the two (Flynn et al., 2000). However, some morphological (Koretsky et al., 2016), biomechanical (Kuhn and Frey, 2012) and biogeographical (Koretsky and Barnes, 2006) studies continue to uncover evidence questioning the monophyly of pinnipeds, though these studies disregard the wealth of molecular evidence in support of monophyly.

Though pinniped monophyly has been well-supported by both molecular and morphological data, it is not apparent how phocids and otariids developed such disparate morphologies and locomotory modes. A diphyletic origin of pinnipeds serves as a compelling argument to explain the contrasting swimming styles and morphological disparity observed between otariids and phocids (Kuhn and Frey, 2012). Otariids are more ursid-like in general appearance and retain the capability of inverting their hindlimbs on land, allowing them to perform some semblance of quadrupedal ambulatory locomotion (Berta, 2012). When submerged, otariids engage in pectoral oscillations, using their enlarged foreflippers to figuratively fly through the water (English, 1976). The more fusiform phocids are unable to revert their hindlimbs forward on land, and swim using pelvic oscillations, producing thrust underwater via lateral undulations of their pelvic region and alternate protraction and retraction of their hindflippers (Fish et al., 1988; Pierce et al., 2011). The highly autapomorphic *Odobenus rosmarus*, the lone extant member of the once speciose Odobenidae, is not easily accommodated into either group. *Odobenus* is more otariid-like in general appearance, but shares a peculiar suite of features with phocids, not found in otariids (Wyss, 1987; Berta and Wyss, 1994). *Odobenus* displays an intermediary swimming style, using pectoral oscillations at slow speeds and pelvic undulations at high speeds (Gordon, 1981). Interpretation of fossil evidence led diphyly proponents to identify *Odobenus* as a highly-derived otarioid, a notion supported by molecular studies, though not within the framework of a diphyletic Pinnipedia (Flynn et al., 2000; Fulton and Strobeck, 2007).

The evolutionary relationships advocated by proponents of pinniped diphyly were initially not subject to the rigor of appropriate cladistic methods (McLaren, 1960). Later studies invoking cladistic methods continued to recover these relationships, though these analyses typically excluded characters

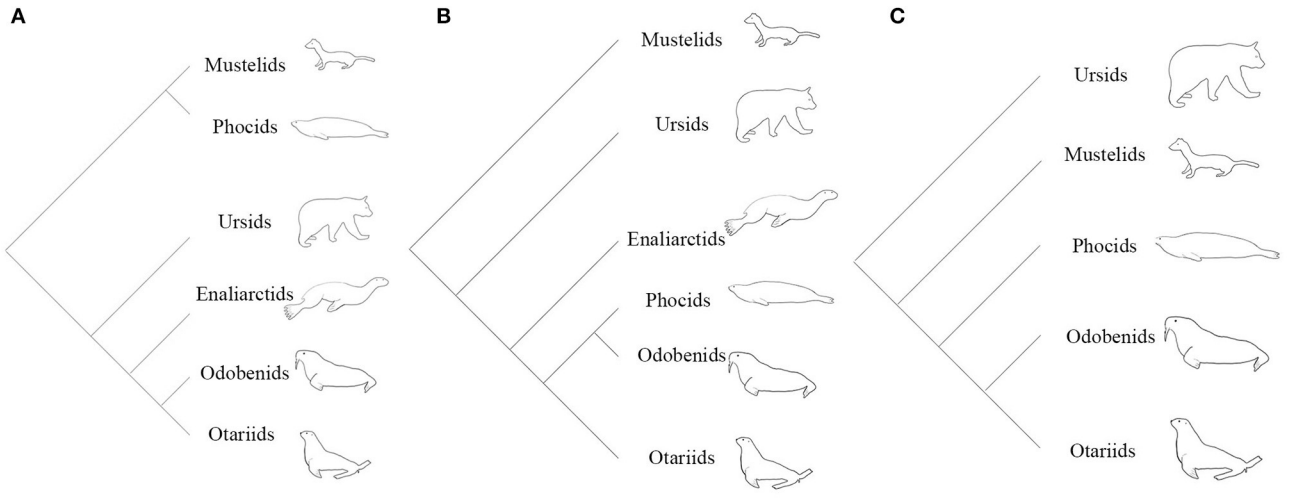


FIGURE 1 | (A) Hypothesis of the evolutionary relationships of arctoids under a diphyletic view. **(B)** Hypothesis of the evolutionary relationships of arctoids under a monophyletic view based on morphological evidence (Wyss, 1987; Berta and Wyss, 1994). **(C)** Hypothesis of the evolutionary relationships of arctoids under a monophyletic view using molecular data (Flynn et al., 2005; Fulton and Strobeck, 2007; Sato et al., 2009).

that may have arisen in parallel due to the functional constraints associated with aquatic specialization (Tedford, 1976; de Muizon, 1982a). Pinniped diphyly is supported by a small number of morphological features shared between otarioids and ursids on the one hand, and phocids and mustelids on the other hand (McLaren, 1960; Tedford, 1976; de Muizon, 1982a). The features shared between otarioids and phocids, largely in the postcranial skeleton and auditory region, were disregarded due to their perceived susceptibility to homoplasy (Tedford, 1976), a notion later shown to be untenable within an appropriate phylogenetic framework (Wyss, 1987, 1988). Cladistic analyses of large morphological datasets have unilaterally supported pinniped monophyly (Berta and Wyss, 1994). However, such analyses have had difficulty establishing the relationship of pinnipeds to other arctoids, as their results are typically incongruent with the relationships identified by molecular analyses (e.g., Berta and Wyss, 1994 vs. Bininda-Emonds et al., 1999; Fulton and Strobeck, 2007). This problem is exacerbated by the paucity of fossil remains representing early-diverging pinnipeds that could fill the gap between them and any putative arctoid clade (Deméré et al., 2003).

Since the establishment of pinniped monophyly (Wyss, 1987, 1988; Berta et al., 1989), the most contentious issue in pinniped systematics has been the placement of pinnipeds within Arctoidea. Phylogenetic analyses of morphological data are in disagreement over the sister group of pinnipeds, with musteloids and ursids both presenting compelling cases. Identification of Ursidae as the sister group to pinnipeds was a minor component of early hypotheses of pinniped monophyly (Wyss, 1987; Berta and Wyss, 1994), as the most recent common ancestor of pinnipeds was envisioned as ursid-like. *Enaliarctos*, the earliest-diverging pinnipedimorph known at the time, possesses many features which were thought to characterize ursids ancestrally, including a shelf-like P^4 protocone, a labiolingually-restricted M^1, and a deep lateral basioccipital embayment for the inferior petrosal sinus (Mitchell and Tedford, 1973; Flynn et al., 1988; Hunt and Barnes, 1994). However, Wolsan (1993) and Kohno (1993, 1994, 1996a) recovered musteloids as the sister group to pinnipeds, a pairing strongly supported by molecular data (**Figure 1C**; Flynn et al., 2005; Fulton and Strobeck, 2007; Sato et al., 2009, 2012; Doronina et al., 2015).

Puijila darwini, uncovered from lacustrine deposits in the High Arctic, was originally proposed to be a transitional pinniped (Rybczynski et al., 2009), filling in the gap between the fully-flippered *Enaliarctos* and terrestrial carnivorans. This hypothesis was based on a relatively small parsimony analysis, employing exclusively morphological characters, investigating the relationships of *Puijila* to terrestrial and semi-aquatic arctoids. However, this interpretation requires a more comprehensive phylogenetic analysis (Diedrich, 2011; Berta, 2012; Boessenecker and Churchill, 2013; Koretsky and Domning, 2014; Koretsky et al., 2016). In the initial analysis (Rybczynski et al., 2009), *Puijila* was recovered in a clade with *Enaliarctos, Potamotherium*, and *Amphicticeps*, thereby supporting the notion *Puijila* is either a pinniped, a pinnipedimorph, or an earlier-diverging member of this clade. However, an extensive sampling of crown-group pinnipeds was not included.

Many comparative analyses of carnivorans, including phylogenetic investigations, have excluded pinnipeds (Bininda-Emonds and Gittleman, 2000), with researchers concerned the extreme aquatic adaptations of pinnipeds may obscure functional patterns across terrestrial carnivorans. Moreover, analyses investigating the interrelationships of pinnipeds and other arctoids have been limited to extant/crown taxa. Morphological analyses have included only broadly-defined "Mustelidae," "Ursidae," and/or "Amphicyonidae" as outgroups (Kohno, 1993, 1994; Berta and Wyss, 1994). Assimilating in-group taxa into a hypothetical common ancestor is indeed more prudent than using exemplary taxa (Bininda-Emonds et al., 1998), but still requires many assumptions (e.g., monophyly), which can be easily violated by incompletely preserved extinct taxa.

Excepting Finarelli (2008), which included only a single pinniped species, phylogenetic analyses of fossil arctoids have omitted extant taxa (Wolsan, 1993; Wang et al., 2005). Such analyses have provided valuable insights into the nuanced differences between the stem-ward taxa of each clade. However, without molecular data constraining the topology, it is difficult to retrieve robustly-supported topologies and reconstruct ancestral nodes.

The present work synthesizes these detached lines of inquiry, which are made less incongruous by the inclusion of numerous extinct taxa from the respective arctoid families and subfamilies. A total-evidence tip-dating approach is employed to test the proposed allocation of *Puijila darwini* to the stem of Pinnipedia, and to investigate the evolutionary relationships of and patterns of character evolution associated with early-diverging pinnipeds. The addition of extinct arctoids, and a broad sampling of extant taxa, including both morphological and molecular data, allows for the reconstruction of ancestral nodes to be made from the phylogenetic method itself.

INSTITUTIONAL ABBREVIATIONS

AMNH, American Museum of Natural History, New York, USA; BSP, Bayerische Staatssammlung für Paleontologie und Historische Geologie, Munich, Germany; CMN, Canadian Museum of Nature, Ottawa, Canada; FMNH, Field Museum of Natural History, Chicago USA; FSL, Departement des Sciences de la Terre, Universit. Claude Bernard, Lyon, France; JODA, John Day Fossil Beds, Kimberly, USA; LACM, Natural History Museum of Los Angeles County, Los Angeles, USA; MGL, Musée Geologique Cantonal, Lausanne, Switzerland; MNHN, Institut de Paleontologie, Museum National d'Histoire Naturelle, Paris, France; NMB, Naturhistorisches Museum Basel, Basel, Switzerland; NMNS, National Museum of Nature and Science, Tsukuba, Japan; NUFV, Nunavut Fossil Vertebrate Collection (housed in CMN); ROM, Royal Ontario Museum, Ottawa, Canada; UOMNH, University of Oregon Museum of Natural History, Eugene, USA; USNM, Smithsonian Institution National Museum of Natural History; UWBM, University of Washington

Burke Museum, Seattle, USA; YPM, Yale Peabody Museum, New Haven, USA.

MATERIALS AND METHODS

Selection of Taxa

Appendix 1 in **Supplementary Material** lists the taxa coded for use in the analyses, and identifies which taxa are included in each analysis. Completeness was considered when selecting extinct taxa. For example, *Eotaria*, the earliest-known otariid (Boessenecker and Churchill, 2015; Velez-Juarbe, 2017), was not included, as it is known only from fragmentary mandibular and dental elements (10.29% scored). *Semantor*, a semi-aquatic arctoid with disputed phylogenetic affinities (Orlov, 1933), is similarly known from only post-cranial elements (14.70% scored overall), and was also excluded from the present analyses. A primary goal in the selection of extinct taxa was to sample early-diverging members of each clade (as identified in Wolsan, 1993; Berta and Wyss, 1994; Wang et al., 2005; Finarelli, 2008), following the methodological reasoning of Wang et al. (2005). A representative taxon of every known family within Arctoidea, extinct or extant, is included. See the **Supplementary Information** file for detailed explanations for each included taxon.

Canidae was identified as an ideal outgroup for this analysis, reflecting the sister group relationship between canids and arctoids (Tomiya and Tseng, 2016), and the availability of molecular data for the group. The earliest diverging canids are *Prohesperocyon* and *Hesperocyon*. The latter is well-represented in the fossil record, and does not display the specializations observed in later-diverging canids, so it is included in the present study. *Canis lupus* is included as the lone extant canid.

Total Evidence Dating

To date the phylogeny and coestimate the times and evolutionary rates, a "total-evidence dating" approach, was employed (Pyron, 2011; Ronquist et al., 2012b, 2016; Gavryushkina et al., 2016; Lee, 2016; Zhang et al., 2016; Wang et al., 2017). Also known as fossil-tip calibration, this approach allows for the simultaneous analysis of extant and extinct taxa, integrating known age ranges of extinct taxa (directly associated with fossil tips) to infer phylogeny, divergence times, and macroevolutionary parameters (Ronquist et al., 2012a). Furthermore, employing a fossilized birth-death model into a tip-dating analysis allows for the incorporation of information concerning speciation, extinction, and sampling processes (Zhang et al., 2016).

Molecular Data

Following Sato et al. (2009), nucleotide sequences of five nuclear coding genes were obtained—APOB (Apolipoprotein B), BRCA1 (Breast cancer 1, early onset), RAG1 (Recombination activating gene 1), IRBP (Interstitial retinol binding protein 3), VWF (Von Willebrand factor)—for a selection of extant arctoids (see Appendix 2 in **Supplementary Material**). The sequence data were downloaded from GenBank (NCBI, and accession numbers are listed in **Table S3**). Sequences were aligned using the MUSCLE alignment tool in Mesquite (Maddison and Maddison, 2015). To create identical lengths of each individual sequence for each taxon, sequences were manually trimmed in regions that were poorly aligned and/or displayed >50% gaps for a specific column. This ensures that when all five genes are analyzed together, homologous sites for each gene are aligned. Each aligned gene was then allocated to a discrete partition within the dataset. Data was partitioned into six units—Morphological, APOB, BRCA1, RAG1, IRBP, VWF. The Mk model was applied to the Morphological partition. To determine the evolutionary models of best fit for each data partition of molecular data, marginal likelihoods for each available model were calculated and compared separately for each partition using the stepping stone method (Xie et al., 2010), a path sampling method which obviates the controversial use of the harmonic mean method to estimate likelihood (Fan et al., 2010). **Table S2** lists the models and variants that were considered, and their marginal likelihood scores when subject to a Markov Chain Monte Carlo analysis of 5,500,000 generations and 50 steps. The models with the highest average marginal likelihood score (averaged between the two runs), were selected. The GTR + gamma model was selected for VWF, APOB, and BRCA1, and the GTR + invgamma model was selected for RAG1 and IRBP. The models-of-best-fit were later applied to their relevant data partitions in the Bayesian analyses.

Molecular data was included for at least one taxon from each extant caniform family—Canidae (*Canis lupus)*, Ursidae (*Ursus arctos)*, Mephitidae (*Mephitis mephitis)*, Ailuridae (*Ailurus fulgens)*, Procyonidae (*Procyon lotor),* Mustelidae *(Gulo gulo, Lontra canadiensis, Enhydra lutris, Neovison vison, Taxidea taxus)*, Odobenidae (*Odobenus rosmarus)*, Phocidae (*Cystophora cristata, Erignathus barbatus, Halichoerus grypus, Mirounga angustirostris, Monachus monachus, Phoca vitulina)*, and Otariidae (*Arctocephalus* sp., *Callorhinus ursinus, Eumetopias jubatus, Otaria byronia, Zalophus californianus)*. The inclusion of taxa was based on the availability of their selected genes in GenBank. Gaps and missing characters were treated as missing data. *Canis lupus* was specified as the outgroup.

Morphological Data

Eighty-two craniomandibular characters, fifty dental characters, and seventy-two postcranial characters were coded when possible for all taxa in the analysis. While some characters are novel, the majority of characters were derived from previous phylogenetic analyses of Carnivora, including those focusing on pinnipeds (Berta and Wyss, 1994; Kohno, 1996a,b, 2006; Deméré and Berta, 2001, 2002, 2005; Boessenecker and Churchill, 2013, 2015, 2018; Amson and de Muizon, 2014; Churchill et al., 2014; Furbish, 2015), musteloids (Bryant et al., 1993; Wolsan, 1993; Ahrens, 2014; Valenciano et al., 2016), arctoids (Tedford et al., 1994; Wang et al., 2005; Finarelli, 2008), and carnivorans more broadly (Wyss and Flynn, 1993; Wesley-Hunt and Flynn, 2005; Spaulding and Flynn, 2012; Tomiya and Tseng, 2016) (see Appendix 1 in **Supplementary Material** for full character list). Many of these characters were then modified, either by editing existing character states or by adding character states, to better reflect the diversity of these characters across arctoids. All changes are noted in the **Supplementary Information File**.

All taxa were specifically coded for this study (Appendix 2 in **Supplementary Material**).

Some phylogenetic analyses of pinnipeds (Deméré and Berta, 2002) have attempted to quantify and discretize characters that gauge nuanced differences in size and/or shape of morphological features, an approach followed in the present analysis. Further work on the morphological characters used in the study of extinct carnivorans should use statistical techniques to test the discretization or "binning" of different character states (Prieto-Marquez, 2010; Zou and Zhang, 2016).

A representative from every known arctoid family was coded, including members of the extinct Amphicyonidae, Oligobuninae and Desmatophocidae (**Supplementary Table 1** in **Supplementary Material**). When possible, adult male specimens were selected for character coding, due to the high sexual dimorphism displayed in many modern (Gittleman and Valkenburgh, 1997) and extinct arctoids (Hunt and Skolnick, 1996; Cullen et al., 2014).

To code their character states, specimens representing the extinct taxa were observed personally. Data for certain characters were also gathered from the literature if not discernible in the specimens available to us. Those concerning the internal cranial anatomy were inferred from video files of CT scans spanning the different axes of the cranium (from www.digimorph.org) and the literature (Ahrens, 2014; Geraads and Spassov, 2016; Grohé et al., 2016).

Analysis

The total-evidence tip-dating analysis was initially run in MrBayes v3.2.6 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003) for 10,000,000 generations, with a sampling frequency of 1,000 and a diagnostic frequency of 1,000. To ensure convergence upon similar results, two independent runs of MCMC (Markov-chain Monte Carlo) were specified for each analysis, with four chains (three hot and one cold) per run. Stationarity and convergence of the posterior probabilities were gauged using Tracer v1.6 (Rambaut, 2007) for OSX. Convergence of parameters was identified when the plot of the log likelihoods ("the Trace line") varied about a constant value. After accepting convergence, runs were considered sufficiently long when the effective sample sizes (ESS's) for each trace rose above 200. AWTY (Wilgenbusch, 2004) was offline as of this analysis, so no test of convergence of tree topologies was used, though analyses were only terminated when the average standard deviation of split frequencies was <0.01. Inspection of the trace plots also allowed us to identify the burn-in of each analysis, which was subsequently discarded. Following termination of the runs and discarding of the burn-in, the consensus tree showing all compatible clades was requested (contype = allcompat).

A relaxed clock model was implemented to account for evolutionary rate variation across branches and over time. The clock rate was associated with a lognormal prior with a mean of -7.0 and a standard deviation of 0.6 on the natural logarithm scale. The analysis specified an independent gamma rates model, and the variance increase parameter had an exponential prior with rate 10. The fossilized birth-death model was employed (Heath et al., 2014), with the following priors: speciationpr = exp(10), extinctionpr = beta(1,1), fossilizationpr = beta(1,1). The extant sample proportion was estimated to be 0.085. Treeagepr was set at 45, following Matzke and Wright (2016). The full script for this analysis is available in Appendix 3 in **Supplementary Material**.

Fossil stratigraphic ages, used to introduce time intervals for fossil samples, are listed in Appendix 4 in **Supplementary Material**. Fossil ages include ranges of uncertainty.

Ancestral character state reconstruction (ACSR) was performed for selected characters (those identified as possible pinniped or pinnipedimorph synapomorphies and homoplastic characters in this and other analyses) under the maximum likelihood approach using Mesquite v. 3.04. The Trace Character History function was applied to the phylogenetic tree recovered from total-evidence tip-dating, with the morphological characters used as the source of stored characters. The mK1 model was selected as the model of character state evolution, with equal probability for any character change. ACSR was used to identy the likely plesiomorphic condition for each character at the base of Pinnipedia, and to identify any possible reversals or instances of parallel trait development within the clade.

Parsimony Analysis of Combined Morphological-Molecular Dataset

Parsimony analysis was performed on the same dataset with the same outgroup using TNT 1.1 (Goloboff et al., 2008). One thousand replicates of random addition sequences were requested, followed by 500 iterations of tree bisection reconnection (TBR) and parsimony ratchet. Bootstrap support values (BS) were also obtained using TNT, using symmetric-resampling frequencies and 1,000 replicates.

RESULTS AND DISCUSSION

Total-Evidence Dating

Musteloids are identified as the sister group to pinnipeds (pp = 0.31), and *Puijila* and *Potamotherium* are identified as the sister clade to Pinnipedimorpha (pp = 0.89). Otariids and Odobenids are recovered as sister taxa (pp = 0.69), and desmatophocids are identified as the sister group to the otariid-odobenid coupling (pp = 0.75), to the exclusion of phocids. The present analysis indicates a divergence date between pinnipeds and musteloids in the Eocene, ranging from the Ypresian to the early Bartonian (95% highest posterior density interval 41.17–50.3 Ma) (**Figure 2**). Divergence time estimates for selected nodes are listed in **Table S3**.

Divergence time estimates within crown group Pinnipedia produced by the present analysis do not stray wildly from those posited by other analyses, and in fact recover a divergence between Phocidae and Otarioidea (∼26.94 Ma) roughly midway between those of Arnason et al. (2006) on the one hand, and Higdon et al. (2007) and Fulton and Strobeck (2010) on the other. The present analysis identifies an early Miocene (∼22.76) split between Otariidae and Odobenidae, in agreement with that of several other analyses (Higdon et al., 2007; Yonezawa et al., 2009; Nyakatura and Bininda-Emonds, 2012), while other analyses using node calibration have recovered a later divergence date inconsistent with the ages of the earliest

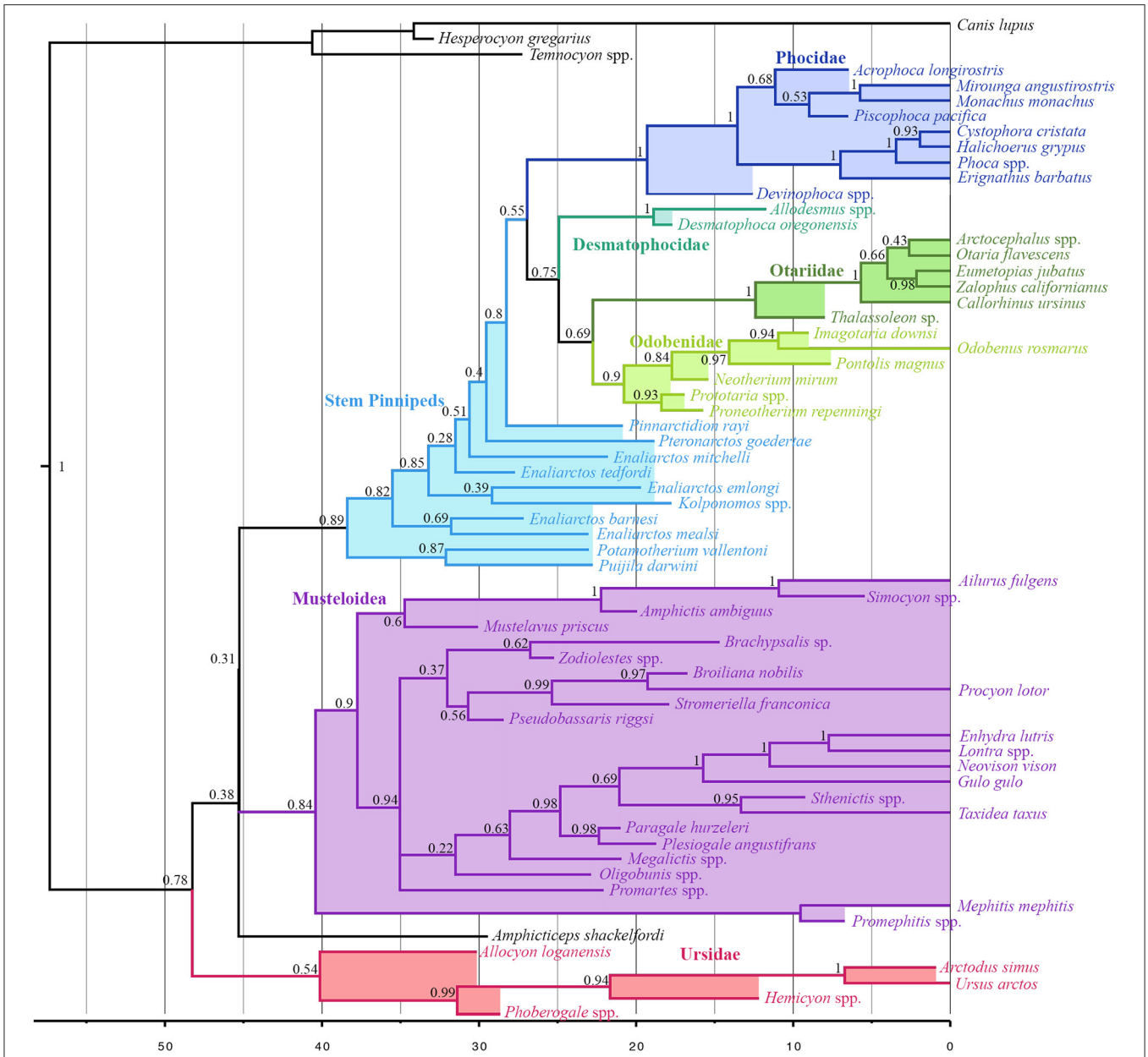


FIGURE 2 | All compatible clades consensus tree of Total-Evidence Tip-Dating analysis using Bayesian inference. Numbers at nodes indicate posterior probability values for that node.

odobenid, *Proneotherium* (Fulton and Strobeck, 2010). Beyond the crown pinniped radiation, the present investigation identifies older divergence times than many other analyses employing node- or tip-dating (excepting Arnason et al., 2006).

On the other side of the pinniped-musteloid split, the oldest incontestable pinniped fossils appear in the earliest Chattian (Berta, 1991; Prothero et al., 2001), far later than both the estimated early-to-mid Eocene split from musteloids, and the estimated ~38.41 Ma divergence time between *Puijila*/*Potamotherium* and all later-diverging pinnipedimorphs (Berta, 1991; Berta and Wyss, 1994). This long gap would allow plenty of time for the establishment of morphological characteristics reflecting the pronounced level of aquatic adaptation exhibited by *Enaliarctos* (Berta and Ray, 1990), ostensibly following a long initial phase of evolution in a non-marine setting represented by *Puijila* and *Potamotherium*, from lacustrine and fluvial deposits, respectively.

Pinnipedimorph Taxonomy

The Pinnipedimorpha was erected by Berta (1991) and includes *Enaliarctos* and all of its descendants, including a monophyletic crown-group Pinnipedia, which is strongly

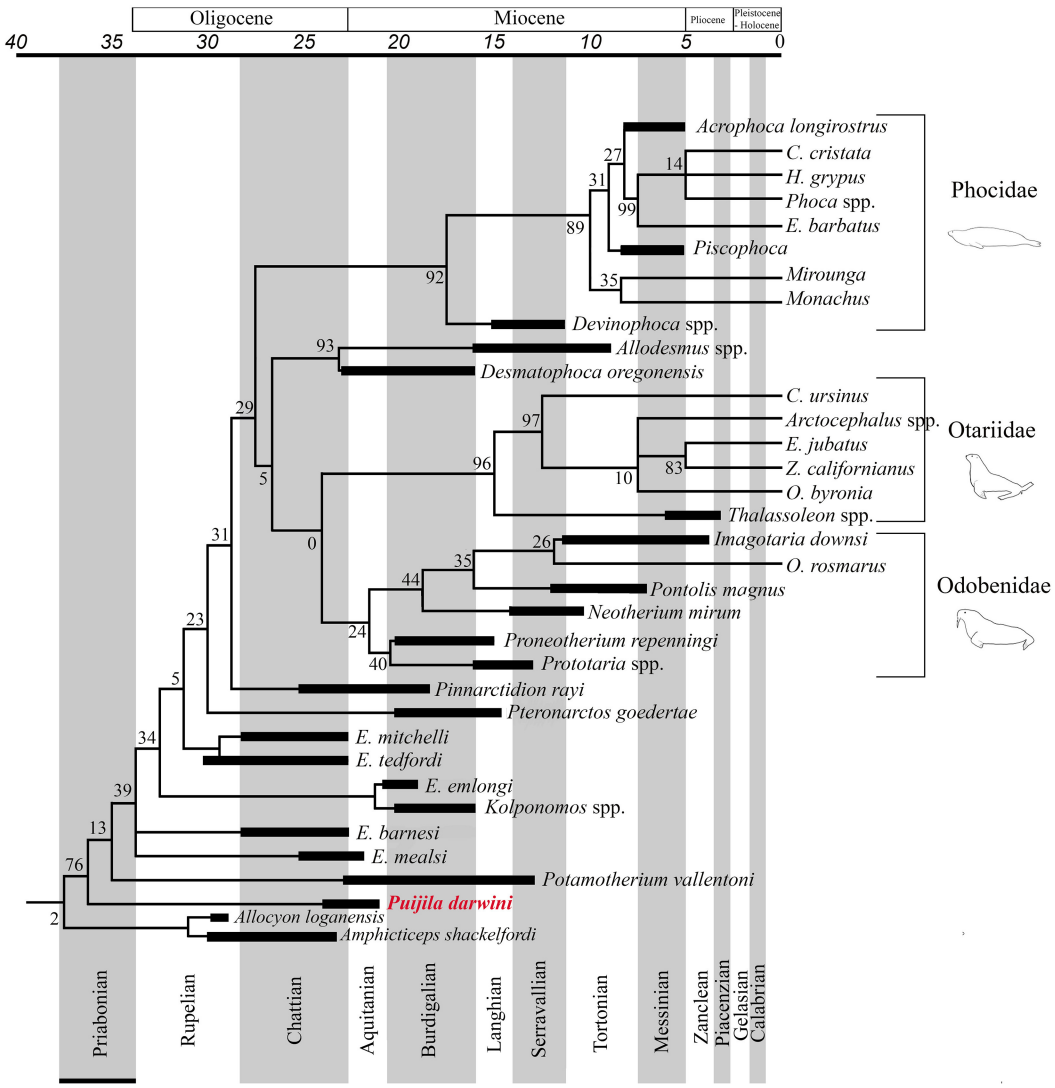


FIGURE 3 | Pinniped branch of the strict consensus tree from the parsimony analysis of the combined molecular and morphological dataset, with a geological time-scale. Black bars indicate age ranges (including range of uncertainty) of extinct taxa.

supported in all phylogenetic analyses performed in the present study. This study finds *Puijila* and *Potamotherium* to be members of this lineage that diverged prior to *Enaliarctos*, suggesting that a more inclusive taxonomic definition is warranted. All phylogenetic analyses in the current study recover a monophyletic group that includes crown group pinnipeds and the stem pinnipeds recognized here (*Puijila, Potamotherium, Kolponomos, Enaliarctos, Pteronarctos, Pinnarctidion*). A monophyletic Pinnipedia is restricted to crown group pinnipeds in both analyses. The reasonably strong support for a monophyletic *Puijila* + *Potamotherium* + Pinnipedimorpha in both phylogenetic analyses suggests *Puijila* and *Potamotherium* may confidently be identified as stem pinnipeds. *Amphicticeps shackelfordi* and *Allocyon loganensis*, a pair of taxa previously identified as possible early-diverging members of the pinniped divergence (Tedford et al., 1994; Rybczynski et al., 2009), are also recovered along the pinniped stem in the parsimony analysis (**Figure 3**), though this placement is very poorly supported (BS = 2).

Puijila and *Potamotherium* do not possess many of the conventional pinniped synapomorphies, but as potential stem pinnipedimorphs, they should not be expected to possess these features. The present Bayesian analysis suggest the features of modern pinnipeds arose sequentially, as observed in many other secondarily aquatic tetrapods (Hall, 1999), rather than as an integrated package.

Ten synapomorphies were identified for a clade of *Potamotherium* + *Puijila* + pinnipedimorphs (**Table 1**), eight of which are identified (>50% likelihood) at the base of the clade in the ACSR analysis. Unsurprisingly, it is the basicranial region that supports the phylogenetic affinities of *Puijila* and *Potamotherium*, as half of these possible synapomophies relate to

TABLE 1 | Synapomorphies of the clade *Puijila* + *Potamotherium* + Pinnipedimorpha, as identified in the parsimony analysis of morphological data.

Synapomorphies of *Puijila* + *Potamotherium* + Pinnipedimorpha (This study)
1. Postorbital constriction
2. Lacrimal, size
3. Postglenoid foramen
4. Pseudosylvian sulcus
5. Round window size
6. P^4/M^1 relative size
7. M^1, postprotocrista
8. M^2, metaconule
9. P_4 accessory cusps
10. M_1, metaconid mesiodistal position

Bolded characters are those that record >50% likelihood for the internode toward the base of Pinnipedia in the ACSR analysis employing maximum likelihood, as applied to the topology recovered from the Bayesian total-evidence tip-dating analysis.

the basicranial region. Unfortunately, data from the basicranial region is incompletely known from many fossil specimens, as the majority of basicranial characters are not preserved or are difficult to access. As these regions are of significant phylogenetic utility, conclusions on the precise phylogenetic relationships of early-diverging fossil pinnipeds may be reinterpreted once such data becomes available. Preservational biases have occasionally drastically misled identifications of fossil specimens (Donoghue and Purnell, 2009; Pattinson et al., 2014), and they can cause directional shifts in phylogenetic analyses (Sansom and Wills, 2013). Fossil specimens may appear to lack later derivations or specializations associated with more crown-ward clades, and thus, may erroneously be shifted toward the stem (Sansom, 2014).

Many previous phylogenetic analyses identifying *Potamotherium* as a mustelid have been limited in their observations to just a handful of specimens (Wyss, 1987; Bininda-Emonds and Russell, 1996; Wang et al., 2005; Finarelli, 2008) or relied on Savage (1957) thorough description of the taxon (Furbish, 2015). The pinnipedimorph affinities of *Potamotherium* become increasingly apparent upon inspection of additional undescribed material of *Potamotherium* from the collections at NMB, MNHN, and FSL. *Potamotherium* displays several polymorphic characters, with the derived states commonly representing pinnipedimorph synapomorphies. The variable appearance of these pinnipedimorph synapomorphies suggests a transitional morphology between terrestrial arctoids and pinnipedimorphs. These features include: presence of large antorbital process, mandible deepest anteriorly, reduction of the fossa for the tensor tympani, and presence of a palatal midline ridge.

Although the group Pinnipedimorpha is well-supported, the details of the evolution of the stem lineages are less well-resolved. For example, there is considerable cranial and dental diversity between different species of *Enaliarctos* (Cullen et al., 2014; the present study), and significant taxonomic revision of the genus, with comparison to other stem pinnipeds overlapping *Enaliarctos* in age, may be warranted. Despite superficially appearing less aquatically-specialized than *Potamotherium*, *Puijila* shares many traits with *Enaliarctos* to the exclusion of *Potamotherium*, though the similarities are not sufficient to recover a closer relationship between the two in our phylogenetic analyses. Many of these traits are also observed in otarioids in general, but not phocids, including: a thin, projecting tympanic crest, an inflated and somewhat rounded caudal entotympanic, and presence of an alisphenoid canal. Specifically, the tympanic crest and tympanic bulla of *Puijila* strongly resemble those of *Enaliarctos emlongi* and *Enaliarctos mealsi.* We also identified the shared presence of an intrabullar septum in *Puijila, Enaliarctos emlongi,* and *Enaliarctos mealsi*. Additional traits found in *Puijila, Enaliarctos,* and later-diverging pinnipedimorphs, to the exclusion of *Potamotherium*, include: absence of I_1, a reduced M_1 metaconid, presence of sharp ventral keel on the axis, and presence of a cylindrical lesser tuberosity of the humerus.

The presence/absence of the alisphenoid canal has been considered as a character of significant phylogenetic utility (Tedford, 1976; Koretsky et al., 2016). The alisphenoid canal carries multiple structures, including a branch of the external carotid artery, likely protecting it from occlusion during contraction of the pterygoideus muscle (Ewer, 1973). Accordingly, the alisphenoid canal may have been lost in taxa whose carotid canals lie closer to the orbit, so that the external carotid artery is largely unexposed to the adjacent musculature. This is borne out by mustelids and mephitids, who display a reduced post-canine dentition, and in phocids, who possess large orbits. While it appears unlikely an organism could have lost the alisphenoid canal and subsequently regained it, we cannot, a priori, rule such a scenario out. Thus, the character was not specified as irreversible in our analysis, though we do admit a reversal to the presence of the canal seems unlikely. Within the topology produced by the TEA, we can assume the loss of the alisphenoid canal represents an autapomophy for *Potamotherium*, rather than a synapomorphy of a hypothetical *Potamotherium* + Phocidae clade (de Muizon, 1982a; Koretsky and Rahmat, 2015). It is likely the alisphenoid canal was lost multiple times in arctoids, including at least once in phocids, mephitids, the clade containing mustelids + procyonids, and *Ailuropoda*, respectively.

Reappraisal of Pinniped Synapomorphies

Pinnipedia (Illiger, 1811) encompasses the last common ancestor of otarioids and phocids, and all of its descendants (Berta and Wyss, 1994). Wyss (1987) and Berta and Wyss (1994), suggested reversals were likely more common than convergence throughout the early evolutionary history of pinnipeds. This view was based on tree topologies in which Phocidae was recovered as the sister group to Odobenidae, and Ursidae was identified as the sister group to pinnipeds. Such a view has become less tenable as molecular analyses have revised our interpretation of these relationships. Many of these putative reversals may now be reinterpreted, in light of the present analysis. Reversals now appear less common than convergences within pinnipeds, as the ACSR analysis identifies over twice as many likely parallelisms

($n = 30$; interpreted as a trait that arises at least once in multiple crown group clades) as likely reversals ($n = 14$), and a few of these reversals are in fact identified as arising in parallel in multiple crown group clades (**Table S4**). This notion is also supported by the morphology of the early-diverging fossil representatives of the three extant pinniped families—*Prototaria* (or *Proneotherium*), *Eotaria,* and *Devinophoca*. In spite of their possession of odobenid, otariid, and phocid synapomorphies, respectively, these taxa also possess a suite of primitive characters, lacking some of the aquatic specializations previously used to define Pinnipedia.

In light of the recent discovery of fossils attributable to *Prototaria, Proneotherium, Eotaria,* and *Devinophoca*, it appears that previously cited pinniped synapomorphies are more variable in distribution among crown group pinnipeds than previously believed, and may not define the base of Pinnipedia. **Table 2** lists all previously proposed synapomorphies of crown pinnipeds (Pinnipedia), and notes if the early-diverging fossil crown members of each pinniped family display the ancestral or derived character state, and the likelihood the synapomorphy is present at the identified at the base of pinnipedia based on the ACSR of the tree derived from the total-evidence tip-dating analysis. The ancestral state is observed at least once in a crown member for 11 of the 16 characters, and seven of 16 of these features are identified as unlikely to characterize the base of Pinnipedia using ACSR. The five characters that do not display retention of an ancestral state are postcranial characters that are poorly sampled for these extinct taxa, many of which are only known from skulls, mandibles, and/or a few postcranial elements. The appearance of an ancestral state in any of these fossil forms could theoretically represent an anomalous derivation or reversal. However, if early-diverging representatives from more than one family retain the ancestral condition, then it is equally parsimonious to assume the derived condition did not characterize the base of Pinnipedia. As **Table 2** demonstrates, this latter scenario describes six of the 16 characters. It must be assumed these numbers would increase if there were less missing data (currently only 44/80 total possible codings).

The following features have been used to delineate a monophyletic Pinnipedia to the exclusion of the stem "pinnipedimorphs" *Enaliarctos*, *Pteronarctos*, and *Pacificotaria*: pit for tensor tympani absent, I^3 lingual cingulum absent, M_{1-2} trigonid suppressed, nasolabialis fossa absent, antorbital process large, P^4 protocone shelf absent, P^4 one- or two-rooted, and M_2 absent (Berta and Wyss, 1994).

The absence of the pit for the tensor tympani does appear to be a synapomorphy of crown pinnipeds (Wyss, 1987; Berta and Wyss, 1994), as the tensor tympani appears to insert on the Eustachian canal in all extant pinnipeds sampled. However, Wesley-Hunt and Flynn (2005) identified a transitional state in *Pteronarctos*, in which the fossa for the tensor tympani is present, but only shallowly excavated into the promontory. Likewise, the fossa for the tensor tympani is greatly reduced in *Puijila*, several specimens of *Potamotherium* (Mitchell and Tedford, 1973; Pers. Obs.), and multiple species of *Enaliarctos* (Berta, 1991). CT data of the inner ear of early-diverging otariids, phocids, and odobenids should clarify the timing of this transition to determine if the fossa for the tensor tympani was lost multiple times.

Lingual cingulum of the I^3 is retained in *Prototaria* (Kohno, 1994; Kohno et al., 1995) and *Thalassoleon* (Deméré and Berta, 2005). An unreduced M_1 metaconid or trigonid is retained in *Eotaria* (Boessenecker and Churchill, 2013). The nasolabialis

TABLE 2 | List of previously proposed synapomorphies of crown group Pinnipedia (Berta and Wyss, 1994), and their occurrence in select fossil pinnipedimorphs.

Previously proposed pinniped synapomorphies	*Enaliarctos*	*Prototaria*	*Proneotherium*	*Thalassoleon*	*Eotaria*	*Devinophoca*	*% likelihood trait appears at base of Pinnipedia (ACSR)*
Loss of pit for tensor tympani	0	?	0	?	?	?	96.68%
I^3 Lingual Cingulum	0	0	1	0	?	?	76.71%
Reduced m1 metaconid	0	?	1	?	0	?	20.63%
(12)Nasolabialis Fossa Absence	0	0	0	0	?	1	4.01%
(13)Fossa muscularis absence	0	0	1	1	?	1	72.29%
Antorbital Process Large	0	1	1	1	?	0	<0.1%
P^4 Protocone Shelf	0	0	0	1	?	0	9.11%
P^4 1- or 2-rooted	0	0	0	1	?	0	1.14%
m2 absence	0	0	0	1	0	1	1.08%
Five Lumbar Vertebrae	0	?	?	?	?	?	51.71%
Olecranon flattened and expanded	0	?	?	1	?	1	86.26%
Radius, expanded distal half	0	?	?	1	?	?	99.39%
Pubic symphysis unfused	0	?	?	1	?	?	87.91%
Fovea for teres femoris	0	?	0	1	?	1	26.40%
Greater femoral trochanter larger and flattened	0	?	1	1	?	1	97.60%
Tibia-Fibula Fusion	0	?	0	0	?	?	99.96%

fossa appears to have been lost multiple times in pinnipeds, as it is present in a stem otariid (*Thalassoleon*), multiple fossil odobenids (*Proneotherium* and *Imagotaria*), and a possible desmatophocid (*Pinnarctidion*). Polarity for a large antorbital process must now be reversed, and the character can be conceived of as a possible synapomorphy for otarioids, with further derivation in both otariids and odobenids. A protocone shelf of the P^4 is retained in *Devinophoca* and *Proneotherium* (Churchill and Clementz, 2016). Reduction of the number of roots of P^4 (<3) appears to be a synapomorphy of pinnipeds, however, *Devinophoca* may display a transitional stage in which the posterior root is bilobed, though the roots are nevertheless well-merged (Koretsky and Holec, 2002). A similar P^4 root morphology is observed in *Proneotherium* and *Neotherium.* The M_2 is retained in *Pontolis* (Deméré, 1994), *Eotaria* (Boessenecker and Churchill, 2013), and even atavistically reappears in extant otariids (Drehmer et al., 2004).

Of these proposed synapomorphies, only absence of the fossa for the tensor tympani and reduction of P^4 roots appear to definitively define the node at the base of Pinnipedia. Berta and Wyss (1994) also listed a set of ambiguous postcranial synapomorphies of Pinnipedia, though their presence or absence in *Pteronarctos* and several species of *Enaliarctos* cannot be asserted without postcranial remains. These are: five lumbar vertebrae present, flattened and posteriorly expanded olecranon process, expanded distal end of radius, pubic symphysis unfused, loss of fovea teres femoris, greater trochanter of femur lost. These synapomorphies have not been discovered in pinnipedimorphs, nor have the plesiomorphic conditions been identified in any fossil pinnipeds. However, some of these are approached by *Potamotherium* and *Enaliarctos*, who display somewhat expanded distal radii, somewhat expanded greater trochanters, and a reduced fovea teres femoris. Unsurprisingly, many of these features are thought to be related to aquatic adaptation—general flattening of the long-bones, posterior expansion of the olecranon process, loss of the fovea teres femoris—as they are observed homoplastically in other clades of aquatic mammals (Gingerich et al., 1990, 1994).

Another more inclusive clade than the Pinnipedimorpha is also recognized—Pinnipediformes (Berta, 1994). This group encompasses *Pteronarctos* and Pinnipedia, to the exclusion of *Enaliarctos* (Berta, 1994). Four additional unambiguous synapomorphies were identified for Pinnipediformes: absence of lacrimal-jugal contact, maxilla contributes heavily to orbit, embrasure pit of P^4-M^1 shallow or absent, mastoid process in close proximity to paroccipital process and connected by a ridge.

The extent of the lacrimal is difficult to ascertain in fossil pinnipeds and their ancestors, as the lacrimal becomes well-fused to surrounding bones in ontogeny. It appears likely this is a pinniped synapomorphy. The discovery of juvenile specimens displaying the presence or absence of a lacrimal in *Puijila* or *Enaliarctos* could confirm or undermine this contention. The maxillary contribution to the orbit is also likely a pinniped or pinnipediforme synapomorphy, but fusion of these elements in *Enaliarctos* and *Puijila* precludes certain determination. While the embrasure pit of *Proneotherium* is reduced (Boessenecker and Churchill, 2013), it is nevertheless present, while all extant pinnipeds observed lack an embrasure pit outright. In phocids, the mastoid and paroccipital processes are widely separated, and the polarity of this character should be reversed, so it is considered a synapomorphy of otarioids.

Another proposed synapomorphy of Pinnipedia is the presence of prominent orbital vacuities (Wyss, 1987; Berta, 1991), which are unossified spaces on the medial orbital wall (Wyss, 1987). The presence of orbital vacuities appears to be unique to crown pinnipeds. The orbital vacuities are similar in phocids and otariids, but are placed far posteriorly in *Odobenus* (Kohno, 2006), though fossil walruses do possess a pinniped-like orbital vacuity (Deméré and Berta, 2001). Though similarly placed, the otariid orbital vacuity is of a different configuration than the phocid orbital vacuity (Wozencraft, 1989; Bininda-Emonds and Russell, 1996), and for that reason, Berta and Wyss (1994) speculate on an independent origin of orbital vacuities in the three pinniped families, a hypothesis supported by the present phylogenetic analysis. Orbital vacuities are absent in *Pinnarctidion*, present in *Allodesmus* (Berta and Wyss, 1994) and possibly incipient in *Desmatophoca* (Deméré and Berta, 2002).

Additional ambiguous synapomorphies of crown pinnipeds of Berta and Wyss (1994) include: enlarged tuberosities of the humerus, and a flattened and enlarged greater trochanter. A somewhat enlarged, though unflattened greater trochanter of the femur is observed in *Thalassoleon.* A somewhat enlarged lesser tubercle of the humerus that continues down the shaft as a cylindrical ridge, is observed in *Puijila, Semantor*, and the desmatophocids.

The earliest-diverging otariid (*Thalassoleon*) and phocid (*Devinophoca*) included in the analysis both display decidedly derived morphologies (Deméré and Berta, 2005; Koretsky et al., 2016). Yet, several character states previously believed to have been lost at the base of crown Pinnipedia are retained even in these taxa (**Table 2**). The earliest-diverging odobenid (Kohno, 1994; Boessenecker and Churchill, 2013) displays only a single purported pinniped synapomorphy of Berta and Wyss (1994)—presence of the antorbital process—a character which is now identified as an otarioid synapomorphy, and the specific configuration of this character in *Prototaria* is more appropriately considered an odobenid synapomorphy (Berta and Wyss, 1994). Deméré and Berta (2002) commented on the lack of the fovea for the teres femoris in *Proneotherium*, suggesting it would be expected given its transitional morphology. However, as this was offered as a pinniped synapomorphy, it should theoretically be present in *Proneotherium*. It would not be unexpected for a transitional odobenid to lack odobenid synapomorphies, but it would be unconceivable to observe a transitional odobenid lacking several pinniped synapomorphies, considering how few synapomorphies are identified (Berta and Wyss, 1994).

From the results of the present study, we reject many of the synapomorphies previously used to diagnose Pinnipedia, as many were founded on hypotheses of sister group relationships of Pinnipedia-Ursidae and Phocidae-Odobenidae. The present analysis recognizes a new suite of possible synapomorphies that support a monophyletic crown Pinnipedia, to the exclusion

TABLE 3 | Synapomorphies of the clade Pinnipedia, as identified in the parsimony analysis of morphological data using TNT.

1. Fossa muscularis
2. Nasals in dorsal view
3. Premaxilla-nasal contact
4. Embayment for the inferior petrosal sinus
5. Auditory bullae, pseudoseptae
6. P^4, protocone
7. P/M relative size
8. P^4, metastyle length
9. M^1, metacone height
10. M^1, length of lingual half
11. M^2, size
12. Humerus, supinator ridge extent/size
13. Humerus, greater tubercle extent/size
14. Femur, ligamentum teres femoris presence
15. Femur, trochanteric fossa

Bolded characters are those that record >50% likelihood for the internode toward the base of Pinnipedia in the ACSR analysis employing maximum likelihood, as applied to the topology recovered from the Bayesian total-evidence tip-dating analysis.

of *Puijila, Potamotherium, Enaliarctos,* and *Pteronarctos*. Admittedly, many of these possible synapomorphies cannot, at the present time, be identified in early-diverging members of the crown families due to inadequate preservation, and may turn out to represent independent derivations. Also, several of these synapomorphies require reversals to the ancestral condition in *Prototaria*, *Proneotherium, Thalassoleon,* and/or *Devinophoca* (synapomorphies 1, 3, 4, 6, 8, 10, 11, 12, 14 in **Table 3**), lessening the confidence with which we can assign them to the base of Pinnipedia. Furthermore, as with those identified by Berta and Wyss (1994), many of the synapomorphies identified here are uniquely approached by stem pinnipedimorphs, among arctoids. The stem pinnipedimorphs identified in the analysis appear to display a transitional or incipient state for many of these characters (synapomorphies 7, 8, 11, and 13 in **Table 3**), suggesting selective pressures favoring those traits were already present in stem pinnipedimorphs.

Thus, while the present analysis identifies a suite of potential pinniped synapomorphies, an investigation of their distribution among extinct taxa indicates that several of these features appear to be incipiently present in stem pinnipeds and others may have arisen in parallel among the three extant pinniped families. In the absence of incontestable transitional fossils, a more thorough investigation into the alpha and beta taxonomy of these known early-diverging taxa may provide additional clues regarding the divergence of otariids and phocids.

Review of Parallel Evolution and Swimming Specialization in Arctoids

Swimming has evolved multiple times in arctoids (Berta et al., 2005). While no other known arctoids have developed the remarkable swimming specializations observed in pinnipedimorphs, many other taxa are considered to be semi-aquatic, including otters, the American mink, the European mink, the polar bear, the robust otter-like fossil musteloid *Mionictis* (Baskin, 1998), and, if its affinities lie outside the pinniped divergence, *Kolponomos*. Otters, minks and *Mionictis* all share with *Puijila* and *Potamotherium* a similar ectomorph bodyplan and several other features related to swimming and aquatic feeding, including webbed digits (Savage, 1957; Berta et al., 2005; Rybczynski et al., 2009). These features are so numerous that they have even previously been used to assert the lutrine affinities of *Mionictis* (known from the middle Miocene) (Baskin, 1998; Tseng et al., 2009) and *Potamotherium*, though neither of these appear possible in the light of divergence time estimate studies consistently identifying a late Miocene divergence of Lutrinae (Sato et al., 2009). Considering these arctoid taxa converged upon similar specializations related to swimming, perhaps similar shifts could have begot flippers independently in otariids, phocids, and odobenids.

Adaptive radiations were thought to characterize the stem of pinnipedimorphs (Simpson, 1944), but recent evidence, based on tests of cranial disparity (Jones et al., 2015) and molecular markers (Higdon et al., 2007), suggest rapid radiations did not occur until the otarioid-phocoid split. Subsequent adaptive radiations appear to have taken place separately as otariids and odobenids diverged. This scenario is supported by the primitive morphologies of the early-diverging members of each extant pinniped family. Most notably, the early diverging odobenids *Prototaria* and *Proneotherium* do not depart significantly, in terms of morphology, from *Enaliarctos* and *Pteronarctos*. However, the results of the total-evidence tip-dating analysis indicate relatively rapid divergence events occurred along the stem, prior to a lull in identified speciations as the clades in crown Pinnipedia began to diverge. This high level of alpha and beta diversity is also reflected in the wide geographic distribution of stem pinnipeds.

Other molecular work has found evidence of genes that appear correlated with aquatic adaptation more broadly across marine mammals (Wang et al., 2008; Zhou et al., 2015), and of parallel shifts accelerating evolutionary rates of genes in multiple marine mammal clades (Chikina et al., 2016). However, convergent evolution of genes has been difficult to detect across marine mammal clades. Parallel substitutions of functionally-enriched genes do not appear to occur more frequently between marine mammals than they do between marine and terrestrial mammals (Irwin and Árnason, 1994; Foote et al., 2015; Zhou et al., 2015). While convergent molecular evolution and convergent phenotypic evolution are both common phenomena, adaptive molecular convergence associated with phenotypic convergence is uncommon (Foote et al., 2015).

Though our phylogenetic analyses identify a number of possible synapomorphies of crown pinnipeds, it is not possible to ascertain whether or not these shared traits arose via common ancestry or due to similar selective pressures on closely-related groups of organisms. Modern cladistics methods are not infallible. While more reliable than phenetic interpretations, phylogenetic inference of morphological data remains imprecise, and thus, susceptible to systematic error. This limits the confidence we can have in topological placements from such studies, including the present study.

The Mk model has monopolized Bayesian Inference of discrete morphological data. While it has been shown to be effective in simple data sets, it may not be able to accurately model morphological evolution. The Mk model employs a continuous-time Markov process running over finite state spaces (Klopfstein et al., 2015). A Markov process is conditioned only on the active state of the process. The past and future of the process are independent of the current state. Furthermore, the Markov process is also assumed to be at stationarity and time-reversible. A process attains stationarity when the current state ceases being dependent on the starting state. A process is time-reversible when its stochasticity does not become inconsistent or ill-defined if time is reversed.

The stationarity and time-reversibility of this Markov process prevents the Mk model from accounting for directional evolution (Klopfstein et al., 2015). Directional evolution, one of the three modes of selection originally proposed by Darwin (1859), involves an extreme phenotype being favored. Such selection is believed to be strong when an organism must adapt to a drastic and sudden shift in ecological pressures. Conceptually, directional selection would be a powerful propeller for an organism beginning to adapt to an aquatic existence. Conversely, in organisms that are secondarily aquatic, or highly specialized in some other manner, it should be unlikely, though not impossible, for organisms to revert to a less specialized form.

Since the Mk model does not allow for the input of assumptions of directional selection, it is unlikely that parallel evolution would reveal itself in a topology using the Mk model. A priori assumptions of directional selection are indeed difficult to integrate into an evolutionary model, but successful work has been carried out on characterizing directional selection in molecular datasets (Merritt and Quattro, 2001; Creevey and McInerney, 2004; Nielsen, 2005; Kosiol et al., 2008; McClellan, 2013; Enard et al., 2014; Bloom, 2017), allowing directional selection to be integrated into models of molecular evolution (Huelsenbeck et al., 2006; Ronquist et al., 2012b). Unfortunately, the non-objective nature of morphological characters and datasets, as outlined earlier in this chapter, makes it difficult to characterize directional selection as it relates to adaptive phenotypes. The extent of parallel evolution in pinnipeds may thus be greater than that suggested by the present study and others. It is possible that the extreme aquatic specializations characterizing extant pinnipeds, including flippers, developed in parallel in otariids, odobenids, and phocids, but our current phylogenetic methods are not capable of detecting such an event.

CONCLUSION

Pinnipedimorpha, as currently defined, includes the most recent common ancestor of *Enaliarctos* and all of its descendents. The phylogenetic analyses of the present study provide support for the placement of additional fossil arctoid taxa at the base of this group. These "proto" pinnipeds come from disparate locations, deposits, and ages, and display a wide range of aquatic adaptations, ranging from the large-bodied molluscivore *Kolponomos* (Stirton, 1960), to the freshwater otter-like forms *Puijila* and *Potamotherium*, to the aquatically-specialized *Semantor*. More detailed study of additional fossil arctoid taxa (including *Amphicticeps*), particularly with regards to their internal cranial architecture, may further explain the complexities associated with the transition of pinnipeds from a terrestrial to an aquatic environment.

At present, evidence overwhelmingly favors a monophyletic origin of pinnipeds. However, otariids, odobenids, and phocids display a startling amount of parallel evolution, as many of their shared features are absent in the early-diverging fossil ancestors of each family. Within a monophyletic framework, we postulate that parallel evolution may be the mechanism explaining their specialization within the aquatic realm, particularly with regards to raptorial feeding and hydrodynamic locomotion. A definitive answer awaits fossil evidence and the advancement of morphological phylogenetic methods. In the mean-time, statistical tests of convergence (Muschick et al., 2012; Ingram and Mahler, 2013; Arbuckle et al., 2014) could be applied to cladistic data sets of arctoids, to further examine the likelihood of parallel evolution within pinnipeds.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the **Appendix 3**.

AUTHOR CONTRIBUTIONS

NR and NK conceived of the study. RP and NK collected data for the phylogenetic analysis. RP and HM ran the analyses. RP wrote the manuscript in consultation with NR, HM, and NK.

FUNDING

This research was supported by a Burke-VPCSG, a Paleontological Society student research grant, and an NSERC CGS-M to RP, and a NSERC discovery grant to HM JSPS-ESA11740286 to NK.

ACKNOWLEDGMENTS

We are grateful to the staff of the Canadian Museum of Nature and the Government of Nunavut for allowing us to study NUFV-405. For assistance and access to fossil collections, we thank: W. Simpson (FMNH), J. Galkin (AMNH), C. Sidor, R. Eng, M. Riven, J. Bradley (Burke), F. Mayer (BNHM), D. Bohaska (NMNH), N. Famoso, G Retallack, G. Davis (UOMNH), J. Samuels (JODA), A. Pritchard, J. Gauthier (YPM), K. Seymour, D. C. Evans, B. Lim, J. Miller (ROM), C. de Muizon, G. Billet (MNHN), R. Marchant (MGL), G. Roessner (BSP), and E. Robert (FSL). For useful discussion on phylogenetics and taxonomy, and their encouragement to tackle this research question, we thank Annalisa Berta, Thomas Cullen, Susumu Tomiya, Takahiro Yonezawa, Alberto Valenciano, and John Flynn. For providing morphological information of Amphicticeps and other basal

musteloids and amphicynodontids (including useful casts) since the initial stage of this study, we thank Xiaoming Wang (LACM). We would also like to thank Morgan Churchill and Olaf Bininda-Emonds for comments that greatly improved the focus and scope of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2019.00457/full#supplementary-material

REFERENCES

Ahrens, H. E. (2014). Morphometric study of phylogenetic and ecologic signals in procyonid (Mammalia: Carnivora) endocasts. *Anat. Rec.* 297, 2318–2330. doi: 10.1002/ar.22996

Amson, E., and de Muizon, C. (2014). A new durophagous phocid (Mammalia: Carnivora) from the late Neogene of Peru and considerations on monachine seals phylogeny. *J. Syst. Palaeontol.* 12, 523–548. doi: 10.1080/14772019.2013.799610

Arbuckle, K., Bennett, C. M., and Speed, M. P. (2014). A simple measure of the strength of convergent evolution. *Methods Ecol. Evol.* 5, 685–693. doi: 10.1111/2041-210X.12195

Arnason, U. (1974). Comparative chromosome studies in Pinnipedia. *Heredita* 76, 179–225. doi: 10.1111/j.1601-5223.1974.tb01340.x

Arnason, Ú., Gullberg, A., Janke, A., Kullberg, M., Lehman, N., Petrov, E. A., et al. (2006). Pinniped phylogeny and a new hypothesis for their origin and dispersal. *Mol. Phylogen. Evol.* 41, 345–354. doi: 10.1016/j.ympev.2006.05.022

Baskin, J. (1998). "Mustelidae," in *Evolution of Tertiary Mammals of North America*, Vol. 1, eds C. M. Janis, K. M. Scott, and L.L. Jacobs (Cambridge: Cambridge University Press), 152–173.

Berta, A. (1991). *New Enaliarctos (Pinnipedimorpha) From the Oligocene and Miocene of Oregon and role of "Enaliarctids" in Pinniped phylogeny*. *Smithsonian Contributions to Paleontology*. Washington: Smithsonian Institution Press.

Berta, A. (1994). New specimens of the Pinnipediform *Pteronarctos* from the Miocene of Oregon. *Smithson. Contrib. Paleobiol.* 78, 1–30.

Berta, A. (2012). *Return to the Sea: The Life and Evolutionary Times of Marine Mammals*. Berkeley: Univ of California Press.

Berta, A., and Ray, C. E. (1990). Skeletal morphology and locomotor capabilities of the archaic pinniped *Enaliarctos mealsi*. *J. Vertebr. Paleontol.* 10, 141–157. doi: 10.1080/02724634.1990.10011803

Berta, A., Ray, C. E., and Wyss, A. R. (1989). Skeleton of the oldest known pinniped, *Enaliarctos mealsi*. *Science* 244, 60–62. doi: 10.1126/science.244.4900.60

Berta, A., Sumich, J. L., and Kovacs, K. M. (2005). *Marine Mammals: Evolutionary Biology*. Elsevier.

Berta, A., and Wyss, A. R. (1994). "Pinniped phylogeny," *Proceedings of the San Diego Society of Natural History*, Vol. 29:33–56.

Bininda-Emonds, O. R., and Gittleman, J. L. (2000). Are pinnipeds functionally different from fissiped carnivores? The importance of phylogenetic comparative analyses. *Evolution* 54, 1011–1023. doi: 10.1111/j.0014-3820.2000.tb00100.x

Bininda-Emonds, O. R. P., Bryant, H. N., and Russell, A. P. (1998). Supraspecific taxa as terminals in cladistic analysis: implicit assumptions of monophyly and a comparison of methods. *Biol. J. Linn. Soc.* 64, 101–133. doi: 10.1111/j.1095-8312.1998.tb01536.x

Bininda-Emonds, O. R. P., Gittleman, J. L., and Purvis, A. (1999). Building large trees by combining phylogenetic information: a complete phylogeny of the extant Carnivora (Mammalia). *Biol. Rev.* 74, 143–175. doi: 10.1017/S0006323199005307

Bininda-Emonds, O. R. P., and Russell, A. P. (1996). *Morphological Perspective on the Phylogenetic Relationships of the Extand Phocid Seals (Mammalia: Carnivora: Phocidae)*. Bonn: Zoologisches Forschungsinstitut.

Bloom, J. D. (2017). Identification of positive selection in genes is greatly improved by using experimentally informed site-specific models. *Biol. Dir.* 12:1. doi: 10.1186/s13062-016-0172-z

Boessenecker, R. W., and Churchill, M. (2013). A reevaluation of the morphology, paleoecology, and phylogenetic relationships of the enigmatic walrus *Pelagiarctos*. *PloS ONE* 8:e54311. doi: 10.1371/journal.pone.0054311

Boessenecker, R. W., and Churchill, M. (2015). The oldest known fur seal. *Biol. Lett.* 11:20140835. doi: 10.1098/rsbl.2014.0835

Boessenecker, R. W., and Churchill, M. (2018). The last of the desmatophocid seals: a new species of *Allodesmus* from the upper Miocene of Washington, USA, and a revision of the taxonomy of Desmatophocidae. *Zool. J. Linn. Soc.* 184, 211–235. doi: 10.1093/zoolinnean/zlx098

Bryant, H. N., Russell, A. P., and Fitch, W. D. (1993). Phylogenetic relationships within the extant Mustelidae (Carnivora): appraisal of the cladistic status of the Simpsonian subfamilies. *Zool. J. Linn. Soc.* 108, 301–334. doi: 10.1111/j.1096-3642.1993.tb00301.x

Chikina, M., Robinson, J. D., and Clark, N. L. (2016). Hundreds of genes experienced convergent shifts in selective pressure in marine mammals. *Mol. Biol. Evol.* 33, 2182–2192. doi: 10.1093/molbev/msw112

Churchill, M., Boessenecker, R., and Clementz, M. T. (2014). Colonization of the southern hemisphere by fur seals and sea lions (Carnivora: Otariidae) revealed by combined evidence phylogenetic and Bayesian biogeographical analysis. *Zool. J. Linn. Soc.* 172, 200–225. doi: 10.1111/zoj.12163

Churchill, M., and Clementz, M. T. (2016). The evolution of aquatic feeding in seals: insights from *Enaliarctos* (Carnivora: Pinnipedimorpha), the oldest known seal. *J. Evol. Biol.* 29, 319–334. doi: 10.1111/jeb.12783

Creevey, C. J., and McInerney, J. O. (2004). Clann: investigating phylogenetic information through supertree analyses. *Bioinformatics* 21, 390–392. doi: 10.1093/bioinformatics/bti020

Cullen, T. M., Fraser, D., Rybczynski, N., and Schröder-Adams, C. (2014). Early evolution of sexual dimorphism and polygyny in Pinnipedia. *Evolution* 68, 1469–1484. doi: 10.1111/evo.12360

Darwin, C. (1859). *On the Origin of Species by Means of Natural Selection Or the Preservation of Favoured Races in the Struggle for Life*, ed H. Milford. Oxford: Oxford University Press.

Davies, J. L. (1958). The Pinnipedia: an essay in zoogeography. *Geograph. Rev.* 48, 474–493. doi: 10.2307/211670

de Jong, W. W., and Goodman, M. (1982). Mammalian phylogeny studied by sequence-analysis of the eye lens protein alpha-crystallin. *Z Saugetierkd Int. J. Mammal. Biol.* 47, 257–276.

de Muizon, C. (1982a). Les relations phylogénétiques des Lutrinae (Mustelidae, Mammalia). *Geobios* 15, 259–277. doi: 10.1016/S0016-6995(82)80118-6

de Muizon, C. (1982b). Phocid phylogeny and dispersal. *Ann. S. Afr. Mus.* 89, 172–213.

Deméré, T. A. (1994). "Two new species of fossil walruses (Pinnipedia: Odobenidae) from the upper Pliocene San Diego Formation, California," in *Proceedings of the San Diego Society of Natural History*, Vol. 29. 77–98.

Deméré, T. A., and Berta, A. (2001). A reevaluation of *Proneotherium repenningi* from the Miocene Astoria Formation of Oregon and its position as a basal odobenid. *J. Vertebr. Paleontol.* 21, 279–310. doi: 10.1671/0272-4634(2001)021[0279:AROPRF]2.0.CO;2

Deméré, T. A., and Berta, A. (2002). "The Miocene pinniped Desmatophoca oregonensis Condon, 1906 (Mammalia: Carnivora) from the Astoria Formation, Oregon. Later Cenozoic Mammals of Land and Sea: tributes to the Career of Clayton E. Ray. *Smithson. Contrib. Paleobiol.* 93, 113–147.

Deméré, T. A., and Berta, A. (2005). New skeletal material of *Thalassoleon* (Otariidae: Pinnipedia) from the Late Miocene-Early Pliocene (Hemiphillian) of California. *Bull. Florida Mus. Nat. Hist.* 45, 379–411.

Deméré, T. A., Berta, A., and Adam, P. J. (2003). Pinnipedimorph evolutionary biogeography. *Bull. Am. Mus. Nat. Hist.*279, 33–76. doi: 10.1206/0003-0090(2003)279<0032:C>2.0.CO;2

Diedrich, C. (2011). The world's oldest fossil seal record. *Nat. Sci.* 3, 914–920. doi: 10.4236/ns.2011.311117

Donoghue, P. C., and Purnell, M. A. (2009). Distinguishing heat from light in debate over controversial fossils. *BioEssays* 31, 178–189. doi: 10.1002/bies.200800128

Doronina, L., Churakov, G., Shi, J., Brosius, J., Baertsch, R., Clawson, H., et al. (2015). Exploring massive incomplete lineage sorting in arctoids (Laurasiatheria, Carnivora). *Mol. Biol. Evol.* 32, 3194–3204. doi: 10.1093/molbev/msv188

Dragoo, J. W., and Honeycutt, R. L. (1997). Systematics of mustelid-like carnivores. *J. Mammal.* 78, 426–443. doi: 10.2307/1382896

Drehmer, C. J., Fabián, M. E., and Menegheti, J. O. (2004). Dental anomalies in the Atlantic population of South American sea lion, *Otaria byronia* (Pinnipedia, Otariidae): evolutionary implications and ecological approach. *Latin Am. J. Aquat. Mammals* 3, 7–18. doi: 10.5597/lajam00044

Enard, D., Messer, P. W., and Petrov, D. A. (2014). Genome-wide signals of positive selection in human evolution. *Genome Res.* 24, 885–895. doi: 10.1101/gr.164822.113

English, A. W. (1976). Limb movements and locomotor function in the California sea lion (Zalophus californianus). *J. Zool.* 178, 341–364. doi: 10.1111/j.1469-7998.1976.tb02274.x

Ewer, R. F. (1973). *The Carnivores*. Ithaca, NY: Cornell University Press.

Fan, Y., Wu, R., Chen, M. H., Kuo, L., and Lewis, P. O. (2010). Choosing among partition models in Bayesian phylogenetics. *Mol. Biol. Evol.* 28, 523–532. doi: 10.1093/molbev/msq224

Fay, F. H., Rausch, V. R., and Felitz, E. T. (1967). Cytogeneic comparison of some pinnipeds (Mammalia: Eutheria). *Can. J. Zool.* 45, 773–778. doi: 10.1139/z67-088

Finarelli, J. A. (2008). A total evidence phylogeny of the Arctoidea (Carnivora: Mammalia): relationships among basal taxa. *J. Mammal. Evol.* 15, 231–259. doi: 10.1007/s10914-008-9074-x

Fish, F. E., Innes, S., and Ronald, K. (1988). Kinematics and estimated thrust production of swimming harp and ringed seals. *J. Exp. Biol.* 137, 157–173.

Flynn, J. J., Finarelli, J. A., Zehr, S., Hsu, J., and Nebdal, M. A. (2005). Molecular phylogeny of the Carnivora (Mammalia): assessing the impact of increased sampling on resolving enigmatic relationships. *Syst. Biol.* 54, 317–337. doi: 10.1080/10635150590923326

Flynn, J. J., Nebdal, M. A., Dragoo, J. W., and Honeycutt, R. L. (2000). Whence the red panda? *Mol. Phylogen. Evol.* 17, 190–199. doi: 10.1006/mpev.2000.0819

Flynn, J. J., and Nedbal, M. A. (1998). Phylogeny of the Carnivora (Mammalia): congruence vs incompatibility among multiple data sets. *Mol. Phylogen. Evol.* 9, 414–426. doi: 10.1006/mpev.1998.0504

Flynn, J. J., Neff, N. A., and Tedford, R. H. (1988). Phylogeny of the Carnivora. *Phylogeny Classif. Tetrapods* 2, 73–116.

Foote, A. D., Liu, Y., Thomas, G. W., Vina,r, T., Alföldi, J., Deng, J., et al. (2015). Convergent evolution of the genomes of marine mammals. *Nat. Genet.* 47:272. doi: 10.1038/ng.3198

Fulton, T. L., and Strobeck, C. (2006). Molecular phylogeny of the Arctoidea (Carnivora): effect of missing data on supertree and supermatrix analyses of multiple gene data sets. *Mol. Phylogen. Evol.* 41, 165–181. doi: 10.1016/j.ympev.2006.05.025

Fulton, T. L., and Strobeck, C. (2007). Novel phylogeny of the raccoon family (Procyonidae: Carnivora) based on nuclear and mitochondrial DNA evidence. *Mol. Phylogen. Evol.* 43, 1171–1177. doi: 10.1016/j.ympev.2006.10.019

Fulton, T. L., and Strobeck, C. (2010). Multiple fossil calibrations, nuclear loci and mitochondrial genomes provide new insight into biogeography and divergence timing for true seals (Phocidae, Pinnipedia). *J. Biogeograp.* 37, 814–829. doi: 10.1111/j.1365-2699.2010.02271.x

Furbish, R. (2015). *Something Old, Something New, Something Swimming in the Blue: An Analysis of the Pinniped Family Desmatophocidae, its Phylogenetic Position and Swimming Mode*. Diss. San Diego, CA: San Diego State University.

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. *Syst. Biol.* 66, 57–73. doi: 10.1093/sysbio/syw060

Geraads, D., and Spassov, N. (2016). Musteloid carnivores from the upper Miocene of South-Western Bulgaria, and the phylogeny of the Mephitidae. *Geodiversitas* 38, 543–559. doi: 10.5252/g2016n4a5

Gingerich, P. D., Raza, S. M., Arif, M., Anwar, M., and Zhou, X. (1994). New whale from the Eocene of Pakistan and the origin of cetacean swimming. *Nature* 368:844. doi: 10.1038/368844a0

Gingerich, P. D., Smith, B. H., and Simons, E. L. (1990). Hind limbs of Eocene *Basilosaurus*: evidence of feet in whales. *Science* 249, 154–157. doi: 10.1126/science.249.4965.154

Gittleman, J. L., and Valkenburgh, B. V. (1997). Sexual dimorphism in the canines and skulls of carnivores: effects of size, phylogency, and behavioural ecology. *J. Zool.* 242, 97–117. doi: 10.1111/j.1469-7998.1997.tb02932.x

Goloboff, P. A., Farris, J. S., and Nixon, K. C. (2008). TNT, a free program for phylogenetic analysis. *Cladistics* 24, 774–786. doi: 10.1111/j.1096-0031.2008.00217.x

Gordon, K. R. (1981). Locomotor behaviour of the walrus (*Odobenus*). *J. Zool.* 195, 349–367. doi: 10.1111/j.1469-7998.1981.tb03470.x

Gregory, W. K. (1910). *The Orders of Mammals*, Vol. 27. New York, NY: Order of the Trustees [American Museum of Natural History].

Grohé, C., Tseng, Z. J., Lebrun, R., Boistel, R., and Flynn, J. J. (2016). Bony labyrinth shape variation in extant Carnivora: a case study of Musteloidea. *J. Anat.* 228, 366–383. doi: 10.1111/joa.12421

Hall, B. K. (1999). *The Neural Crest in Development and Evolution*. New York, NY: Springer Science & Business Media.

Heath, T. A., Huelsenbeck, J. P., and Stadler, T. (2014). The fossilized birth–death process for coherent calibration of divergence-time estimates. *Proc. Natl. Acad. Sci.* 111, 2957–E966. doi: 10.1073/pnas.1319091111

Higdon, J. W., Bininda-Emonds, O. R., Beck, R. M., and Ferguson, S. H. (2007). Phylogeny and divergence of the pinnipeds (Carnivora, Mammalia) assessed using a multigene dataset. *BMC Evol. Biol.* 7:216. doi: 10.1186/1471-2148-7-216

Huelsenbeck, J. P., Jain, S., Frost, S. W., and Pond, S. L. K. (2006). A Dirichlet process model for detecting positive selection in protein-coding DNA sequences. *Proc. Natl. Acad. Sci.* 103, 6263–6268. doi: 10.1073/pnas.0508279103

Huelsenbeck, J. P., and Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17, 754–755. doi: 10.1093/bioinformatics/17.8.754

Hunt, R. M. Jr., and Barnes, L. G. (1994). Basicranial evidence for ursid affinity of the oldest pinnipeds. *Proceedings of San Diego Society of Natural History,* Vol. 29, 57–67.

Hunt, R. M. Jr., and Skolnick, R. (1996). The giant mustelid *Megalictis* from the early Miocene carnivore dens at Agate Fossil Beds National Monument, Nebraska; earliest evidence of dimorphism in New World Mustelidae (Carnivora, Mammalia). *Rocky Mt. Geol.* 31, 35–48.

Illiger, J. K. W. (1811). *Prodomus Systematis Mammalium et Avium Additis Terminis Zoographicii Utriusque Classis Eorumque Versione Germanica*. Berlin: sumptibus C. Salfeld.

Ingram, T., and Mahler, D. L. (2013). SURFACE: detecting convergent evolution from comparative data by fitting Ornstein-Uhlenbeck models with stepwise Akaike Information Criterion. *Methods Ecol. Evol.* 4, 416–425. doi: 10.1111/2041-210X.12034

Irwin, D. M., and Árnason, Ú. (1994). Cytochromeb gene of marine mammals: phylogeny and evolution. *J. Mammal. Evol.* 2, 37–55. doi: 10.1007/BF01464349

Jones, K. E., Smaers, J. B., and Goswami, A. (2015). Impact of the terrestrial-aquatic transition on disparity and rates of evolution in the carnivoran skull. *BMC Evol. Biol.* 15:1. doi: 10.1186/s12862-015-0285-5

Klopfstein, S., Vilhemlsen, L., and Ronquist, F. (2015). A nonstationary Markov model detects directional evolution in hymenopteran morphology. *Syst. Biol.* 64, 1089–1103. doi: 10.1093/sysbio/syv052

Kohno, N. (1993). "Have the pinnipeds been derived from the ursids," in *Abstracts of the Annual Meeting of the Palaeontological Society of Japan* (Tsukuba). 79.

Kohno, N. (1994). A new Miocene pinniped in the genus *Prototaria* (Carnivora: Odobenidae) from the Moniwa Formation, Miyagi, Japan. *J. Vertebr. Paleontol.* 14, 414–426. doi: 10.1080/02724634.1994.10011568

Kohno, N. (1996a). "The Oligo-Miocene aquatic arctoid carnivore *Potamotherium*, and its bearing on the relationships of pinnipeds," in *Resumes Readaptations Milieu Aquatique*, Vol. 1 (Poitiers), 22–23.

Kohno, N. (1996b). Miocene pinniped *Allodesmus* with special reference to the "Mito seal" from Ibaraki Prefecture, Central Japan. *Transact. Proc. Palaeontol. Soc. Jpn.* 181, 388–404.

Kohno, N. (2006). A new Miocene odobenid (Mammalia: Carnivora) from Hokkaido, Japan, and its implications for odobenid phylogeny. *J. Vertebr. Paleontol.* 26, 411–421. doi: 10.1671/0272-4634(2006)26[411:ANMOMC]2.0.CO;2

Kohno, N., Barnes, L. G., and Hirota, K. (1995). Miocene fossil pinnipeds of the genera *Prototaria* and *Neotherium* (Carnivora; Otariidae; Imagotariinae) in the north Pacific Ocean: evolution, relationships, and distribution. *Island Arc.* 3, 285–308. doi: 10.1111/j.1440-1738.1994.tb00117.x

Koretsky, I. A., and Barnes, L. G. (2006). "Pinniped evolutionary history and paleobiogeography," in *Mesozoic and Cenozoic Vertebrates and*

Paleoenvironments: Tributes to the Career of Professor Dan Grigorescu (Bucharest: Ars Docendi), 143–153.

Koretsky, I. A., Barnes, L. G., and Rahmat, S. J. (2016). Re-Evaluation of Morphological Characters Questions Current Views of Pinniped Origins. *Vestnik zoologii* 50, 327–354. doi: 10.1515/vzoo-2016-0040

Koretsky, I. A., and Domning, D. P. (2014). One of the oldest seals (Carnivora, Phocidae) from the Old World. *J. Vertebr. Paleontol.* 34, 224–229. doi: 10.1080/02724634.2013.787428

Koretsky, I. A., and Holec, P. (2002). A primitive seal (Mammalia: Phocidae) from the early middle Miocene of Central Paratethys. *Smithson. Contrib. Paleobiol.* 93, 163–178.

Koretsky, I. A., and Rahmat, S. J. (2015). A new species of the subfamily Devinophocinae (Carnivora, Phocidae) from the Central Paratethys. *Res. Paleontol. Stratigrap.* 121, 31–47. doi: 10.13130/2039-4942/6399

Kosiol, C., Vinar, T., da Fonseca, R. R., Hubisz, M. J., Bustamante, C. D., Nielsen, R., et al. (2008). Patterns of positive selection in six mammalian genomes. *PloS Genet.* 4:e1000144. doi: 10.1371/journal.pgen.1000144

Kuhn, C., and Frey, E. (2012). Walking like caterpillars, flying like bats—pinniped locomotion. *Palaeobiodivers. Palaeoenviron.* 92, 197–210. doi: 10.1007/s12549-012-0077-5

Lee, M. S. (2016). Multiple morphological clocks and total evidence tip-dating in mammals. *Biol. Lett.* 12:20160033. doi: 10.1098/rsbl.2016.0033

Maddison, W. P., and Maddison, D. R. (2015). *Mesquite: A Modular System for Evolutionary Analysis. Version 2.75. 2011*.

Matzke, N. J., and Wright, A. (2016). Inferring node dates from tip dates in fossil canidae: the importance of tree priors. *Bio. Lett.* 12:20160328. doi: 10.1098/rsbl.2016.0328

McClellan, D. A. (2013). Directional darwinian selection in proteins. *BMC Bioinform.* 14 (Suppl. 13):S6. doi: 10.1186/1471-2105-14-S13-S6

McLaren, I. A. (1960). Are the pinnipedia biphyletic? *System. Zoolog.* 9, 18–28. doi: 10.2307/2411537

Merritt, T. J. S., and Quattro, J. M. (2001). Evidence for a period of directional selection following gene duplication in a neurally expressed locus of triosephosphate isomerase. *Genetics* 159:689697.

Mitchell, E., and Tedford, R. H. (1973). The enaliarctinae: a new group of extinct aquatic carnivora and a consideration of the origin of the otariidae. *Bull. AMNH* 151, 201–284.

Mivart, S. G. (1885). On the anatomy, classification, and distribution of the arctoidea. *Proc. Zoolog. Soc. Lond.* 53, 340–404. doi: 10.1111/j.1096-3642.1885.tb02915.x

Miyamoto, M. M., and Goodman, M. (1986). Biomolecular systematics of eutherian mammals: phylogenetic patterns and classification. *Syste.Bio.* 35, 230–240. doi: 10.1093/sysbio/35.2.230

Muschick, M., Indermaur, A., and Salzburger, W. (2012). Convergent evolution within an adaptive radiation of cichlid fishes. *Curr.Bio.* 22, 2362–2368. doi: 10.1016/j.cub.2012.10.048

Nielsen, R. (2005). Molecular signatures of natural selection. *Annu. Rev. Genet.* 39, 197–218. doi: 10.1146/annurev.genet.39.073003.112420

Nojima, T. (1990). A morphological consideration of the relationships of pinnipeds to other carnivorans based on the bony tentorium and bony falx. *Mar. Mammal Sci.* 6, 54–74. doi: 10.1111/j.1748-7692.1990.tb00226.x

Nyakatura, K., and Bininda-Emonds, O. R. (2012). Updating the evolutionary history of Carnivora (Mammalia): a new species-level supertree complete with divergence time estimates. *BMC biol.* 10:12. doi: 10.1186/PREACCEPT-5398900576110216

Orlov, J. A. (1933). *Semantor macrurus* (ordo Pinnipedia, fam. semantoridae fam. nova) aus den neogen-ablagerungen westsibiriens. *Trav.l'Instit. Paléozool. Acad. Sci. URSS* 2, 165–268.

Pattinson, D. J., Thompson, R. S., Piotrowski, A. K., and Asher, R. J. (2014). Phylogeny, paleontology, and primates: do incomplete fossils bias the tree of life? *Syst.Biol.* 64, 169–186. doi: 10.1093/sysbio/syu077

Pierce, S. E., Clack, J. A., and Hutchinson, J. R. (2011). Comparative axial morphology in pinnipeds and its correlation with aquatic locomotory behaviour. *J. Anato.* 219, 502–514. doi: 10.1111/j.1469-7580.2011.01406.x

Prieto-Marquez, A. (2010). Global phylogeny of hadrosauridae (Dinosauria: Ornithopoda) using parsimony and bayesian methods. *Zool. J. Linn. Soc.* 159, 435–502. doi: 10.1111/j.1096-3642.2009.00617.x

Prothero, D. R., Bitboul, C. Z., Moore, G. W., and Niem, A. R. (2001). *Magnetic Stratigraphy and Tectonic Rotation of the Oligocene Alsea, Yaquina, and Nye Formations*. Lincoln County.

Pyron, R. A. (2011). Divergence time estimation using fossils as terminal taxa and the origins of lissamphibia. *Syst. Biol.* 60, 466–481. doi: 10.1093/sysbio/syr047

Rambaut, A. (2007). *FigTree, A Graphical Viewer of Phylogenetic Trees*.

Ray, C. E. (1976). Geography of phocid evolution. *Syst. Biol.* 25, 391–406. doi: 10.2307/2412513

Repenning, C. A., and Tedford, R. H. (1977). *Otarioid Seals of the Neogene*. No. 992. Washington: US Govt. Print. Off.

Romero-Herrera, A. E., Lehmann, H., Joyseyet, K. A., and Friday, A. E. (1978). On the evolution of myoglobin. *Philos. Trans. Royal Soc. Lond. B* 283, 61–163. doi: 10.1098/rstb.1978.0018

Ronquist, F., and Huelsenbeck, J. P. (2003). MrBayes 3: bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572–1574. doi: 10.1093/bioinformatics/btg180

Ronquist, F., Klopfstein, S., Vilhelmsen, L., Schulmeister, S., Murray, D. L., and Rasnitsyn, A. P. (2012a). A total-evidence approach to dating with fossils, applied to the early radiation of the hymenoptera. *Syst. Biol.* 61, 973–999. doi: 10.1093/sysbio/sys058

Ronquist, F., Lartillot, N., and Phillips, M. J. (2016). Closing the gap between rocks and clocks using total evidence dating. *Philos. Trans. Royal Soc. B* 371:20150136. doi: 10.1098/rstb.2015.0136

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., et al. (2012b). MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61, 539–542. doi: 10.1093/sysbio/sys029

Rybczynski, N., Dawson, M. R., and Tedford, R. H. (2009). A semi-aquatic arctic mammalian carnivore from the miocene epoch and origin of pinnipedia. *Nature* 458, 1021–1024. doi: 10.1038/nature07985

Sansom, R. S. (2014). Bias and sensitivity in the placement of fossil taxa resulting from interpretations of missing data. *Syst. Biol.* 64, 256–266. doi: 10.1093/sysbio/syu093

Sansom, R. S., and Wills, M. A. (2013). Fossilization causes organisms to appear erroneously primitive by distorting evolutionary trees. *Sci. Rep.* 3:2545. doi: 10.1038/srep02545

Sarich, V. M. (1969). Pinniped evolution and the rate of evolution of carnivore albumins. *U. S. Geol. Surv. Profes. Pap.* 992, 1–93.

Sato, J. J., Wolsan, M., Minami, S., Hosoda, T., Sinaga, M. H., Hiyama, K., et al. (2009). Deciphering and dating the red panda's ancestry and early adaptive radiation of musteloidea. *Mol.Phylogen. Evolu.* 53, 907–922. doi: 10.1016/j.ympev.2009.08.019

Sato, J. J., Wolsan, M., Prevosti, F. J., D'Elía, G., Begg, C., Begg, K., et al. (2012). Evolutionary and biogeographic history of weasel-like carnivorans (Musteloidea). *Mol. Phylogen. Evol.* 63, 745–757. doi: 10.1016/j.ympev.2012.02.025

Sato, J. J., Wolsan, M., Suzuki, H., Hosoda, T., Yamaguchi, Y., Hiyama, K., et al. (2006). Evidence from nuclear DNA sequences sheds light on the phylogenetic relationships of pinnipedia: single origin with affinity to musteloidea. *Zool. Sci.* 23, 125–146. doi: 10.2108/zsj.23.125

Savage, R. J. G. (1957). "October. The anatomy of Potamotherium an Oligocene lutrine," in *Proceedings of the Zoological Society of London*, Vol. 129 (Oxford: Blackwell Publishing Ltd), 151–244.

Simpson, G. G. (1944). *Tempo and Mode in Evolution*. New York, NY: Columbia University Press.

Spaulding, M., and Flynn, J. J. (2012). Phylogeny of the carnivoramorpha: the impact of postcranial characters. *J. Syst. Palaeontol.* 10, 653–677. doi: 10.1080/14772019.2011.630681

Stirton, R. A. (1960). *A Marine Carnivore From the Clallam Miocene Formation, Washington. A Review of the Sirenia and Desmostylia.* 345.

Tedford, R. H. (1976). Relationship of pinnipeds to other carnivores (Mammalia). *Syst. Biol.* 25, 363–374. doi: 10.2307/2412511

Tedford, R. H., Barnes, L. G., and Ray, C. E. (1994). "The early Miocene littoral ursoid carnivoran *Kolponomos*: systematics and mode of life," in *Proceedings of the San Diego Society of Natural History*, Vol. 29, 11–32.

Tomiya, S., and Tseng, Z. J. (2016). Whence the beardogs? Reappraisal of the Middle to Late Eocene 'Miacis' from Texas, USA, and the origin of

Amphicyonidae (Mammalia, Carnivora). *Royal Soc. open sci.* 3:160518. doi: 10.1098/rsos.160518

Tseng, Z. J., Wang, X., and Stewart, J. D. (2009). A new immigrant mustelid (Carnivora, Mammalia) from the middle Miocene Temblor Formation of central California. *PaleoBios.* 29, 13–23.

Valenciano, A., Baskin, J. A., Abella, J., Pérez-Ramos, A., Álvarez-Sierra, M. Á., Morales, J., et al. (2016). *Megalictis*, the Bone-Crushing Giant Mustelid (Carnivora, Mustelidae, Oligobuninae) from the Early Miocene of North America. *PLoS ONE* 11:e0152430. doi: 10.1371/journal.pone.0152430

Velez-Juarbe, J. (2017). *Eotaria citrica*, sp. nov., a new stem otariid from the "Topanga" formation of Southern California. *PeerJ* 5:e3022. doi: 10.7717/peerj.3022

Wang, X., Grohé, C., Su, D. F., White, S. C., Ji, X., Kelley, J., et al. (2017). A new otter of giant size, *Siamogale melilutra* sp. nov. (Lutrinae: Mustelidae: Carnivora) from the latest miocene shuitangba site in north-eastern Yunnan, south-western China, and a total evidence analysis of lutrines. *J. Syst. Palaeontol.* 16, 39–65. doi: 10.1080/14772019.2016.1267666

Wang, X., McKenna, M. C., and Dashzeveg, D. (2005). *Amphicticeps* and *Amphicynodon* (Arctoidea, Carnivora) from Hsanda Gol Formation, central Mongolia and phylogeny of basal arctoids with comments on zoogeography. *Am. Mus. Novitates* 3483, 1−60. doi: 10.1206/0003-0082(2005)483[0001:AAAACF]2.0.CO;2

Wang, Z., Yuan, L., Rossiter, S. J., Zuo, X., Ru, B., Zhong, H., et al. (2008). Adaptive evolution of 5′ HoxD genes in the origin and diversification of the cetacean flipper. *Mol. Biol. Evol.* 26, 613–622. doi: 10.1093/molbev/msn282

Weber, M. W. C. (1904). *Die Säugetiere: Einführung in die Anatomie und Systematik der recenten und fossilen Mammalia*, ed J. G. Fischer.

Wesley-Hunt, G. D., and Flynn, J. J. (2005). Phylogeny of the Carnivora: basal relationships among the carnivoramorphans, and assessment of the position of 'Miacoidea'relative to Carnivora. *J. Syst. Palaeontol.* 3, 1–28. doi: 10.1017/S1477201904001518

Wilgenbusch, J. C. (2004). *AWTY: A System for Graphical Exploration of MCMC Convergence in Bayesian Phylogenetic Inference*. Available online at: http://ceb.csit.fsu.edu/awty

Wolsan, M. (1993). Phylogeny and classification of early European Mustelida (Mammalia: Carnivora). *Acta Theriol.* 38, 345–384. doi: 10.4098/AT.arch.93-29

Wozencraft, W. C. (1989). "The phylogeny of the recent Carnivora," in *Carnivore Behavior, Ecology, and Evolution* (New York City, NY: Springer US), 495–535.

Wyss, A. R. (1987). The walrus auditory region and the monophyly of pinnipeds. *Amer. Mus. Novitates* 2871, 1–31.

Wyss, A. R. (1988). On "retrogression" in the evolution of the Phocinae and phylogenetic affinities of the monk seals. *Am. Museum Nat. Hist.* 2924, 1–38.

Wyss, A. R., and Flynn, J. J. (1993). "A phylogenetic analysis and definition of the Carnivora," in *Mammal Phylogeny: Placentals*, eds F. S. Szalay, M. J. Novacek, and M. C. McKenna. New York, NY: Springer, 32–52.

Xie, W., Lewis, P. O., Fan, Y., Kuo, L., and Chen, M. H. (2010). Improving marginal likelihood estimation for Bayesian phylogenetic model selection. *Syst. Biol.* 60, 150–160. doi: 10.1093/sysbio/syq085

Yonezawa, T., Kohno, N., and Hasegawa, M. (2009). The monophyletic origin of sea lions and fur seals (Carnivora; Otariidae) in the Southern Hemisphere. *Gene* 441, 89–99. doi: 10.1016/j.gene.2009.01.022

Zhang, C., Stadler, T., Klopfstein, S., Heath, T. A., and Ronquist, F. (2016). Total-evidence dating under the fossilized birth-death process. *Syst. Biol.* 65, 228–249. doi: 10.1093/sysbio/syv080

Zhou, X., Seim, I., and Gladyshev, V. N. (2015). Convergent evolution of marine mammals is associated with distinct substitutions in common genes. *Sci. Rep.* 5:16550. doi: 10.1038/srep16550

Zou, Z., and Zhang, J. (2016). Morphological and molecular convergences in mammalian phylogenetics. *Nat. Commun.* 7:12758. doi: 10.1038/ncomms12758

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Paterson, Rybczynski, Kohno and Maddin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genetics

BRIEF RESEARCH REPORT
published: 11 March 2020
doi: 10.3389/fgene.2020.00182

A Cambrian–Ordovician Terrestrialization of Arachnids

Jesus Lozano-Fernandez[1,2†], Alastair R. Tanner[1], Mark N. Puttick[3], Jakob Vinther[1,2], Gregory D. Edgecombe[4*] and Davide Pisani[1,2*]*

[1] School of Biological Sciences, University of Bristol, Bristol, United Kingdom, [2] School of Earth Sciences, University of Bristol, Bristol, United Kingdom, [3] Department of Biology and Biochemistry, Milner Centre for Evolution, University of Bath, Bath, United Kingdom, [4] Department of Earth Sciences, Natural History Museum, London, United Kingdom

Understanding the temporal context of terrestrialization in chelicerates depends on whether terrestrial groups, the traditional Arachnida, have a single origin and whether or not horseshoe crabs are primitively or secondarily marine. Molecular dating on a phylogenomic tree that recovers arachnid monophyly, constrained by 27 rigorously vetted fossil calibrations, estimates that Arachnida originated during the Cambrian or Ordovician. After the common ancestor colonized the land, the main lineages appear to have rapidly radiated in the Cambrian–Ordovician boundary interval, coinciding with high rates of molecular evolution. The highest rates of arachnid diversification are detected between the Permian and Early Cretaceous. A pattern of ancient divergence estimates for terrestrial arthropod groups in the Cambrian while the oldest fossils are Silurian (seen in both myriapods and arachnids) is mirrored in the molecular and fossil records of land plants. We suggest the discrepancy between molecular and fossil evidence for terrestrialization is likely driven by the extreme sparseness of terrestrial sediments in the rock record before the late Silurian.

Keywords: Arachnida, Chelicerata, terrestrialization, Cambrian, molecular clocks, diversification

OPEN ACCESS

Edited by:
Rachel C. M. Warnock, ETH Zürich, Switzerland

Reviewed by:
Joanna Wolfe, Harvard University, United States
David Marjanović, Leibniz Institut für Evolutions und Biodiversitätsforschung, Germany

****Correspondence:***
Jesus Lozano-Fernandez
jesus.lozano@ibe.upf-csic.es
Gregory D. Edgecombe
g.edgecombe@nhm.ac.uk
Davide Pisani
davide.pisani@bristol.ac.uk

†Present address:
Jesus Lozano-Fernandez, Institute of Evolutionary Biology (CSIC-UPF), Barcelona, Spain

Specialty section:
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Genetics

Received: *04 September 2019*
Accepted: *14 February 2020*
Published: *11 March 2020*

Citation:
Lozano-Fernandez J, Tanner AR, Puttick MN, Vinther J, Edgecombe GD and Pisani D (2020) A Cambrian–Ordovician Terrestrialization of Arachnids. Front. Genet. 11:182. doi: 10.3389/fgene.2020.00182

INTRODUCTION

Arachnids are an important group of terrestrial arthropods, including the familiar ticks, mites, spiders, and scorpions, together with pseudoscorpions, camel spiders, vinegaroons, whip spiders, and a few other groups. Arachnids are important predatory arthropods across almost every conceivable terrestrial habitat. While ticks are ectoparasites that affect humans and livestock, spiders are ecologically the most successful arachnids and as predators consume vast quantities of insects. Thus, understanding when arachnids colonized land and diversified is of interest from a macroevolutionary and macroecological perspective.

Arachnids are chelicerates, together with the marine horseshoe crabs (Xiphosura) and sea spiders (Pycnogonida). They are the most speciose clade in Chelicerata, with more than 112,000 described extant species. Together with hexapods and myriapods, arachnids represent one of three distinct and ancient events of arthropod terrestrialization (terrestrial isopods are a younger addition to the continental arthropod biota). While arachnids have traditionally been considered monophyletic and terrestrial (apart from secondarily marine mites: Pepato et al., 2018) this picture has been challenged at different times. Scorpions were long thought to be the sister group of all other extant arachnids (Weygoldt and Paulus, 1979) or most closely allied to the aquatic "sea scorpions," the eurypterids

(Dunlop and Webster, 1999). Early fossil scorpions have been interpreted as aquatic, and in some cases even as marine. These phylogenetic hypotheses and interpretations of fossil ecology have been seen as requiring independent events of terrestrialization. Some of these views have been overturned by strong molecular (Regier et al., 2010; Sharma et al., 2014; Sharma and Wheeler, 2014; Leite et al., 2018; Ballesteros and Sharma, 2019; Lozano-Fernandez et al., 2019) and morphological (Garwood and Dunlop, 2014; Klußmann-Fricke and Wirkner, 2016) evidence for scorpions being nested within the Arachnida as the sister group of the other arachnids with book lungs, the Tetrapulmonata. Indeed, detailed correspondences in book lung morphology between scorpions and tetrapulmonates support their homology (Scholtz and Kamenz, 2006). The supposed aquatic mode of life of various fossil scorpions has also been questioned on both morphological and geological grounds (Dunlop et al., 2008; Kühl et al., 2012). Another challenge to a single terrestrialization event in arachnids came from analyses of phylogenomic datasets, which have often recovered the marine Xiphosura to be nested *within* Arachnida (Ballesteros and Sharma, 2019). This remains a contentious issue, as other phylogenomic analyses have yielded trees in which Arachnida is monophyletic (Lozano-Fernandez et al., 2019).

Most chelicerate lineages are predatory components of a diverse range of ecosystems, and the rock record attests to their presence in both earlier Paleozoic marine settings (Legg, 2014) and through into the Mesozoic and Cenozoic, which witnessed a prolific diversification of spiders and other terrestrial arachnids (Penney, 2003; Selden et al., 2009). The terrestrial rock record prior to the Silurian is very sparse (Kenrick et al., 2012) and has presented some apparent discordances when investigating myriapod (Fernández et al., 2018), hexapod (Lozano-Fernandez et al., 2016), and plant divergence times. While the body fossil record of terrestrial plants and arthropods does not extend much further back than the Silurian (∼443–419 Ma), molecular clock estimates go back to the Ordovician (485–443 Ma) and Cambrian (538–485 Ma) (Morris et al., 2018b; Lozano-Fernandez et al., 2016). However, likely plant spores with desiccation-resistant adaptations extend back to the middle Cambrian (Gensel, 2008). The fossil record presents no unequivocal evidence for crown-group arachnids before the Silurian. The oldest crown-group arachnids are of Silurian age (stem-group Scorpiones in the Llandovery), followed by the extinct Trigonotarbida in the late Silurian (early Přídolí), Acariformes and Opiliones in the Early Devonian (Pragian), and Pseudoscorpiones in the Middle Devonian (Givetian). Several other arachnid orders first appear in the Carboniferous, including Araneae, Uropygi, Amblypygi, and Ricinulei. In contrast to a picture of scattered branches of the arachnid crown-group first appearing in the Siluro–Devonian, older representatives of the arachnid lineage are stem-group Arachnida, and are marine shoreline or brackish water/estuarine forms rather than being terrestrial. These include Chasmataspida and Eurypterida, the earliest members of which date to the Miaolingian Series of the Cambrian (Drumian Stage) and the Late Ordovician (Sandbian), respectively (Dunlop et al., 2003; Dunlop, 2010; Wolfe et al., 2016). Xiphosura-like chelicerates have a good fossil record, showing considerable morphological stasis, with marine stem-group representatives of Xiphosura such as *Lunataspis* being documented from the Late Ordovician, ca. 445 Ma (Rudkin et al., 2008), and a species from the Early Ordovician (Tremadocian) of Morocco (Van Roy et al., 2010) extends the lineage's history even deeper. With such a deep history revealed by the fossil record, any inferred phylogenetic position for Xiphosura within terrestrial arachnids would imply that the marine ecology of this lineage should be a secondary acquisition. Although such a scenario is paleontologically unlikely, molecular studies have often recovered horseshoe crabs in highly derived clades of arachnids, such as sister groups to Opiliones or Palpigradi (Pepato and Klimov, 2015), Ricinulei (Sharma et al., 2014; Ballesteros and Sharma, 2019), or Scorpiones and Araneae (von Reumont et al., 2012; Roeding et al., 2009; Sanders and Lee, 2010).

As with other terrestrial groups, molecular dating has recovered old dates for the origin and main diversification of arachnids. As part of wider campaigns investigating arthropods using just a few arachnid representatives, Rota-Stabelli et al. (2013), Wheat and Wahlberg (2013), and Lozano-Fernandez et al. (2016) recovered dates for the origin of Arachnida with credibility intervals bracketed between the Cambrian, in the first two studies, and Ordovician in the latter. Recently, Ballesteros and Sharma (2019) reported a chelicerate molecular phylogeny in which when they constrained Arachnida to be monophyletic inferred an Ediacaran origin for the group. Consequently, there are significant geochronological discrepancies, particularly for terrestrial lineages, between the molecular clock-based studies and the younger dates suggested by the first appearances of fossils. As fossils do not inform on the age of origin of clades (Signor and Lipps, 1982), but rather provide minimum ages of divergence (Donoghue and Benton, 2007), clock-based methods are required to approach an accurate evolutionary timescale.

Focusing on the favored topology of Lozano-Fernandez et al. (2019) that recovers both Arachnida and Acari as monophyletic groups, we here estimate the divergence time of arachnids. To calibrate the molecular clock, we use a carefully selected and expanded set of 27 fossil constraints across the tree.

MATERIALS AND METHODS

Phylogenetic Reconstruction

The molecular supermatrix used here is composed of 89 species, 75 of them being chelicerates, with 14 other panarthropod species as outgroups. This matrix (Lozano-Fernandez et al., 2019, matrix A after exclusion of six unstable taxa) is a concatenation of 233 highly conserved and slow-evolving genes retrieved from transcriptomic data (45,939 amino acid positions and 78,1% complete). To evaluate the robustness of the results to an alternative topology, we also performed a divergence-date analysis in which Arachnida was non-monophyletic, with Xiphosura nested inside the arachnids (Lozano-Fernandez et al., 2019, matrix B containing 95 taxa). The phylogenetic trees were inferred using PhyloBayes MPI v.4.1 (Lartillot et al., 2013) under the site-heterogeneous CAT–GTR + Γ model of amino acid substitution (Lartillot and Philippe, 2004). Convergence was

assessed by running two independent Markov chains and using the bpcomp and tracecomp tools from PhyloBayes to monitor the maximum discrepancy in clade support (maxdiff), the effective sample size (effsize), and the relative difference in posterior mean estimates (rel_diff) for several key parameters and summary statistics of the model. We ran the analysis for 10,000 cycles and discarded as "burn-in" the first 3,000 generations.

Molecular Clock Analyses

Divergence time estimation was performed using PhyloBayes 3.3f (serial version) (Lartillot et al., 2009). We fixed the topology following Lozano-Fernandez et al. (2019), see previous section. We compared the fit of alternative, autocorrelated (CIR model – Lepage et al., 2006, 2007) and uncorrelated (uncorrelated gamma multipliers model; UGAM – Drummond et al., 2006), relaxed molecular clock models generating ten different random splits replicates and performing cross-validation analyses (see PhyloBayes manual for details). The tree was rooted on the Onychophora–Euarthropoda split. A set of 27 fossil calibrations and 1 node constrained by a maximum age (see **Table 1** and **Supplementary Data Sheet S4** for justifications) was used. We imposed a soft maximum of 559 Ma for the onychophoran–euarthropod split based on trace fossils in the White Sea/South Australian Ediacaran. This uses the radiometric date of 558+/− 1 Ma from Martin et al. (2000) for strata at which body fossils such as *Kimberella*, a putative total-group bilaterian metazoan (Martin et al., 2000; Benton et al., 2015), occur. Metazoan trace fossils in the White Sea/South Australian Ediacaran indicate suitable preservation for arthropod traces, were they present. We regard this to be a conservative soft maximum, as there is no body or trace fossil evidence for arthropods in the Ediacaran. A minimum for the divergence of onychophorans and arthropods is set by the earliest *Rusophycus* traces (total-group Arthropoda), dated to a minimum of 528.8 Ma following Wolfe et al. (2016). To allow the analysis to explore younger ages and prevent having posterior ages being much older than the fossil record, we also set maximum constraints on a few of the deepest calibrations within Euarthropoda (Ho and Phillips, 2009; Wolfe et al., 2016; Morris et al., 2018a). We infer that crown-group Mandibulata and Chelicerata do not predate the oldest fossil evidence for arthropods (*Rusophycus*; see above) and set soft maxima for each of this pair of sister taxa at the base of the Cambrian (538.8 Ma following Linnemann et al. (2019). Within Arachnida, we infer Acari and Arachnopulmonata do not predate the oldest body fossils of crown-group Chelicerata, using *Wisangocaris barbarahardyae* and its date of 509 Ma as a soft maximum following Wolfe et al. (2016). The amino acid substitution model used to estimate branch lengths was the CAT–GTR + Γ model, as in the phylogenetic analyses of Lozano-Fernandez et al. (2019). All analyses were conducted using soft bounds with 5% of the probability mass outside the calibration interval. A birth–death model was used to define prior node ages. Analyses were run under the priors to evaluate the effective joint priors induced by our choice of calibrations and root maxima. Convergence was considered achieved with *tracecomp* statistics dropping below 1 for all relative difference scores, and all effective sample sizes being above 50, for all chain parameters. The time-scaled phylogenies were plotted using the package MCMCtreeR, which allows the display of full posterior distributions on nodes and the inclusion of the geological timescale (Puttick, 2019). We included as supplementary data the chronograms, the guiding trees and the calibration file used in PhyloBayes, the subset of sampled timescaled trees used to generate the posterior distributions shown on the figures ("datedist" PhyloBayes file), and the two molecular matrices (**Supplementary Data Sheet S5**).

Rate of Molecular Evolution and Diversification

We estimated rates of molecular evolution within Chelicerata using two different methods. For the first method, we modeled the rates of molecular evolution on a fixed tree topology constrained to the timetree relationships. On this tree we estimated relative branch lengths under the C60 model + Γ in IQTree (Nguyen et al., 2015). We then divided these relative branch lengths by the timetree lengths to provide an estimate of absolute molecular rates through time. For the second method, we inferred ancestral estimates of the amino acid sequence on the fixed timetree, again using the C60 model + Γ in IQTree. We divided the sum of gross amino acid changes between ancestral and descendant nodes by absolute time to obtain per-branch rates of change.

We estimated speciation and extinction rates on the fixed timetree by using a Bayesian episodic diversification rate model in RevBayes 1.0.10 (Höhna et al., 2016). This model estimates piece-wise rates of speciation and extinction on a phylogeny through time (Stadler, 2011; Höhna, 2015). Within each bin, rates of speciation are equal but can differ between bins. The initial episodic speciation and extinction rate was sampled from a log-uniform distribution $U(-10,10)$. Moving backward in time for each distinct time bin, the model samples speciation and extinction rate from a normal distribution with the mean inherited from the value of the previous bin so rates are auto-correlated. Each normal distribution has a standard deviation inferred from an exponential hyper-prior of mean 1. In this manner, the model follows a Brownian motion pattern of rate change through time. To incorporate incomplete sampling in the model, we provided estimates of the known extant species numbers to complement the diversity shown in the tree using empirical taxon sampling by providing estimate diversity represented by each tip on our incomplete time tree. This empirical taxon-sampling approach is believed to produce less biased estimates of speciation and extinction parameters compared to diversified taxon sampling (Höhna, 2014). We used the values of the described extant species from Zhang (2013): Pycnogonida (1346); Xiphosura (4); Ricinulei (77); Opiliones (6571); Solifugae (1116); Acariformes (42233); Parasitiformes (12385); Pseudoscorpiones (3574); Scorpiones (2109); Uropygi (119); Amblypygi (172); and Araneae (44863). As we tested for the presence of early high rates compared to later times rather than differences in geological time units

TABLE 1 | Set of numbered fossil calibrations with minima and maxima.

Number	Clade constrained	Calibration	Maxima	Minima
1	Chelicerata	*Wisangocaris barbarahardyae*	538.8	509
2	Euchelicerata	*Chasmataspis*-like resting traces		500.5
3	Arachnopulmonata	*Palaeophonus loudonensis*	509	432.6
4	Opiliones	*Eophalangium sheari*	509	405
5	Acari	*Protocarus crani*	509	405
6	Sarcoptiformes	*Protochthonius gilboa*		382.7
7	Pedipalpi	*Parageralinura naufraga*		319.9
8	Palpatores	*Macrogyrion cronus*		298.75
9	Araneae	*Palaeothele montceauensis*		298.75
10	Avicularoidea	*Rosamygale grauvogeli*		240.5
11	Xiphosurida	*Tachypleus gadeai*		236
12	Synspermiata	*Eoplectreurys gertschi*		158.1
13	Entelegynae	*Mongolarachne jurassica*		158.1
14	Araneoidea	Unnamed Linyphiinae		129.41
15	Bipectina	*Cretamygale chasei*		125
16	Bothriuoidea + Scorpionoidea + "Chactoidea"	*Protoischnurus axelrodurum*		112.6
17	Cyphopthalmi	*Palaeosiro burmanicum*		98.17
18	Laniatores	*Petrobunoides sharmai*		98.17
19	Metastriata	*Amblyomma birmitum*		98.17
20	Buthida	*Uintascorpio halandrasorum*		48.5
21	Lobopodia	*Rusophycus avalonensis*	559	528.82
22	Altocrustacea	*Yicaris dianensis*		514
23	Progoneata	*Casiogrammus ichthyeros*		426.9
24	Chilopoda	*Crussolum* sp.		416
25	Hexapoda	*Rhyniella praecursor*		405
26	Aparaglossata	*Westphalomerope maryvonnae*		313.7
27	Onychophora	*Cretoperipatus burmiticus*		100
	Mandibulata		538.8	

(e.g., Period), we assumed there were 10 equally sized time intervals which can potentially possess distinct speciation and extinction rates. We included as supplementary data the input files used for the diversification analyses (**Supplementary Data Sheet S5**).

RESULTS

Molecular Divergence Time Estimation

The topology used for the molecular clock analyses here is obtained from Lozano-Fernandez et al. (2019). In this tree Chelicerata and Euchelicerata are monophyletic, with the horseshoe crabs retrieved as the sister group of monophyletic Arachnida. Bayesian cross-validation indicates that the autocorrelated CIR model (Lepage et al., 2006, 2007) most optimally fits the data – cross-validation score = 32.00 ± 9.44 – against UGAM (with all ten replicates supporting CIR as best). Accordingly, divergence time estimation was performed using the Autocorrelated CIR model and results within Chelicerata are presented in **Figure 1**, the full chronogram in **Supplementary Data Sheet S1** and the retrieved ages in **Table 2**.

The age of the Euarthropoda root, given the taxonomic sample, is recovered near the end of the Ediacaran, 546 million years ago (Ma), with the 95% highest posterior density (HPD) lying between 551 and 536 Ma. Chelicerata are inferred to originate at 535 Ma (with HPD 540 – 527 Ma), similar to the age retrieved for Mandibulata 535 Ma (539 – 526 Ma). The origin of Myriapoda comprises ages centered on the early Cambrian 516 Ma (524 – 505 Ma) and precedes that of Pancrustacea at 486 Ma (501 – 471 Ma). Hexapods are inferred to be much younger in age than myriapods, ranging through the Late Ordovician to Early Devonian, 422 Ma (448 – 400 Ma).

Arachnid terrestrialization is inferred to date to the Cambrian to Ordovician, crown-group Arachnida having a mean at 485 Ma (494 – 475 Ma). Therefore, our results support a Cambrian or Early Ordovician origin of two of the three main terrestrial arthropod lineages (myriapods and arachnids). The upper limit is consistent with fossil evidence for stem-group Arachnida (the chasmataspidid trackways noted above) but is substantially older than any crown-group fossils. Within Arachnida, rapid cladogenesis then occurred during the 20 million years that followed their origin, with several crown-group supra-ordinal clades becoming established in this time interval. By around 450 Ma, all 10 stem groups leading to extant orders of chelicerates included in this analysis (out of 12 in total – Palpigrada and Schizomida are unsampled) were already established. Further cladogenesis is inferred to have involved a more gradual tempo of evolution, in particular for Arachnopulmonata and Acari, which originated at ∼ 470 Ma but greatly expanded after the start of the

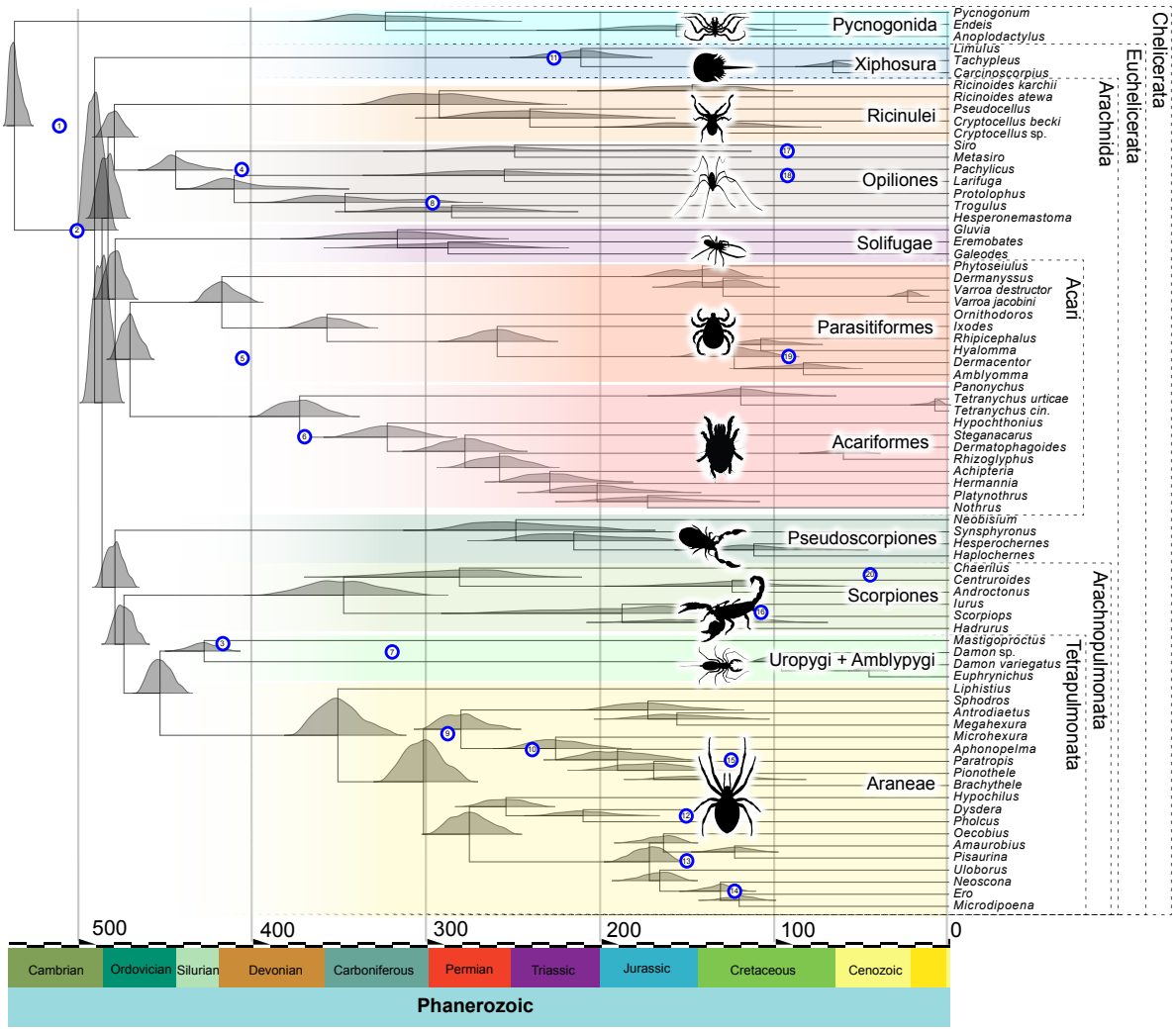


FIGURE 1 | Chelicerate divergence times in the molecular clock analysis (outgroups not shown). Divergence times shown are obtained under the CIR autocorrelated, relaxed molecular clock model. Nodes in the tree represent average divergence times. The density plots represent the posterior distributions from the considered node. The numbered blue circles represent the age of the fossil calibrations and are located at a height corresponding to the node they are calibrating (see **Table 1**). In the timescale on the *X* axis, numbers represent millions of years before the present.

TABLE 2 | Molecular divergence times for arthropod crown-group lineages.

Taxon	Mean age (Ma)	95% Credibility interval	Taxon	Mean age (Ma)	95% Credibility interval
Lobopodia	553	559–544	Pycnogonida	316	359–261
Euarthropoda	546	551–536	Xiphosura	210	239–179
Mandibulata	535	539–526	Ricinulei	294	338–236
Myriapoda	516	524–505	Opiliones	440	457–412
Chilopoda	429	460–389	Solifugae	318	361–258
Diplopoda	506	518–489	Acari	469	481–475
Pancrustacea	486	501–471	Parasitiformes	417	434–399
Hexapoda	422	448–400	Acariformes	374	396– 347
Chelicerata	535	540–527	Pseudoscorpiones	244	290–172
Euchelicerata	489	497–479	Scorpiones	354	395–300
Arachnida	485	494–475	Amblypygi	97	114–82
Arachnopulmonata	473	483–461	Araneae	352	375–328

Mesozoic (252 Ma to 66 Ma). Our dating suggests that the oldest crown-group arachnid orders are Opiliones and Parasitiformes, with Silurian and Devonian origins, respectively. Crown-group scorpions have a Devonian to Carboniferous origin, with the sampled extant lineages splitting more recently. For Araneae, the crown-group age is centered on the Devonian–Carboniferous boundary, with most extant mygalomorph and araneomorph lineages diversifying after the Jurassic.

In general, these Paleozoic age estimates for deep nodes within the most intensely sampled arachnid orders are similar to those inferred in other recent molecular dating analyses. For example, our estimates for crown-group Araneae is consistent with the Late Devonian (Fernández et al., 2018) or Early Carboniferous dates (Garrison et al., 2016) retrieved in other transcriptome-based analyses; likewise, a Carboniferous mean age for crown-group opistothele spiders is found in each of these studies. Our estimates encompassing a Late Ordovician median age for crown-group Opiliones corresponds to that estimated using tip dating by Sharma and Giribet (2014), whereas node calibration in that study recovered a Silurian median. In the case of Scorpiones, a Late Devonian to Carboniferous origin of the crown group is closely comparable to the date for the same node by Sharma et al. (2018), but older than the strictly Carboniferous ages estimated by Howard et al. (2019). However, in all cases mentioned the credibility intervals substantially overlap, indicating that these independent studies found results that, despite some differences, are not significantly different and corroborate each other. One exception is from a recent phylotranscriptomic study of Pseudoscorpiones, which retrieved an Ordovician to Carboniferous origin for the group (Benavides et al., 2019), significantly older than the Permian ages retrieved here. This may reflect the much more complete taxonomic coverage of pseudoscorpion diversity in the Benavides et al. (2019) analysis than in ours.

To assess whether our joint prior assumptions were driving our posterior estimates, we also ran the analysis under the priors (i.e., we performed analyses without data) and found that the joint priors allowed a wide possible distribution of ages, for the most part encompassing but not enforcing the posteriors (see **Supplementary Data Sheet S2**). We also performed a molecular clock analysis from a different matrix that resulted in a topology in which Xiphosura was nested within Arachnida, specifically as the sister group of Arachnopulmonata + Pseudoscorpiones. Overall, the result it is in general agreement with the main analysis, with most significant discrepancies concerning the age of Pycnogonida, which encompasses Silurian to Devonian ages, whereas in the main analysis are centered on the Carboniferous (see **Supplementary Data Sheet S3**).

Rate of Molecular Evolution and Diversification

We conducted estimations of molecular evolution and diversification rates based on our chelicerate timetree in an attempt to clarify whether the explosive cladogenesis at the onset of the arachnid radiation early in the Phanerozoic was matched by an increase in either of these rates. The analyses of rates of molecular evolution along the branches show a very high rate early in chelicerate history, including at the origin of Euchelicerata (**Figure 2**). These rates remain high during the early radiation of the Arachnida until the end of the Cambrian. Molecular rate estimations using branch lengths or ancestral sequences under a non-clock model gave nearly identical results. Using the episodic model of speciation and extinction rates through time, we found no evidence of high rates of speciation during the Cambrian, the period that presents the highest rates of molecular evolution. Instead, there is evidence for higher rates of cladogenesis later, bracketed between the Permian and Early Cretaceous, but especially high in the Permian and Triassic (**Figure 2**).

DISCUSSION

Molecular clocks allow the reconstruction of evolutionary timescales, but the reliability of these timescales depends on a variety of assumptions, which includes fossil data that have robust stratigraphic and phylogenetic justification, the use of a robust phylogenetic framework for the extant taxa, and the use of well-fitting models of both amino acid substitution and change in rate of molecular evolution. In this context, fossil calibrations then provide minimum ages for the origin of crown groups. Based on current best practice, our analysis uses fossil calibrations following the guidelines set out by Parham et al. (2012) (see **Table 1** and **Supplementary Data Sheet S4**). Furthermore, we report divergence times on a well-sampled phylogeny (Lozano-Fernandez et al., 2019), using the best-fitting molecular substitution and relaxed molecular clock models (see Lozano-Fernandez et al., 2019). We therefore contend that our findings provide the currently most robust insights into early chelicerate evolution (**Figure 1**). In our analysis, the ancestral pycnogonid divergence from Euchelicerata is inferred to have happened early in the Cambrian. This does not greatly predate the oldest unequivocal total-group pycnogonid, *Cambropycnogon klausmuelleri* (Waloszek and Dunlop, 2002), from the late Cambrian Orsten Konservat-Lagerstätte. The pycnogonid–euchelicerate divergence date suggests cryptic evolution of the euchelicerate stem group in the early Cambrian. Chelicerate – and, indeed, arthropod – body fossils are lacking in the earliest Cambrian, the Fortunian, the arthropod fossil record in the first 20 million years of the Cambrian being limited to trace fossils. Subsequently, it is estimated that xiphosurids diverged from arachnids in the late Cambrian, followed soon after by the radiation of crown-group Arachnida. While revising this paper, a new study on spider fossil calibrations (Magalhaes et al., 2019) came out and suggests that a few of the shallower calibrations used within Araneae treated as crown-groups may instead be stem-groups, so we add this caveat when interpreting the age of spiders.

There remain major geochronological discrepancies between the inferred molecular and fossil age of the various terrestrial arthropod groups. While these discrepancies may be thought to question the accuracy of molecular clocks, the differences need

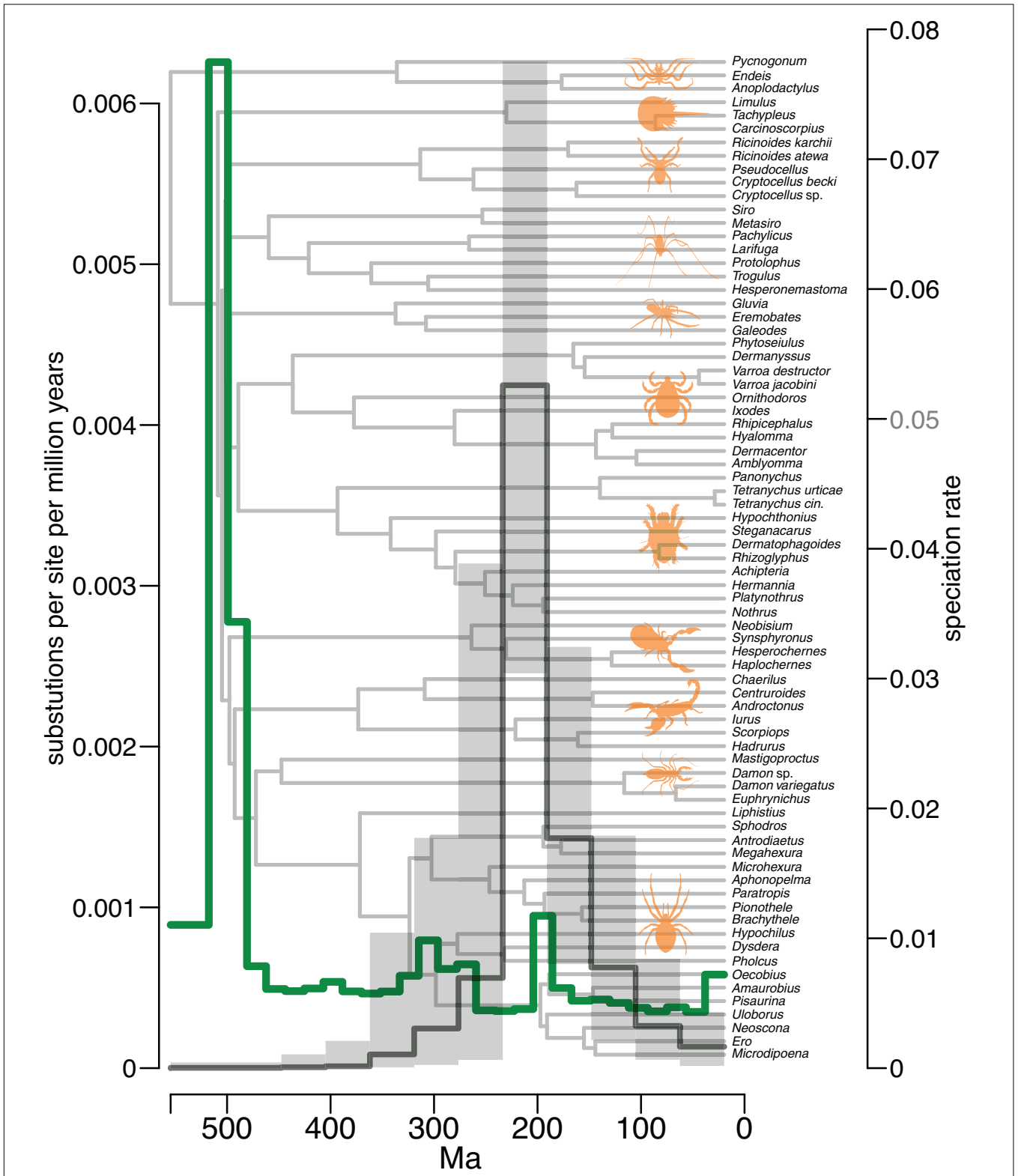


FIGURE 2 | Inferred rates of molecular evolution and diversification over time across chelicerates. Shown on the left *Y* axis are inferred amino acid substitutions per site per million years. On the right *Y* axis are diversification rates estimated with speciation rate as a proxy. Rates of molecular evolution are marked in green. Median speciation rates are marked in black with the highest posterior density of estimates shown in gray. In the timescale, numbers represent millions of years before the present.

to account for pervasive biases in the terrestrial sedimentary rock record. It has been noted that in Euramerica (from which much of the data on early terrestrial arthropods and early plant megafossils are derived), terrestrial sediments are rare before the late Silurian, and first become widespread in the Early Devonian (Kenrick et al., 2012). This temporal bias in the rock record almost certainly affects the fossil records of terrestrial organisms, and likely accounts for a major component of the discordance between molecular and fossil dates. The common recovery of horseshoe crabs as ingroup arachnids is perhaps unsurprising, given the short molecular branch lengths among these nodes in the tree, which also suggest short divergence times. With a reasonably good xiphosurid and pycnogonid fossil record, the molecular clock is well constrained among euchelicerates. This is evidenced by the short credibility intervals among deep arachnid nodes (**Figure 1**, **Table 2**, and **Supplementary Data Sheet S1**).

From an ecological context, it has been suggested that appreciably complex terrestrial ecosystems may have existed as far back as 1 billion years ago (Clarke et al., 2011), with molecular dating suggesting that crown-group land plants were already present by the middle Cambrian (Morris et al., 2018b). If it is indeed the case that myriapods and arachnids were on land so early, we speculate that the animals may have been early grazers on littoral bacterial mats, or predated on other amphibious or terrestrial organisms (Clarke et al., 2011). These ecologies represent habitats highly unfavorable to fossilization, such as high-energy environments characterized by erosion rather than deposition (Parry et al., 2018). It is unsurprising that paleontological insight is thus limited, and inference of the molecular kind as used here becomes more important as an investigative tool.

We estimate that arachnids colonized the land near the Cambrian–Ordovician boundary, and diversified soon after. Rates of molecular evolution were high at the onset of Arachnida, coinciding with rapid cladogenesis. High rates are concentrated on the branches leading to the major clades within Arachnida, representing major morphological and ecological partitions within the group. In unusually large and ancient clades, such as chelicerates, it is expected to find high rates of molecular evolution in their early lineages (Budd and Mann, 2018, 2019), and we retrieved results in agreement with that expectation (**Figure 2**). In order to avoid biases related to that fact, we imposed on the molecular clock analyses several maxima on the deepest nodes to account for possible overestimations of divergence times. Arachnids are predominantly predators, which must reflect the presence of an already diverse ecosystem, which the slightly older divergence times for myriapods and embryophytes established in the middle Cambrian. We therefore suspect that arachnids primitively represent carnivorous arthropods rather than having adapted to this mode of life convergently several times. The carnivorous centipede (Chilopoda) crown group is separated from other myriapods by a long branch estimated to be much younger in age (Fernández et al., 2018) than the detritivorous and/or fungal-feeding progoneate myriapods, allowing arachnids to be potentially the first carnivorous animals on land with early myriapod lineages as a likely source of prey. Hexapod divergence estimates are generally younger, suggesting a colonization of land no earlier than the Ordovician (Lozano-Fernandez et al., 2016; Schwentner et al., 2017).

There is a clear contrast in evolutionary tempo after the explosive radiation of the Cambrian and Ordovician. More gradual cladogenesis characterizes later Phanerozoic macroevolutionary dynamics of chelicerates, as is seen in the origins of ordinal clades. Our diversification studies reveal an increase in speciation rates bracketed between the Permian and the Early Cretaceous, in the origin of most sub-ordinal clades, with no evidence of higher speciation rates coinciding with the early rapid arachnid cladogenesis. A heightened diversification of spiders during the Cretaceous has previously been detected, suggested to result from the rise of angiosperms, stimulated by a warmer climate that led to the proliferation of spiders' main prey, insects (Shao and Li, 2018). Interestingly, we did not observe an early burst of diversification at the origin of chelicerates followed by a slowdown toward the present, a statistical bias usually found in large clades that survive to the present, the so-called "push of the past" (Nee et al., 1994). Instead, it seems that speciation rates are decoupled from the rates of molecular change. The common origin of arachnids giving rise to a plethora of adaptations, together with high molecular rates on the short internodes at the origin of the group suggests an ancient adaptive radiation shortly after colonizing the land, but our diversification analyses have not detected higher speciation rates at that time, one of the key features signaling an adaptive radiation (Glor, 2010). We acknowledge that the taxon sampling may not be the most adequate to infer speciation rates, as it was originally designed to maximize diversity, particularly at the deepest nodes to resolve the splits at the ordinal level.

CONCLUSION

Our analysis corroborates euchelicerates having radiated in the Cambrian and arachnids having diversified rapidly in the latest Cambrian–Early Ordovician. While this radiation was rather fast (see **Figure 1**), we found no evidence that the speciation rates that underpinned it were explosive. The late Cambrian to Early Ordovician emergence of arachnid stem groups onto land was soon followed by a rapid radiation near that same geological boundary, cladogenesis coinciding with high rates of molecular evolution during that time. A later phase of diversification within Arachnida is detected between the Permian and Early Cretaceous, during which the living arachnid orders exhibit heightened rates of speciation.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

JL-F, AT, DP, GE, and JV designed the experiments. JL-F, GE, and DP authored the main text with further suggestions from all other authors. JL-F carried out the sequencing laboratory work. JL-F, AT, and MP carried out the matrix compilation and computational analyses. AT, JL-F, and MP designed the figures.

FUNDING

JL-F was supported by a Marie Skłodowska-Curie Fellowship (655814) and a postdoctoral contract funded by the Beatriu de Pinós Programme of the Generalitat de Catalunya (2017-BP-00266). AT was supported by a University of Bristol (STAR) Ph.D. studentship. MP was funded by the Royal Commission for the Exhibition of 1851. JV and DP were supported by a NERC BETR grant (NE/P013678/1).

ACKNOWLEDGMENTS

We thank the journal's referees for insightful, constructive reviews.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.00182/full#supplementary-material

DATA SHEET S1 | Like **Figure 1**, but with the outgroups shown.

DATA SHEET S2 | Comparison for selected nodes of the 95% credibility intervals of the estimated divergence times generated by the joint priors (in blue) versus the posterior estimates (in red). The circles represent the average ages.

DATA SHEET S3 | Chelicerate divergence times in the molecular clock analysis not recovering Arachnida (outgroups shown) Divergence times shown are obtained under the CIR autocorrelated, relaxed molecular clock model. Nodes in the tree represent average divergence times estimated. The density plots represent the posterior distributions from the considered node. The numbered blue circles represent the age of the fossil calibrations and are located at a height corresponding to the node they are calibrating (see **Table 1**). In the timescale on the *X* axis, numbers represent millions of years before the present, and the geological period noted.

DATA SHEET S4 | Fossil justifications.

DATA SHEET S5 | Two molecular matrices, the guiding trees and the calibration file used in PhyloBayes analyses with the retrieved chronograms, and the subset of sampled timescaled trees used to generate the posterior distributions shown on the figures. The input files used for the diversification analyses are also included.

REFERENCES

Ballesteros, J. A., and Sharma, P. P. (2019). A critical appraisal of the placement of Xiphosura (Chelicerata) with account of known sources of phylogenetic error. *Syst. Biol.* 68, 896–917. doi: 10.1093/sysbio/syz011

Benavides, L. R., Cosgrove, J. G., Harvey, M. S., and Giribet, G. (2019). Phylogenomic interrogation resolves the backbone of the Pseudoscorpiones tree of life. *Mol. Phyl. Evol.* 139:106509. doi: 10.1016/j.ympev.2019.05.023

Benton, M. J., Donoghue, P. C. J., Asher, R. J., Friedman, M., Near, T. J., and Vinther, J. (2015). Constraints on the timescale of animal evolutionary history. *Palaeontol. Electron.* 18, 1–106. doi: 10.26879/424

Budd, G. E., and Mann, R. P. (2018). History is written by the victors: the effect of the push of the past on the fossil record. *Evolution* 72, 2276–2291. doi: 10.1111/evo.13593

Budd, G. E., and Mann, R. P. (2019). Survival and selection biases in early animal evolution and a source of systematic overestimation in molecular clocks. *PaleorXiv* [preprint]. doi: 10.31233/osf.io/j45yf

Clarke, J. T., Warnock, R., and Donoghue, P. C. J. (2011). Establishing a time-scale for plant evolution. *New Phytol.* 192, 266–301. doi: 10.1111/j.1469-8137.2011.03794.x

Donoghue, P. C., and Benton, M. J. (2007). Rocks and clocks: calibrating the tree of life using fossils and molecules. *Trends Ecol. Evol.* 22, 424–431. doi: 10.1016/j.tree.2007.05.005

Drummond, A. J., Ho, S. Y., Phillips, M. J., and Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4:e88. doi: 10.1371/journal.pbio.0040088

Dunlop, J. A. (2010). Geological history and phylogeny of Chelicerata. *Arthropod Struct. Dev.* 39, 124–142. doi: 10.1016/j.asd.2010.01.003

Dunlop, J. A., Anderson, L. I., and Braddy, S. J. (2003). A redescription of *Chasmataspis laurencii* Caster and Brooks, 1956 (Chelicerata: Chasmataspidida) from the Middle Ordovician of Tennessee, USA, with remarks on chasmataspid phylogeny. *Earth Environ. Sci. Trans. R. Soc. Edinb.* 94, 207–225. doi: 10.1017/S0263593300000626

Dunlop, J. A., Tetlie, O. E., and Prendini, L. (2008). Reinterpretation of the Silurian scorpion Proscorpius osborni (Whitfield): integrating data from Palaeozoic and Recent scorpions. *Palaeontology* 51, 303–320. doi: 10.1111/j.1475-4983.2007.00749.x

Dunlop, J. A., and Webster, M. (1999). Fossil evidence, terrestrialization and arachnid phylogeny. *J. Arachnol.* 27, 86–93.

Fernández, R., Edgecombe, G. D., and Giribet, G. (2018). Phylogenomics illuminates the backbone of the Myriapoda Tree of Life and reconciles morphological and molecular phylogenies. *Sci. Rep.* 8:83. doi: 10.1038/s41598-017-18562-w

Garrison, N. L., Rodriguez, J., Agnarsson, I., Coddington, J. A., Griswold, C. E., Hamilton, C. A., et al. (2016). Spider phylogenomics: untangling the spider tree of life. *PeerJ* 4:e1719. doi: 10.7717/peerj.1719

Garwood, R. J., and Dunlop, J. (2014). Three-dimensional reconstruction and the phylogeny of extinct chelicerate orders. *PeerJ* 2:e641. doi: 10.7717/peerj.641

Gensel, P. G. (2008). The earliest land plants. *Annu. Rev. Ecol. Evol. S.* 39, 459–477. doi: 10.1146/annurev.ecolsys.39.110707.173526

Glor, R. E. (2010). Phylogenetic insights on adaptive radiation. *Annu. Rev. Ecol. Evol. S.* 41, 251–270. doi: 10.1146/annurev.ecolsys.39.110707.173447

Ho, S. Y. W., and Phillips, M. Y. (2009). Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Syst. Biol.* 58, 367–380. doi: 10.1093/sysbio/syp035

Höhna, S. (2014). Likelihood Inference of non-constant diversification rates with incomplete taxon sampling. *PLoS One* 9:e84184. doi: 10.1371/journal.pone.0084184

Höhna, S. (2015). The time-dependent reconstructed evolutionary process with a key-role for mass-extinction events. *J. Theor. Biol.* 380, 321–331. doi: 10.1016/j.jtbi.2015.06.005

Höhna, S., Landis, M. J., Heath, T. A., Boussau, B., Lartillot, N., Moore, B. R., et al. (2016). RevBayes: bayesian phylogenetic inference using graphical models and an interactive model-specification language. *Syst. Biol.* 65, 726–736. doi: 10.1093/sysbio/syw021

Howard, R. J., Edgecombe, G. D., Legg, D. A., Pisani, D., and Lozano-Fernandez, J. (2019). Exploring the evolution and terrestrialization of scorpions (*Arachnida: Scorpiones*) with rocks and clocks. *Org. Divers. Evol.* 19, 71–86. doi: 10.1007/s13127-019-00390-7

Kenrick, P., Wellman, C. H., Schneider, H., and Edgecombe, G. D. (2012). A timeline for terrestrialization: consequences for the carbon cycle in the Palaeozoic. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 519–536. doi: 10.1098/rstb.2011.0271

Klußmann-Fricke, B. J., and Wirkner, C. S. (2016). Comparative morphology of the hemolymph vascular system in Uropygi and Amblypygi (*Arachnida*): complex correspondences support Arachnopulmonata. *J. Morphol.* 277, 1084–1103. doi: 10.1002/jmor.20559

Kühl, G., Bergmann, A., Dunlop, J., Garwood, R. J., and Rust, J. E. S. (2012). Redescription and palaeobiology of *Palaeoscorpius devonicus* Lehmann, 1944 from the Lower Devonian Hunsrück Slate of Germany. *Palaeontology* 55, 775–787. doi: 10.1111/j.1475-4983.2012.01152.x

Lartillot, N., Lepage, T., and Blanquart, S. (2009). PhyloBayes 3: a bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics* 25, 2286–2288. doi: 10.1093/bioinformatics/btp368

Lartillot, N., and Philippe, H. (2004). A Bayesian mixture model for across-site heterogeneities in the amino- acid replacement process. *Mol. Biol. Evol.* 21, 1095–1109. doi: 10.1093/molbev/msh112

Lartillot, N., Rodrigue, N., Stubbs, D., and Richer, J. (2013). PhyloBayes MPI: phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. *Syst. Biol.* 62, 611–615. doi: 10.1093/sysbio/syt022

Legg, D. A. (2014). *Sanctacaris uncata*: the oldest chelicerate (*Arthropoda*). *Naturwissenschaften* 101, 1065–1073. doi: 10.1007/s00114-014-1245-4

Leite, D. J., Baudouin-Gonzalez, L., Iwasaki-Yokozawa, S., Lozano-Fernandez, J., Turetzek, N., Akiyama-Oda, Y., et al. (2018). Homeobox gene duplication and divergence in arachnids. *Mol. Biol. Evol.* 35, 2240–2253. doi: 10.1093/molbev/msy125

Lepage, T., Bryant, D., Philippe, H., and Lartillot, N. (2007). A general comparison of relaxed molecular clock models. *Mol. Biol. Evol.* 24, 2669–2680. doi: 10.1093/molbev/msm193

Lepage, T., Lawi, S., Tupper, P., and Bryant, D. (2006). Continuous and tractable models for the variation of evolutionary rates. *Math. Biosci.* 199, 216–233. doi: 10.1016/j.mbs.2005.11.002

Linnemann, U., Ovtcharova, M., Schaltegger, U., Gärtner, A., Hautmann, M., Geyer, G., et al. (2019). New high-resolution age data from the Ediacaran–Cambrian boundary indicate rapid, ecologically driven onset of the Cambrian explosion. *Terra Nova* 31, 49–58. doi: 10.1111/ter.12368

Lozano-Fernandez, J., Carton, R., Tanner, A. R., Puttick, M. N., Blaxter, M., Vinther, J., et al. (2016). A molecular palaeobiological exploration of arthropod terrestrialization. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 371:20150133. doi: 10.1098/rstb.2015.0133

Lozano-Fernandez, J., Tanner, A. R., Giacomelli, M., Carton, R., Vinther, J., Edgecombe, G. D., et al. (2019). Increasing species sampling in chelicerate genomic-scale datasets provides support for monophyly of Acari and Arachnida. *Nat. Commun.* 10:2295. doi: 10.1038/s41467-019-10244-7

Magalhaes, I. L., Azevedo, G. H., Michalik, P., and Ramírez, M. J. (2019). The fossil record of spiders revisited: implications for calibrating trees and evidence for a major faunal turnover since the Mesozoic. *Biol. Rev* 95, 184–217. doi: 10.1111/brv.12559

Martin, M. W., Grazhdankin, D. V., Bowring, S. A., Evans, D. A. D., Fedonkin, M. A., and Kirschvink, J. L. (2000). Age of Neoproterozoic bilatarian body and trace fossils, White Sea, Russia: implications for metazoan evolution. *Science* 288, 841–845. doi: 10.1126/science.288.5467.841

Morris, J. L., Puttick, M. N., Clark, J. W., Edwards, D., Kenrick, P., Pressel, S., et al. (2018a). Reply to Hedges et al.: accurate timetrees do indeed require accurate calibrations. *PNAS* 115, E9512–E9513. doi: 10.1073/pnas.1812816115

Morris, J. L., Puttick, M. N., Clark, J. W., Edwards, D., Kenrick, P., Pressel, S., et al. (2018b). The timescale of early land plant evolution. *PNAS* 115, E2274–E2283. doi: 10.1073/pnas.1719588115

Nee, S., Holmes, E. C., May, R. M., and Harvey, P. H. (1994). Extinction rates can be estimated from molecular phylogenies. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 344, 77–82. doi: 10.1098/rstb.1994.0054

Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/molbev/msu300

Parham, J. F., Donoghue, P. C., Bell, C. J., Calway, T. D., Head, J. J., Holroyd, P. A., et al. (2012). Best practices for justifying fossil calibrations. *Syst. Biol.* 61, 346–359. doi: 10.1093/sysbio/syr107

Parry, L. A., Smithwick, F., Nordén, K. K., Saitta, E. T., Lozano−Fernandez, J., Tanner, A. R., et al. (2018). Soft-Bodied fossils are not simply rotten carcasses–toward a holistic understanding of exceptional fossil preservation: exceptional fossil preservation is complex and involves the interplay of numerous biological and geological processes. *BioEssays* 40:1700167. doi: 10.1002/bies.201700167

Penney, D. (2003). Does the fossil record of spiders track that of their principal prey, the insects? *Earth Environ. Sci. Trans. R. Soc. Edinb.* 94, 275–281. doi: 10.1017/S0263593300000675

Pepato, A. R., and Klimov, P. B. (2015). Origin and higher-level diversification of acariform mites–evidence from nuclear ribosomal genes, extensive taxon sampling, and secondary structure alignment. *BMC Evol. Biol.* 15:178. doi: 10.1186/s12862-015-0458-2

Pepato, A. R., Vidigal, T. H., and Klimov, P. B. (2018). Molecular phylogeny of marine mites (*Acariformes*: *Halacaridae*), the oldest radiation of extant secondarily marine animals. *Mol. Phylogenet. Evol.* 129, 182–188. doi: 10.1016/j.ympev.2018.08.012

Puttick, M. N. (2019). MCMCtreeR: functions to prepare MCMCtree analyses and visualize posterior ages on trees. *Bioinformatics* 35, 5321–5322. doi: 10.1093/bioinformatics/btz554

Regier, J. C., Shultz, J. W., Zwick, A., Hussey, A., Ball, B., Wetzer, R., et al. (2010). Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences. *Nature* 463, 1079–1083. doi: 10.1038/nature08742

Roeding, F., Borner, J., Kube, M., Klages, S., Reinhardt, R., and Burmester, T. (2009). A 454 sequencing approach for large scale phylogenomic analysis of the common emperor scorpion (*Pandinus imperator*). *Mol. Phylogenet. Evol.,* 53, 826–834. doi: 10.1016/j.ympev.2009.08.014

Rota-Stabelli, O., Daley, A. C., and Pisani, D. (2013). Molecular timetrees reveal a Cambrian colonization of land and a new scenario for ecdysozoan evolution. *Curr. Biol.* 23, 392–398. doi: 10.1016/j.cub.2013.01.026

Rudkin, D. M., Young, G. A., and Nowlan, G. S. (2008). The oldest horseshoe crab: a new xiphosurid from Late Ordovician Konservat-Lagerstätten deposits. Manitoba Canada. *Palaeontology* 51, 1–9. doi: 10.1111/j.1475-4983.2007.00746.x

Sanders, K. L., and Lee, M. S. Y. (2010). Arthropod molecular divergence times and the Cambrian origin of pentastomids.*Syst. Biodivers.* 8, 63–74. doi: 10.1080/14772000903562012

Scholtz, G., and Kamenz, C. (2006). The book lungs of Scorpiones and Tetrapulmonata (Chelicerata, Arachnida): evidence for homology and a single terrestrialisation event of a common arachnid ancestor. *Zoology* 109, 2–13. doi: 10.1016/j.zool.2005.06.003

Schwentner, M., Combosch, D. J., Nelson, J. P., and Giribet, G. (2017). A phylogenomic solution to the origin of insects by resolving crustacean-hexapod relationships. *Curr. Biol.* 27, 1–7. doi: 10.1016/j.cub.2017.05.040

Selden, P. A., Anderson, H. M., and Anderson, J. M. (2009). A review of the fossil record of spiders (Araneae) with special reference to Africa, and description of a new specimen from the Triassic Molteno Formation of South Africa. *Afr. Invertebr.* 50, 105–116. doi: 10.5733/afin.050.0103

Shao, L., and Li, S. (2018). Early Cretaceous greenhouse pumped higher taxa diversification in spiders. *Mol. Phylogenet. Evol.* 127, 146–155. doi: 10.1016/j.ympev.2018.05.026

Sharma, P. P., Baker, C. M., Cosgrove, J. G., Johnson, J. E., Oberski, J. T., Raven, R. J., et al. (2018). A revised dated phylogeny of scorpions: phylogenomic support for ancient divergence of the temperate Gondwanan family Bothriuridae. *Mol. Phylogenet. Evol.* 122, 37–45. doi: 10.1016/j.ympev.2018.01.003

Sharma, P. P., and Giribet, G. (2014). A revised dated phylogeny of the arachnid order Opiliones. *Front. Genet.* 5:255. doi: 10.3389/fgene.2014.00255

Sharma, P. P., Kaluziak, S. T., Perez-Porro, A. R., Gonzalez, V. L., Hormiga, G., Wheeler, W. C., et al. (2014). Phylogenomic interrogation of Arachnida reveals systemic conflicts in phylogenetic signal. *Mol. Biol. Evol.* 31, 2963–2984. doi: 10.1093/molbev/msu235

Sharma, P. P., and Wheeler, W. C. (2014). Cross-bracing uncalibrated nodes in molecular dating improves congruence of fossil and molecular age estimates. *Front. Zool.* 11:57. doi: 10.1186/s12983-014-0057

Signor, P. W., and Lipps, J. H. (1982). Sampling bias, gradual extinction patterns, and catastrophes in the fossil record. *Spec. Pap. Geol. Soc. Am.* 190, 291–296. doi: 10.1130/SPE190-p291

Stadler, T. (2011). Mammalian phylogeny reveals recent diversification rate shifts. *PNAS* 108, 6187–6192. doi: 10.1073/pnas.1016876108

Van Roy, P., Orr, P. J., Botting, J. P., Muir, L. A., Vinther, J., Lefebvre, B., et al. (2010). Ordovician faunas of Burgess Shale type. *Nature* 465, 215–218. doi: 10.1038/nature09038

von Reumont, B. M., Jenner, R. A., Wills, M. A., Dell'Ampio, E., Pass, G., Ebersberger, I., et al. (2012). Pancrustacean phylogeny in the light of new phylogenomic data: support for Remipedia as the possible sister group of Hexapoda. *Mol. Biol. Evol.* 29, 1031–1045. doi: 10.1093/molbev/msr270

Waloszek, D., and Dunlop, J. A. (2002). A larval sea spider (*Arthropoda*: *Pycnogonida*) from the upper Cambrian 'Orsten' of Sweden, and the phylogenetic position of pycnogonids. *Palaeontology* 45, 421–446. doi: 10.1111/1475-4983.00244

Weygoldt, P., and Paulus, H. F. (1979). Untersuchungen zur Morphologie, Taxonomie und Phylogenie der Chelicerata: II. Cladogramme und die Entfaltung der Chelicerata. *Z. Zool. Syst. Evol.* 17, 177–200. doi: 10.1111/j.1439-0469.1979.tb00699.x

Wheat, C. W., and Wahlberg, N. (2013). Phylogenomic insights into the Cambrian explosion, the colonization of land and the evolution of flight in Arthropoda. *Syst Biol.* 62, 93–109. doi: 10.1093/sysbio/sys074

Wolfe, J. M., Daley, A. C., Legg, D. A., and Edgecombe, G. D. (2016). Fossil calibrations for the arthropod Tree of Life. *Earth-Sci. Rev.* 160, 43–110. doi: 10.1016/j.earscirev.2016.06.008

Zhang, Z. Q. (2013). Phylum Arthropoda. *Zootaxa* 3703, 17–26. doi: 10.11646/zootaxa.3703.1.6

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Lozano-Fernandez, Tanner, Puttick, Vinther, Edgecombe and Pisani. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genetics

ORIGINAL RESEARCH
published: 20 March 2020
doi: 10.3389/fgene.2020.00252

Quantifying the Error of Secondary vs. Distant Primary Calibrations in a Simulated Environment

Christopher Lowell Edward Powell[1], *Sydney Waskin*[1] *and Fabia Ursula Battistuzzi*[1,2]*

[1] Department of Biological Sciences, Oakland University, Rochester, MI, United States, [2] Center for Data Science and Big Data Analytics, Oakland University, Rochester, MI, United States

Using calibrations to obtain absolute divergence times is standard practice in molecular clock studies. While the use of primary (e.g., fossil) calibrations is preferred, this approach can be limiting because of their rarity in fast-growing datasets. Thus, alternatives need to be explored, such as the use of secondary (molecularly-derived) calibrations that can anchor a timetree in a larger number of nodes. However, the use of secondary calibrations has been discouraged in the past because of concerns in the error rates of the node estimates they produce with an apparent high precision. Here, we quantify the amount of errors in estimates produced by the use of secondary calibrations relative to true times and primary calibrations placed on distant nodes. We find that, overall, the inaccuracies in estimates based on secondary calibrations are predictable and mirror errors associated with primary calibrations and their confidence intervals. Additionally, we find comparable error rates in estimated times from secondary calibrations and distant primary calibrations, although the precision of estimates derived from distant primary calibrations is roughly twice as good as that of estimates derived from secondary calibrations. This suggests that increasing dataset size to include primary calibrations may produce divergence times that are about as accurate as those from secondary calibrations, albeit with a higher precision. Overall, our results suggest that secondary calibrations may be useful to explore the parameter space of plausible evolutionary scenarios when compared to time estimates obtained with distant primary calibrations.

Keywords: molecular clocks, secondary calibrations, simulation, divergence times, timetree, confidence interval

OPEN ACCESS

Edited by:
Michel Laurin, UMR 7207 Centre de Recherche sur la Paléobiodiversité et les Paléoenvironnements (CR2P), France

Reviewed by:
David Buckley, Autonomous University of Madrid, Spain
David Marjanović, Museum für Naturkunde Leibniz Institut für Evolutions und Biodiversitätsforschung, Germany

***Correspondence:**
Fabia Ursula Battistuzzi
battistu@oakland.edu

Specialty section:
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Genetics

Received: *31 October 2019*
Accepted: *02 March 2020*
Published: *20 March 2020*

Citation:
Powell CLE, Waskin S and Battistuzzi FU (2020) Quantifying the Error of Secondary vs. Distant Primary Calibrations in a Simulated Environment. Front. Genet. 11:252. doi: 10.3389/fgene.2020.00252

INTRODUCTION

The use of calibrations to estimate absolute times in a phylogeny is a source of controversy for many reasons; among these are that (i) few are available from independent sources (e.g., fossil record), (ii) their phylogenetic placement can be incorrect, especially in cases of uncertain fossil identification or phylogenetic position, and that (iii) calibration constraints (and the internal distributions between them) are heavily debated, although new methods to estimate probability densities of node ages are being developed (Marjanović and Laurin, 2007, 2008; Ho and Phillips, 2009; Inoue et al., 2010; Sauquet et al., 2012; Sterli et al., 2013; Heath et al., 2014; Hipsley and Müller, 2014; Warnock et al., 2015; dos Reis et al., 2016; Kumar and Hedges, 2016; Didier et al., 2017; Didier and Laurin, 2018; Bromham, 2019; Marshall, 2019). Despite these issues, molecular clock analyses cannot avoid using calibrations if absolute time estimates are the ultimate

goal. Alternative approaches that have been explored include the direct use of substitution rates to estimate divergence times, but these present other challenges such as their applicability when, in many cases, these rates are measured from laboratory-grown strains and may not accurately represent rates of strains in the environment (Ho, 2007; Hipsley and Müller, 2014). A recently developed new strategy, instead, suggests the use of horizontal gene transfer events as time-calibrated constraints in a phylogeny (Magnabosco et al., 2018; Wolfe and Fournier, 2018). While this method holds promise, it currently has not been applied widely and the reconstruction of horizontal gene transfer events useful as calibration constraints is a challenging endeavor as it relies on clear gene exchange patterns between groups, one of which should also have primary calibrations. Thus, challenges for these alternative calibration strategies are equally difficult to overcome and, for now, limited improvements have been made. Therefore, the process of assigning time constraints to some nodes in a phylogeny still remains the primary source of information to obtain absolute node times, despite the unbalance between the amount of information (i.e., number of nodes that can be calibrated) available and what would be required to obtain accurate time estimates. This unbalance between availability and need has led researchers to test a variety of alternatives to meet the new needs by increasing calibration sources. These new strategies are especially important in very large phylogenies due to the increase in the ratio of unknown to calibrated nodes and, thus, the potential increase in error propagation of estimates for nodes that are far from a calibration point often caused by rate variation among branches (Britton, 2005; Hug and Roger, 2007; Perie and Doyle, 2012; dos Reis and Yang, 2013; Hipsley and Müller, 2014).

Some examples of alternative strategies to improve the number and quality of available calibrations include (i) the definition of boundaries is either based on the time boundaries of the geologic stratum in which the fossil is found or based on an accurate phylogenetic placement and timing of extinct taxa (Marjanović and Laurin, 2007; Marshall, 2008; Sterli et al., 2013; Warnock et al., 2017) (ii) the selection of representative taxa, which effectively decreases the unknown/calibrated node ratio (Britton, 2005; Perie and Doyle, 2012), and (iii) the use of secondary calibrations (Sauquet et al., 2012; Hipsley and Müller, 2014; Schenk, 2016). In this study we focus on secondary calibrations, which are molecular time estimates obtained from previous molecular clock analyses that were calibrated using independent evidence (primary calibrations). The strongest advantage of using derived (e.g., secondary, tertiary) calibrations is that it effectively provides an infinite source of calibration constraints, only constrained by the number of steps (nodes) a researcher chooses between calibrated and unknown nodes. However, this advantage is dependent on one fundamental question: does the use of derived calibrations result in accurate timetree estimates?

Two past studies have addressed this question with simulation and empirical data analyses, using both Bayesian and maximum-likelihood-based methods (Sauquet et al., 2012; Schenk, 2016). Both found similar outcomes, with secondary calibration estimates being younger than expected and with overly narrow confidence intervals leading to small uncertainties around inaccurate estimates. These results supported previous evidence against the use of secondary calibrations reinforcing the practice of avoiding them, if possible (e.g., Graur and Martin, 2004; Reisz and Müller, 2004; Hug and Roger, 2007; Hipsley and Müller, 2014). However, questions about the performance of secondary calibrations remain. For example, how is the error from primary calibrations compounded into estimates based on secondary calibrations? Can the performance of secondary calibrations be predicted based on uncertainties in the primary calibrations? Is the performance of secondary calibrations worse than that of a small number of primary calibrations that are phylogenetically distant from the nodes of interest (distant primary calibrations)?

To address these questions, we designed a simple simulated scenario with the aim of testing if there are predictable patterns when secondary calibrations are used. For this purpose we used RelTime to estimate times because of its minimal assumption requirements and speed of analyses (Tamura et al., 2012, 2018). In our scenario we used two nested trees that share one overlapping ingroup node that is used as the secondary calibration. In one of the two trees we also selected three nodes to act as primary calibrations. These were used with increasingly larger uncertainties in their boundaries (from 0 to 20% departure from the true, simulated time) and biases (either balanced around the true time or skewed toward younger or older times). Based on the results from previous studies, we expected that estimates based on secondary calibrations would be consistently and precisely underestimated relative to the use of primary calibrations and the true times. Instead, we found that secondary time estimates are generally overestimated by approximately 10% but with low precision (large confidence intervals) and with overall patterns that are clearly predictable. These results suggest that our understanding of the accuracy of secondary calibrations is still incomplete and more comprehensive testing is required to determine their effect if used in empirical datasets.

MATERIALS AND METHODS

Dataset

We started from a main tree of 248 species represented in a tree of life (Hedges and Kumar, 2009). This main tree was split into two subtrees (**Figure 1**), tree A (173 species) and tree B (71 species), that represent two clades and maximize the size of the dataset in each tree (see **Supplementary Data Sheet S5** for NWK formatted tree files). We then added to these clades two shared lineages which were arbitrarily chosen and an outgroup. This setup created two nested phylogenies that were used to test hypotheses on the performance of the calibrations. To simulate multiple genes, we used a set of 446 empirical parameters (e.g., length, GC content, initial evolutionary rate) and altered the main timetree according to an autocorrelated model ($\nu = 1$) that resulted in estimated rates of up to $\pm$ 25% of the mean rate (Thorne and Kishino, 2002; Rosenberg and Kumar, 2003) (**Supplementary Presentation S1** and **Supplementary Figure S1**). This effectively created 446 phylogenies with different branch lengths but same topology. These parameters were given to SeqGen to simulate genes under a Hasegawa-Kishino-Yano

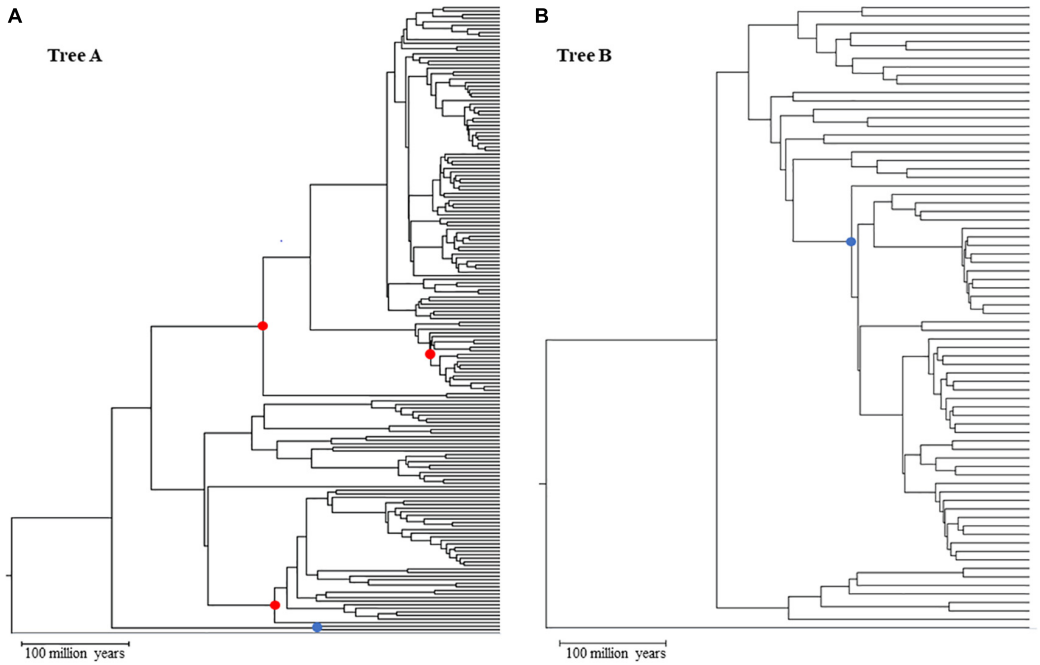


FIGURE 1 | Topologies used in simulated analyses for Tree A **(A)** and Tree B **(B)** red dots primary calibration nodes; blue dot: overlapping node between trees A and B used as secondary calibration in tree **(B)**. Gray branch outgroup.

(HKY) model (Rambaut and Grassly, 1997). Ten random sets of individual genes were then concatenated to reach a length of at least 30,000 sites (30,029–30,725). In addition, we also created one concatenation with all genes (approximately 604,000 sites) and two concatenations of half the number of genes (223 genes per concatenation) with lengths of 273,812 and 330,187. Each of these concatenations were used independently in downstream analyses. Patterns between the 30k, half, and full concatenations were similar. Therefore, we discuss results from the 30k concatenations because they allow us to evaluate the variance of estimates among datasets. For primary calibrations, three nodes from tree A were chosen: a relatively shallow node at 63.9 million years ago (mya), and two that were deeper in the tree but in two different clades (209.4 and 220.2 mya). The overlapping node between tree A and B has an intermediate depth (167 mya) within tree A and is centrally placed within the topology of tree B (**Figure 1**). These primary and secondary calibrations were chosen to minimize the effect that biased location (e.g., all within one clade) and divergence times (e.g., all young nodes) may have on the accuracy of estimations.

Time Estimation

Time estimates for each concatenation were calculated using RelTime as implemented in the command-line version of MEGA X (Kumar et al., 2018). Each analysis was run on the Michigan State University HPCC-ICER cluster using a HKY model, uniform rates among sites, all sites, a maximum likelihood estimator, and local clocks. Our goal was to explore the accuracy of time estimates for the nodes in tree B when different types of calibration were used. Thus, we used three combinations of calibrations, all with minimum and maximum constraints: tree A with three primary calibrations (**Figure 1**, red dots) + tree B with one secondary calibration (this is the overlapping node between tree A and B; **Figure 1**, blue dot) [B_secondary]; tree B with one primary calibration (same node as the previous combination) (**Figure 1**, blue dot) [B_primary]; tree AB (the combination of trees A and B) with the same three primary calibrations used in tree A (**Figure 1**, red dots) which are distant from the nodes in tree B [B_distant_primary]. Additionally, each of these combinations was tested for seven different scenarios that were meant to account for increasing uncertainty on primary calibrations: three of these scenarios had increasingly larger uncertainties, from 0 to 20%, but spread evenly around the true time [0 balanced (0B), 10 balanced (10B), and 20 balanced (20B)]; two scenarios had the error skewed toward younger times [10 low (10L) and 20 low (20L)] and two with the error skewed toward older times [10 high (10H) and 20 high (20H)]. This set up allowed us to test if and how the error in primary calibrations propagates to estimates based on secondary calibrations. Using a subset of our dataset, we also compared our results obtained from tree A using RelTime to results obtained from MCMCTree and found the two to be highly correlated (less than 3% different) (Yang, 2007; Rambaut et al., 2018). Additionally, we tested the potential effect of using a non-partitioned concatenation on the accuracy of the RelTime results by re-analyzing one of our concatenations, partitioning it by gene. Also, in this case, the results are comparable (less than 3% difference), suggesting that the use of a single partition is not biasing the results (see **Supplementary Material**).

Measures of Accuracy

We assessed the outcomes of our molecular clock analyses with a number of measures. *First*, we measured the percent departure of the estimated times (ET) from the known (simulated) true times (ET accuracy). *Second*, we calculated the frequency of confidence intervals (CIs) that included the true times (TT) (CI accuracy). *Third*, we calculated the range of each CI normalized to the depth of the node [(maximum–minimum boundary)/TT] (CI precision). *Fourth*, we measured the distance of each CI boundary to the true time [| (CI boundary – TT)/TT] (CI skewness). We applied the latter measure only to the overlapping node between tree A and B to evaluate the effect of skewness on estimates based on secondary calibrations. Each of these measures was applied to all analyses. We report the averages across the 10 concatenations with ± 1 standard deviation in parenthesis because no significant difference was detected among individual concatenations (see **Supplementary Data Sheets S1–S4** for concatenation specific estimates).

TABLE 1 | Scenarios of varying uncertainty around the true time (TT) for primary calibration boundaries.

Scenario	Calibration boundaries	
	Minimum time	Maximum time
0B	−1 my	+1 my
10B	−5% of TT	+5% of TT
20B	−10% of TT	+10% of TT
10L	−10% of TT	+1 my
10H	−1 my	+ 10% of TT
20L	−20% of TT	+1 my
20H	−1 my	+20% of TT

Each scenario was applied to all three primary calibrations at once and run independently from the other scenarios. The uncertainty associated with the boundaries varies from a minimum of 0% (±1 million year of the true time) to a maximum of 20%. my: million years.

RESULTS

Generally, secondary calibrations are used when primary calibrations are not available or are believed to be too few to provide accurate estimates on nodes that are distantly related. While many studies have used this approach as a last-resort method, a systematic evaluation of their performance might open up the use of secondary calibration to more (and larger) phylogenies, thus expanding the applicability of molecular clocks to complex datasets. To investigate and quantify the amount of error associated with secondary calibration, we used a simulated approach by creating two nested trees (A and B) in which one node estimate from primary calibrations in A is used as a secondary calibration in B. We then quantified the accuracy of estimated times (ET) vs. simulated true times (TT) and the properties of the confidence intervals (CIs) in both trees relative to the type of calibration used. Each analysis was repeated for a series of scenarios with variable levels of uncertainty in the primary calibrations to investigate how estimates derived from secondary calibrations may be affected (see section "Materials and Methods").

In practice, we simulated 100s of nucleotide alignments with a variety of empirically derived parameters (length, initial evolutionary rate, transition/transversion ratio, GC content; **Supplementary Figure S1**) for a phylogeny of 248 lineages that was split into two nested trees (A with 176 species and B with 74 species). We used variable numbers of genes in concatenation to obtain the alignments used to estimate divergence times. On three nodes in A, we applied primary calibrations with varying levels of uncertainty around their true time (see **Table 1**). Then, the CI of the molecular time estimate for the common node between A and B was used as secondary calibration for tree B.

Our evaluations of time estimates' accuracy rely on two basic measures: similarity of the estimated time to the simulated true time (ET accuracy) and the general properties of the CIs (accuracy, precision, skewness). In addition, we also compared estimated times from secondary calibrations to those obtained from B_distant_primary calibrations when the two trees were combined.

Assessing Accuracy of Primary Calibrations

The first step was to measure the trends in accuracy of estimated times and CIs from primary calibrations in tree A. On average, the ET differed from TT by approximately 3% with a tendency to underestimate the results (average of slopes = 0.97, SD = 0.06) which reflects a generally higher confidence on younger (minimum) calibration boundaries than older (maximum) ones (**Figure 2**, black triangles; **Supplementary Figure S2** and **Supplementary Table S1**). The observed trends in accuracy were as expected relative to the error associated with calibration boundaries: balanced calibrations (0, 10, and 20B) performed the best, boundaries skewed toward younger times (10 and 20L) produced underestimated times and boundaries skewed toward older times (10 and 20H) produced overestimated times.

On average, the CI ranges included the true time in 87% of the cases with the 0B scenario being the most likely to fail (78% of CIs include the TT) (**Table 2**). This is expected as narrower calibration boundaries generate narrower CI ranges that are less likely to include the TT in light of the underestimated divergence times. Similarly, in an average of 87% of all the cases the CI included the true time for the node overlapping in trees A and B and, in those cases in which the TT was not included, the minimum boundary was < 5% older than the TT (the maximum boundary was never younger than the TT). In only one scenario, 20H, the difference between TT and minimum boundary was 6.5%, which is expected given the large uncertainty skewed toward deeper times. Moreover, in 93% of the cases the CI of the overlapping node was skewed toward older times which means that, when used as a secondary calibration, this node is expected to behave similarly to the "high" simulated scenarios (see below). Overall, the trends observed from different calibration scenarios in tree A are as expected and provide an accurate basis for the estimation of times using a secondary calibration.

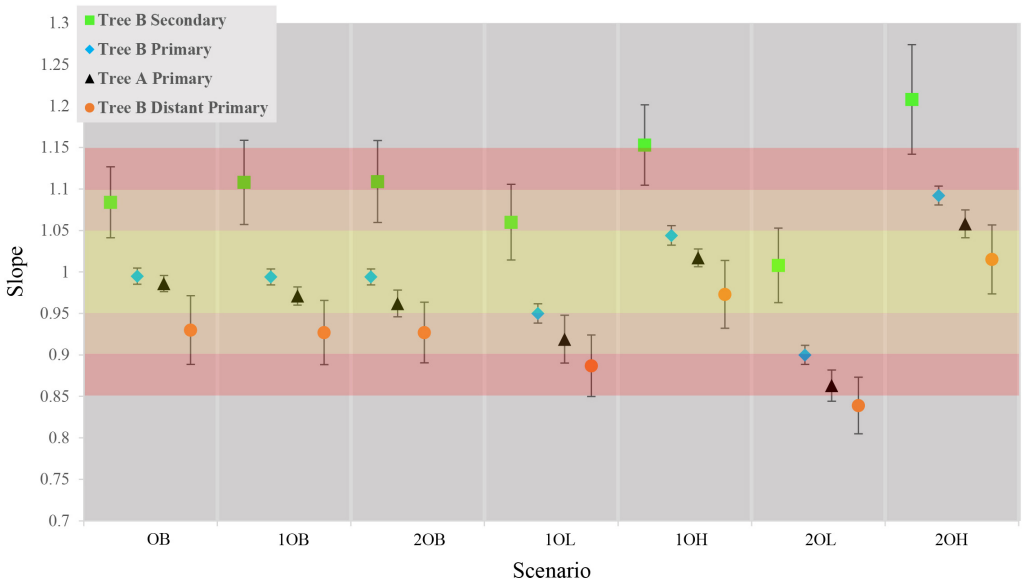


FIGURE 2 | Comparison of average slopes for each of the seven scenarios in the four callibration setting. Gradient of color represents the accuracy of the estimates yellow is ± 5%, Orange ± 10% red = ± 15%. Each data point represents the average slope of the 10 concatensions of true time vs. estimate time with ±1 standard deviation.

Assessing Accuracy of Secondary Calibrations

Using the CI estimated for the overlapping node between tree A and tree B, we obtained secondary divergence times and measured the accuracy of these secondary node ages relative to true times and to estimated times from primary calibrations on distant nodes (we refer to this scenario as B_distant_primary). The first measure is unrealistic in real cases but allows us to quantify the overall error produced by secondary calibrations relative to the true times, while the second measure leads to a quantification of the error introduced by secondary calibrations compared to distant primary ones.

TABLE 2 | Confidence intervals (CIs) accuracy (proportion of CIs that include the simulated true time) for each of the seven scenarios.

Scenario	CI accuracy			
	Tree A primary	Tree B secondary	Tree B distant primary	Tree B primary
0B	0.78 (0.1495)	1.0 (0.0044)	0.98 (0.0222)	0.91 (0.0287)
10B	0.87 (0.1333)	1.0 (0.0044)	0.97 (0.0306)	1.0 (0.0048)
20B	0.95 (0.0704)	1.0 (0.0044)	0.98 (0.0181)	1.00 (0.0000)
10L	0.86 (0.0608)	1.0 (0.000)	0.89 (0.1139)	0.99 (0.0000)
10H	0.84 (0.1590)	1.0 (0.0059)	0.99 (0.0097)	0.99 (0.0125)
20L	0.88 (0.1223)	1.00 (0.0000)	0.85 (0.1528)	1.0 (0.0042)
20H	0.89 (0.1658)	0.99 (0.0240)	1.0 (0.0014)	0.99 (0.0097)
Average	0.87 (0.0525)	1.0 (0.0044)	0.95 (0.0589)	0.98 (0.0313)

Values shown are averages of all 30k concatenations with 1 standard deviation in parenthesis.

Contrary to previous studies, our results show an average 10% (±6%) overestimation of molecular time estimates against true times with the strongest overestimation in the scenarios with the largest inaccuracy on the maximum boundary (10, 20H) (**Figure 2**, green squares; **Supplementary Figure S3** and **Supplementary Table S1**). This is predicted from the CI boundaries of the secondary calibrations that are most strongly skewed toward older times in the "high" scenarios. Interestingly, the amount of inaccuracy produced by secondary calibrations is comparable to the average 7% (±6%) departure of the estimates from the true times produced by distant primary calibrations, although in these cases the node ages are generally underestimated (**Figure 2**, orange circles; **Supplementary Figures S4, S5** and **Supplementary Table S1**).

Despite the overestimation produced by the secondary calibrations, > 99% of CIs include the true time (**Table 2**). The high probability of the true time being included in the CI of each node is due to the large CI range estimates (78–95% of the true time) (**Table 3**). This is approximately double the size of the CIs obtained from distant primary calibrations (42–53% of the true time) and from primary calibrations. The larger size of the CIs when secondary calibrations are used reflects the larger uncertainty in the calibrating range. Indeed, while the primary calibrations were allowed to have at most a 20% uncertainty, the CIs of the node used as secondary calibration have, on average,

TABLE 3 | Confidence interval (CI) precision relative to the simulated true time.

Scenario	CI Precision			
	Tree A primary	Tree B secondary	Tree B distant primary	Tree B primary
0B	0.29 (0.0507)	0.78 (0.1409)	0.45 (0.1420)	0.27 (0.0548)
10B	0.32 (0.0606)	0.78 (0.1740)	0.43 (0.0826)	0.38 (0.0516)
20B	0.45 (0.0614)	0.89 (0.1756)	0.49 (0.0943)	0.51 (0.0474)
10L	0.32 (0.0543)	0.78 (0.1626)	0.42 (0.0777)	0.37 (0.0487)
10H	0.33 (0.0655)	0.82 (0.1867)	0.45 (0.0886)	0.40 (0.0540)
20L	0.44 (0.0567)	0.82 (0.1316)	0.48 (0.0827)	0.50 (0.0353)
20H	0.47 (0.0670)	0.95 (0.1780)	0.53 (0.1045)	0.54 (0.0525)
Average	0.37 (0.0749)	0.83 (0.0065)	0.46 (0.0039)	0.42 (0.0955)

Averages of the 30k concatenations with 1 standard deviation are shown in parentheses.

double that amount, thus producing twofold larger CIs in the estimated nodes.

It is possible that the results obtained in the estimation of node ages for tree B with secondary calibrations could have been driven by anomalies in branch lengths and evolutionary rates specific to this phylogeny rather than the type of calibration used. To identify these potentially confounding factors we first applied a primary calibration on the same node that was used as secondary. If the location and branch lengths associated with the calibration node were biasing the results, the accuracy of estimated times should have been lower even with primary calibrations. Instead, we observed similar accuracy and precision to what was obtained with three calibrations in tree A [**Figure 2**, blue diamonds; **Supplementary Figure S6** and **Supplementary Table S1**; ET on average within 5% of TT, >98% of CIs include the TT (**Table 2**), and the CI precision is approximately 40% (**Table 3**)]. Then, we compared the relative times (before calibrations are applied) of nodes represented only in tree B, only in tree A, and in the combined tree AB. If trees A and B differed substantially in the relative rates of their branches, we would expect that the relative times would not be comparable to those obtained with the combined tree AB. Again, we found the opposite result, which suggests that evolutionary rates do not differ significantly between trees (**Supplementary Figure S7**). These results show that the estimated times obtained are driven primarily by the choice of calibration and that, therefore, they are a valid measure of calibration performance.

These results show predictable trends for the estimates of secondary calibrations that closely mirror the uncertainty of the primary calibrations. Additionally, they show that, at least in this simulated scenario, the absolute accuracy of using a secondary calibration is similar to that of using distant primary calibrations, although the precision is approximately half.

DISCUSSION

Despite the potentially broad applications of secondary calibrations, their use has been hindered by concerns over: (i) the process of implementation of time uncertainties from primary calibrations, (ii) the predictability of their performance, and (iii) their overall accuracy and precision. A few studies in the past 15 years have evaluated these three points and generally agreed that secondary calibrations produce systematically biased but precise estimates, effectively attributing to secondary calibrations the worst kind of error: wrongly precise (Graur and Martin, 2004; Hug and Roger, 2007; Sauquet et al., 2012; Hipsley and Müller, 2014; Schenk, 2016). However, some key aspects of secondary calibration assessment are still missing, such as a systematic analysis of how errors in primary calibrations are compounded in the estimates from secondary calibrations and how secondary vs. primary but distant calibrations perform relative to true (simulated) times. Understanding these key aspects would allow us to determine if, and under which conditions, secondary calibrations might produce informative results.

However, because absolute time estimates are the result of the entanglement of evolutionary rates, branch lengths, and calibrations, another fundamental property of a molecular clock assessment analysis is being able to identify the source of observed errors and, if possible, predict the behavior of model parameterizations based on specific scenarios. This approach can be difficult in methods, such as Bayesian, that produce estimates based on many interacting priors. Instead, a theoretically more straightforward approach, such as RelTime, allows to analyze each parameter independently and isolate the source or sources of errors.

Using this approach, we applied secondary calibrations to a suite of simulated alignments with the goal of analyzing three aspects: (i) the overall accuracy of secondary time estimates compared to true times, (ii) the relative accuracy of secondary vs. primary time estimates, (iii) the trends in the errors for the secondary time estimates relative to uncertainties in primary calibrations. By using the same substitution model and topology as in the simulations, we limited issues from phylogenetic uncertainty, and by using uniform, flat distributions for the calibrations we minimized the need to account for decreasing probabilities in the tails of non-uniform calibration distributions. Despite this, our study design has some limitations and caveats that should be taken into consideration when interpreting our results. For example, the use of uniform distributions is not common in empirical data. In real data, calibration constraints are considered more likely to be close to the earliest known fossil evidence for the lineage, thus favoring the use of lognormal, normal, or exponential distributions that are expected to weigh estimates toward younger times (Hedges and Kumar, 2003; Ho and Duchêne, 2014; Ware and Barden, 2016; Didier and Laurin, 2018). Given our overestimated times in the simulations, using lognormal or exponential distributions would likely improve the performance of secondary calibrations. Thus, our use of uniform distributions was more conservative (more likely to highlight estimation biases produced by secondary calibrations). Second, all our primary calibrations have minimum and maximum

constraints. Molecular clock methods are known to perform better when both boundaries are provided but this is often not possible in empirical data analysis (Marjanović and Laurin, 2007; Parham et al., 2012; Warnock et al., 2012, 2017). Because one of our goals was to evaluate error propagation from primary to secondary calibrations, providing min-max boundaries allowed us to simulate the exact amount of uncertainties in calibrations and, thus, to track their effect on derived time estimates. The predictable correlation between errors in primary and secondary calibrations could be used to investigate the effect of removing one boundary on a primary node.

The measures we used to determine the effects of the use of secondary calibrations are the typical ones of molecular clock assessment studies: the similarity of true times and estimated times, the frequency of CIs that include the true time, and the precision of CI (their range relative to the age of the node). In addition to these, we also considered the relative error (TT vs. ET) and CI precision of secondary calibrations vs. distant primary calibrations and the predictability of estimates based on secondary calibrations based on the simulated scenario. Surprisingly, our results are opposite to those found by two recent studies: our results show that estimates based on secondary calibrations are, in general, overestimated (by approximately 10%) with poor precision (large CIs). While these results are far from optimal, they show that our understanding of estimates based on secondary calibrations is still incomplete and that their dismissal might be premature. Perhaps more interestingly, we also found that the magnitude of the error in estimates based on secondary calibrations is approximately the same as that produced by the use of distant primary calibrations but in the opposite direction (secondary calibrations overestimate, distant primary ones underestimate). This result is significant because one of the strategies commonly adopted to avoid using secondary calibrations is to increase the dataset size to obtain one or more primary calibrations (Perie and Doyle, 2012). These will inevitably be far away, in the phylogenetic sense, from the nodes of interest, potentially leading to errors, as we see in our simulations. However, an advantage of using distant primary calibrations would be two-fold higher precision (narrower CI ranges) that does not come at the expense of a lower probability of including the true time. It should be noted that in our simulated scenario primary calibrations are given with minimum and maximum boundaries which are expected to increase accuracy and precision. It is possible that the precision of distant primary calibrations would be negatively affected in empirical analyses that do not use maximum constraints (Marjanović and Laurin, 2007, 2008; Marshall, 2019).

A deeper analysis of estimates from proximal primary, distant primary, and secondary calibrations can explain the trends observed. First, the large CI ranges from secondary calibrations are caused by the large uncertainty in the boundaries derived from the primary calibrations. Indeed, these boundaries include a 30–40% error, which is almost double the maximum amount of error assigned to primary calibrations. Therefore, these results suggest that estimates based on secondary calibrations incorporate the error present in the primary estimate in addition to their own. Second, the directionality of the secondary estimated errors (over- or underestimation) is also clearly dependent on the skewness of the primary calibration boundaries. For example, in our simulated cases, we saw that CIs for the overlapping node were almost always skewed toward older times by approximately 30%, driving the observed overestimation. Thus, careful choice of accurate primary calibrations is key when secondary calibrations are to be used. Unfortunately, errors associated with primary calibrations are often unknown in empirical data (but see Didier and Laurin, 2018); but knowing that these errors are included in the estimates based on secondary calibrations with predictable trends makes it possible to test the plausibility of different evolutionary hypotheses based on what is known of the primary calibrations.

A question that remains open is why our results differ so strikingly from those of previous studies. While additional analyses will be necessary to provide an answer, a few hypotheses are possible: first, previous studies used primarily Bayesian methods that are known to depend strongly on priors (dos Reis and Yang, 2013; dos Reis et al., 2016). It is possible that the priors affected the results; but the magnitude of this effect, if present, is unknown. Second, the two previous studies did not take into consideration the error associated with the primary calibrations. In one case (Schenk, 2016) it was not simulated and in the other (Sauquet et al., 2012) it was not known because an empirical dataset was used. From the analyses of multiple scenarios (from many different studies) that the authors carried out it is obvious that the youngest estimates based on secondary calibrations are obtained when only young proximal primary calibrations are used, which is the same trend we observe in our analyses and in other, earlier, studies (Brochu, 2004a,b).

Overall, our simulation study shows that our current understanding of the performance of secondary calibrations is still incomplete and that their dismissal from implementation in favor of other solutions (e.g., distant primary) might not produce the desired increase in accuracy. Additionally, our results show that the performance of secondary calibrations can be predicted based on the uncertainty of the primary calibrations (in our case the different scenarios). Thus, secondary calibrations can be used as a testing tool for different evolutionary scenarios. Therefore, we suggest that rather than avoiding secondary calibrations, they should be used and compared with distant primary ones to test the limits of the parameter space of plausible evolutionary scenarios of divergence times.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**. All simulated sequences have been uploaded on Dryad (doi: 10.5061/dryad.1zcrjdfp5).

AUTHOR CONTRIBUTIONS

CP, SW, and FB designed the project. CP and SW performed the analyses. CP and FB wrote the manuscript.

FUNDING

Funding to support this research was provided by NASA (NNX16AJ30G) and Michigan Space Grant Consortium to FB.

ACKNOWLEDGMENTS

We thank the two reviewers and the editor for the many valuable comments on the previous versions of this manuscript. We thank Yagya Sharma, Alex L'Espérance, Brandon Hanna, Deeksha Pai, and Alyson Light for their contribution to an early version of this project.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.00252/full#supplementary-material

REFERENCES

Britton, T. (2005). Estimating divergence times in phylogenetic trees without a molecular clock. *Syst. Biol.* 54, 500–507. doi: 10.1080/10635150590947311

Brochu, C. A. (2004a). Calibration age and quartet divergence date estimation. *Evolution* 58, 1375–1382. doi: 10.1111/j.0014-3820.2004.tb01715.x

Brochu, C. A. (2004b). Patterns of calibration age sensitivity with quartet dating methods. *J. Paleontol.* 78, 7–30. doi: 10.1666/0022-33602004078<0007:pocasw<2.0.co;2

Bromham, L. (2019). Six impossible things before breakfast: assumptions, models, and belief in molecular dating. *Trends Ecol. Evol.* 34, 474–486. doi: 10.1016/j.tree.2019.01.017

Didier, G., Fau, M., and Laurin, M. (2017). Likelihood of tree topologies with fossils and diversification rate estimation. *Syst. Biol.* 66, 964–987. doi: 10.1093/sysbio/syx045

Didier, G., and Laurin, M. (2018). Exact distribution of divergence times from fossil ages and tree topologies. *bioRxiv* [Preprint]. doi: 10.1101/490003

dos Reis, M., Donoghue, P. C. J., and Yang, Z. (2016). Bayesian molecular clock dating of species divergences in the genomics era. *Nat. Rev. Genet.* 17, 71–80. doi: 10.1038/nrg.2015.8

dos Reis, M., and Yang, Z. (2013). The unbearable uncertainty of Bayesian divergence time estimation. *J. Syst. Evol.* 51, 30–43. doi: 10.1111/j.1759-6831.2012.00236.x

Graur, D., and Martin, W. (2004). Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision. *Trends Genet.* 20, 80–86. doi: 10.1016/j.tig.2003.12.003

Heath, T. A., Huelsenbeck, J. P., and Stadler, T. (2014). The fossilized birth-death process for coherent calibration of divergence-time estimates. *Proc. Natl. Acad. Sci. U.S.A.* 111, E2957–E2966. doi: 10.1073/pnas.1319091111

Hedges, S. B., and Kumar, S. (2003). Genomic clocks and evolutionary timescales. *Trends Genet.* 19, 200–206. doi: 10.1016/S0168-9525(03)00053-2

Hedges, S. B., and Kumar, S. (2009). *The Timetree of Life*, 1st Edn, eds S. B. Hedges, and S. Kumar. New York, NY: Oxford University Press.

Hipsley, C. A., and Müller, J. (2014). Beyond fossil calibrations: realities of molecular clock practices in evolutionary biology. *Front. Genet.* 5:138. doi: 10.3389/fgene.2014.00138

Ho, S. Y. M. (2007). Calibrating molecular estimates of substitution rates and divergence times in birds. *J. Avian Biol.* 38, 409–414. doi: 10.1111/j.0908-8857.2007.04168.x

Ho, S. Y. W., and Duchêne, S. (2014). Molecular-clock methods for estimating evolutionary rates and timescales. *Mol. Ecol.* 23, 5947–5965. doi: 10.1111/mec.12953

Ho, S. Y. W., and Phillips, M. J. (2009). Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Syst. Biol.* 58, 367–380. doi: 10.1093/sysbio/syp035

Hug, L. A., and Roger, A. J. (2007). The Impact of fossils and taxon sampling on ancient molecular dating analyses. *Mol. Biol. Evol.* 24, 1889–1897. doi: 10.1093/molbev/msm115

Inoue, J., Donoghue, P. C. J., and Yang, Z. (2010). The impact of the representation of fossil calibrations on bayesian estimation of species divergence times. *Syst. Biol.* 59, 74–89. doi: 10.1093/sysbio/syp078

Kumar, S., and Hedges, S. B. (2016). Advances in time estimation methods for molecular data. *Mol. Biol. Evol.* 33, 863–869. doi: 10.1093/molbev/msw026

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096

Magnabosco, C., Moore, K. R., Wolfe, J. M., and Fournier, G. P. (2018). Dating phototrophic microbial lineages with reticulate gene histories. *Geobiology* 16, 179–189. doi: 10.1111/gbi.12273

Marjanović, D., and Laurin, M. (2007). Fossils, molecules, divergence times, and the origin of Lissamphibians. *Syst. Biol.* 56, 369–388. doi: 10.1080/10635150701397635

Marjanović, D., and Laurin, M. (2008). Assessing confidence intervals for stratigraphic ranges of higher taxa: the case of Lissamphibia. *Acta Palaeont. Pol.* 53, 413–432. doi: 10.4202/app.2008.0305

Marshall, C. R. (2008). A simple method for bracketing absolute divergence times on molecular phylogenies using multiple fossil calibration points. *Am. Nat.* 171, 726–742. doi: 10.1086/587523

Marshall, C. R. (2019). Using the fossil record to evaluate timetree timescales. *Front. Genet.* 10:1049. doi: 10.3389/fgene.2019.01049

Parham, J. F., Donoghue, P. C. J., Bell, C. J., Calway, T. D., Head, J. J., Holroyd, P. A., et al. (2012). Best practices for justifying fossil calibrations. *Syst. Biol.* 61, 346–359. doi: 10.1093/sysbio/syr107

Perie, M. D., and Doyle, J. A. (2012). Dating clades with fossils and molecules: the case of Annonaceae. *Bot. J. Linn. Soc.* 169, 84–116. doi: 10.1111/j.1095-8339.2012.01234.x

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., and Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* 67, 901–904. doi: 10.1093/sysbio/syy032

Rambaut, A., and Grassly, N. C. (1997). Seq-Gen: an application for the Monte Carlo simulation of protein sequence evolution along phylogenetic trees. *Bioinformatics* 13, 559–560. doi: 10.1093/bioinformatics/13.5.559

Reisz, R. R., and Müller, J. (2004). Molecular timescales and the fossil record: a paleontological perspective. *Trends Genet.* 20, 237–241. doi: 10.1016/j.tig.2004.03.007

Rosenberg, M. S., and Kumar, S. (2003). Heterogeneity of nucleotide frequencies among evolutionary lineages and phylogenetic inference. *Mol. Biol. Evol.* 20, 610–621. doi: 10.1093/molbev/msg067

Sauquet, H., Ho, S. Y. W., Gandolfo, M. A., Jordan, G. J., Wilf, P., Cantrill, D. J., et al. (2012). Testing the impact of calibration on molecular divergence times using a fossil-rich group: the case of *Nothofagus* (Fagales). *Syst. Biol.* 61, 289–313. doi: 10.1093/sysbio/syr116

Schenk, J. J. (2016). Consequences of secondary calibrations on divergence time estimates. *PLoS One* 11:e0148228. doi: 10.1371/journal.pone.0148228

Sterli, J., Pol, D., and Laurin, M. (2013). Incorporating phylogenetic uncertainty on phylogeny-based palaeontological dating and the timing of turtle diversification. *Cladistics* 29, 233–246. doi: 10.1111/j.1096-0031.2012.00425.x

Tamura, K., Battistuzzi, F. U., Billing-Ross, P., Murillo, O., Filipski, A., and Kumar, S. (2012). Estimating divergence times in large molecular phylogenies. *Proc. Natl. Acad. Sci. U.S.A.* 109, 19333–19338. doi: 10.1073/pnas.1213199109

Tamura, K., Tao, Q., and Kumar, S. (2018). Theoretical foundation of the reltime method for estimating divergence times from variable evolutionary rates. *Mol. Biol. Evol.* 35, 1770–1782. doi: 10.1093/molbev/msy044

Thorne, J. L., and Kishino, H. (2002). Divergence time and evolutionary rate estimation with multilocus data. *Syst. Biol.* 51, 689–702. doi: 10.1080/10635150290102456

Ware, J. L., and Barden, P. (2016). Incorporating fossils into hypotheses of insect phylogeny. *Curr. Opin. Insect Sci.* 18, 69–76. doi: 10.1016/j.cois.2016.10.003

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. *Proc. R. Soc. B Biol. Sci.* 282:20141013. doi: 10.1098/rspb.2014.1013

Warnock, R. C. M., Yang, Z., and Donoghue, P. C. J. (2012). Exploring uncertainty in the calibration of the molecular clock. *Biol. Lett.* 8, 156–159. doi: 10.1098/rsbl.2011.0710

Warnock, R. C. M., Yang, Z., and Donoghue, P. C. J. (2017). Testing the molecular clock using mechanistic models of fossil preservation and molecular evolution. *Proc. R. Soc. B Biol. Sci.* 284, 20170227. doi: 10.1098/rspb.2017.0227

Wolfe, J. M., and Fournier, G. P. (2018). Horizontal gene transfer constrains the timing of methanogen evolution. *Nat. Ecol. Evol.* 2, 897–903. doi: 10.1038/s41559-018-0513-7

Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24, 1586–1591. doi: 10.1093/molbev/msm088

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Powell, Waskin and Battistuzzi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genetics

HYPOTHESIS AND THEORY
published: 27 May 2020
doi: 10.3389/fgene.2020.00526

Rates and Rocks: Strengths and Weaknesses of Molecular Dating Methods

*Stéphane Guindon**

Laboratoire d'Informatique de Robotique et de Microélectronique de Montpellier, CNRS and Université Montpellier (UMR 5506), Montpellier, France

I present here an in-depth, although non-exhaustive, review of two topics in molecular dating. Clock models, which describe the evolution of the rate of evolution, are considered first. Some of the shortcomings of popular approaches—uncorrelated clock models in particular—are presented and discussed. Autocorrelated models are shown to be more reasonable from a biological perspective. Some of the most recent autocorrelated models also rely on a coherent treatment of instantaneous and average substitution rates while previous models are based on implicit approximations. Second, I provide a brief overview of the processes involved in collecting and preparing fossil data. I then review the main techniques that use this data for calibrating the molecular clock. I argue that, in its current form, the fossilized birth-death process relies on assumptions about the mechanisms underlying fossilization and the data collection process that may negatively impact the date estimates. Node-dating approaches make better use of the data available, even though they rest on paleontologists' intervention to prepare raw fossil data. Altogether, this study provides indications that may help practitioners in selecting appropriate methods for molecular dating. It will also hopefully participate in defining the contour of future methodological developments in the field.

Keywords: fossils, calibration, Bayesian inference, relaxed clock models, fossilized-birth-death process, total-evidence, tip-dating

OPEN ACCESS

Edited by:
Miguel Arenas, University of Vigo, Spain

Reviewed by:
Andrew M. Ritchie, Australian National University, Australia
David Marjanović, Museum für Naturkunde Leibniz Institut für Evolutions und Biodiversitätsforschung, Germany
Sebastian Duchene, The University of Melbourne, Australia
Gregory Fournier, Massachusetts Institute of Technology, United States

***Correspondence:**
Stéphane Guindon
guindon@lirmm.fr

Specialty section:
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Genetics

Received: *18 September 2019*
Accepted: *30 April 2020*
Published: *27 May 2020*

Citation:
Guindon S (2020) Rates and Rocks: Strengths and Weaknesses of Molecular Dating Methods. Front. Genet. 11:526. doi: 10.3389/fgene.2020.00526

1. INTRODUCTION

Telling apart the rate of molecular substitution from the time, measured in calendar units, that define periods of evolution, is the main endeavor of molecular dating techniques. The basic idea underlying these techniques is straightforward. The comparison of a set of homologous genetic sequences provides information about the number of (nucleotide, amino acid, or codon) substitutions that took place along the edges of the phylogeny connecting these sequences. If information is available about either the rate at which these substitutions take place or the actual timing of particular events in the phylogeny, then one may express the length of an edge as the product of the rate of molecular evolution by a calendar time elapsed along this edge. The application of this approach, in its simplest form (Zuckerkandl and Pauling, 1965), has led to spectacular findings—the reappraisal of the timing of divergence between African apes and humans (Sarich and Wilson, 1967) being perhaps the most emblematic. Since their first use more than five decades ago, molecular dating methods have considerably increased in sophistication. Heightened complexity indeed arose at many different levels of the analysis, going from the collection of genetic and fossil data to the reconstruction of phylogenetic trees.

This review article leaves aside many important aspects of modern techniques in molecular dating. In particular, it does not touch on the preparation of data, may it be the various algorithms for aligning homologous genetic sequences or the techniques used in the exploration of geological strata for extracting fossil data. Despite being central in Bayesian methods for the inference of node ages (see e.g., Condamine et al., 2015 for an illustration), the details of the tree-generating processes will also be largely omitted. Furthermore, computational considerations will not be discussed and I will not provide a list of available software implementing the most up-to-date techniques for molecular dating. Simulation techniques used to assess the accuracy and precision with which node ages are inferred, including the generation of phylogenies (Stadler, 2019) and genetic sequences (Fletcher and Yang, 2009; Currat et al., 2019), will also be ignored. I refer the keen reader to dos Reis et al. (2016) and Bromham et al. (2018) that give a broader overview of the various methodological and practical aspects pertaining to molecular dating.

The present work focuses instead on two specific aspects of molecular dating. It first provides an in-depth presentation of the models describing the variation of the rate of molecular evolution along a phylogenetic tree. This presentation serves as a basis to assess clock models, revealing some of the weaknesses of the most popular approaches. Note that the probabilistic models presented here focus exclusively on the evolution of the rate of substitution between nucleotides, amino-acids or codons. Yet, variations in the rate of evolution manifest themselves at other levels in molecules. For instance, the secondary structure of some proteins has been shown to evolve in a non-clock-like manner (Pascual-García et al., 2019). The mode and tempo of evolution of secondary and tertiary protein structures is beyond the scope of this study, however, and I will only deal with the evolution of primary sequences. The second part of this review deals with the techniques used for calibrating molecular dating analyses based on fossil data. Here again, statistical and biological arguments are presented that help evaluate the relevance of the main techniques, including the most recent developments such as the fossilized-birth-death model and the "total-evidence" approach.

2. RATES OF MOLECULAR EVOLUTION ALONG PHYLOGENIES

A substitution between two nucleotides at a particular position in a genome is the outcome of two distinct events. For this reason, it is useful to distinguish between a *mutation*, which is the outcome of a biochemical process, and a *substitution*, which involves a series of population-level events leading to the fixation of a mutation, as detailed below. The mutation that substitutes a nucleotide $i \in \mathcal{A} := \{\text{A,T,G,C}\}$ by another nucleotide $j \neq i$ may be modeled as a "uniformly at random" event, i.e., given i, all $j \neq i$ have the same probability of replacing i. It is clear however that the biological reality is more subtle than that simple model. For instance, mutations generally favor transitions over transversions. Although part of this bias is the consequence of natural selection acting on proteins, it has been shown that transitions are over-represented in pseudogenes (Gojobori et al., 1982), suggesting that the biochemical processes involved in mutations are also responsible for the observed bias. Other biochemical events, such as biased gene conversion (Duret and Galtier, 2009) for instance, invalidate to a certain extant the "uniformly at random" assumption, at least in some parts of the genomes (i.e., the regions prone to recombination) in mammals and yeast. Furthermore, mutation rates are most likely influenced by species-specific characteristics such as generation time, metabolic rate and DNA repair efficiency (Gillespie, 1994; Baer et al., 2007), such that these rates are also likely to vary extensively across lineages in the tree of life.

The second event involved in the making of a substitution is fixation. Although a mutation arises in a single genome at a given point in time, its frequency in the population, and in the species this population belongs to, may increase until it completely replaces the original allele. Note that we assume here that mutations are rare such that a mutant allele reaches fixation or vanishes from the population before a new mutation arises. The process of fixation of a mutation is complex as it is governed by various evolutionary forces such as natural selection (beneficial mutations will, on average, reach fixation more frequently and quickly than mutations leading to a decrease in fitness), genetic drift (the fixation of an allele occurs more quickly in small vs. large populations), and migrations. These three forces constitute the main focus of studies in classical population genetics and will not be discussed in more detail here. Most phylogenetic analysis techniques rely on a phenomenological approach for modeling substitutions whereby mutation and fixation are not distinguished explicitly. Note however that a substantial body of work has focused on deriving models of substitution from the basic principles of population genetics (Halpern and Bruno, 1998; Nielsen and Yang, 2003; Thorne et al., 2007; Cartwright et al., 2011).

The accumulation of substitutions between nucleotides during the course of evolution is thus generally assumed to be governed by a continuous-time Markov process. Individual sites are here considered as independent and identically distributed (iid), i.e., a simulation of the same Markov process runs along the phylogeny, at each position along the genome, to give rise to the observed data at the tips of the tree. The iid assumption is of course problematic when dealing with coding sequences. Indeed, through the action of natural selection, a non-synonymous change in a given region of the sequence may cause another non-synonymous substitution in a remote region in order to compensate for the first one. Yet, substitutions taking place at third coding positions may be considered as approximately iid and the same approximation can be made for non-coding regions of the genome or for pseudo-genes.

2.1. The Strict Clock Model and an Extension

At a given point in time t during the course of evolution, in a particular ancestral lineage labeled with index l and at a particular site s, one considers that substitutions accumulate randomly, following a Poisson point process of intensity $\mu(l,t,s)$. The substitution rate is generally decomposed as follows: $\mu(l,t,s) = r(s)\mu(l,t)$, where $r(s)$ describes the variability of rates across sites. This random variable generally follows a discrete gamma distribution (Yang, 1994) although the use of non-parametric

distributions is now commonplace (Soubrier et al., 2012). In the following, I will ignore this extra layer in the model describing the rate at which substitutions accumulate. I will thus focus on the term $\mu(l,t)$ here on.

A first approach for modeling the fluctuation of the rate of substitution when considering the evolution of multiple species is to assume that $\mu(l,t)$ is constant throughout, i.e., $\mu(l,t) = \mu$. This simplification corresponds to the well-known "strict molecular clock" model pioneered by Zuckerkandl and Pauling (1965). Note that the actual (or realized) number of substitutions in a given time interval, along a particular lineage, may vary from one site to another because of the inherent stochasticity of the underlying process. Yet, these numbers of substitutions are all considered as random draws from the same Poisson distribution.

Molecular sequences can be considered as snapshots of molecular evolution at a single or a few point(s) in time. Hence, detailed information about evolutionary events at *all* points in time is forever lost and only *average* trends can be recovered from the observation. In the following, I will focus on the relationship between clock models, such as the strict clock model cited above, and average substitution rates in the context of date inference. More specifically, the (pathwise) average substitution rate, λ_t, is defined as follows:

$$\lambda_t := \frac{1}{t}\int_0^t \mu(l,x)\mathrm{d}x. \tag{1}$$

Hence, λ_t is proportional to the integral over the rate trajectory $\{\mu(l,x), 0 \leq x \leq t\}$ that gives the value of the substitution rate at all points in time in the time interval $[0,t]$. As we will see below, some clock models (uncorrelated ones in particular) focus solely on the distribution of the average rate λ_t. Other approaches model instead the instantaneous rate $\mu(l,x)$, even though only the average can be inferred from the analysis of molecular data.

Bayesian inference of divergence times rests on the joint posterior density of model parameters given the observed data. I consider for now that data simply consists in two contemporaneous sequences, corresponding to random variables U and V, displaying sequences u and v, respectively. Beside molecular data, one also observes fossil data, noted as I and defining time constraints, i, on the age of the most recent common ancestor of the species with sequences U and V. When using a time-reversible, homogeneous and stationary Markov process describing substitutions between genetic character states, with stationary probabilities noted as $\pi_\cdot$, the joint posterior density of interest is then expressed as follows:

$$\begin{aligned} p_{\Lambda_t,M_t,M_0,T}(\lambda_t,\mu_t,\mu_0,t|U=u,V=v,I=i) \\ &\propto \Pr(U=u|V=v,\Lambda_t=\lambda_t,M_t=\mu_t,M_0=\mu_0,T=t) \\ &\times p_{\Lambda_t}(\lambda_t|M_t=\mu_t,M_0=\mu_0,T=t) \\ &\times p_{M_t}(\mu_t|M_0=\mu_0,T=t) \\ &\times p_{M_0}(\mu_0) \\ &\times p_T(t|I=i) \\ &\times \pi_v \end{aligned} \tag{2}$$

The first term to the right of the "proportional to" ($\propto$) sign is the probability of transitioning from state v (random variable: V) to state u (R.V.: U) along an evolutionary path that lasted t calendar units of times (R.V.: T), with instantaneous rates at the start and at the end of that path being equal to μ_0 and μ_t respectively (R.V.: M_0 resp. M_t), and average rate (as defined in Equation 1) equal to λ_t (R.V.: Λ_t).

In the expression above, the transition probability between observed characters (nucleotide or amino-acids generally) is a function of the instantaneous substitution rates at times 0 and t, plus the average rate in the corresponding time interval. Yet, knowing the instantaneous rates is not required. Indeed, one has:

$$\begin{aligned} &\Pr(U=u|V=v,\Lambda_t=\lambda_t,M_t=\mu_t,M_0=\mu_0,T=t) \\ &\quad=\sum_{k=0}^{\infty}\Pr(U=u,N_t=k|V=v,\Lambda_t=\lambda_t,M_t=\mu_t,M_0=\mu_0,T=t) \\ &\quad=\sum_{k=0}^{\infty}\Pr(U=u|V=v,N_t=k,T=t)\times\Pr(N_t=k|\Lambda_t=\lambda_t), \end{aligned}$$

where N_t is the random variable giving the number of substitutions that took place in $[0,t]$. The key observation here is that the distribution of N_t is determined by a non-homogeneous Poisson process (the parameter of this Poisson process is defined by the rate trajectory). This distribution is unaffected by the specifics of the rate trajectory. It is a function of the average rate along the trajectory only (i.e., λ_t), thereby explaining why μ_t and μ_0 vanish in the equation above. The transition probability thus simplifies to give the following expression (in a simplified notation):

$$\Pr(u|v,\lambda_t,\mu_t,\mu_0,t) = \Pr(u|v,\lambda_t,t) \tag{3}$$

$$= [e^{-\lambda_t t\mathbf{Q}}]_{v,u}, \tag{4}$$

where $\mathbf{Q}$ is the generator of the Markov chain governing substitutions ($[\mathbf{Q}]_{v,u\neq v}$ gives the rate of change from state u to v).

One may then envisage various instances of the clock model. The strict clock model corresponds to the case where the rate trajectories are deterministic such that $p_{\Lambda_t}(\lambda_t|\mu_t,\mu_0)\mathrm{d}\lambda_t = p_{M_t}(\mu_t|\mu_0)\mathrm{d}\mu_t = p_{M_0}(\mu_0)\mathrm{d}\mu_0 = 1$ when $\lambda_t = \mu_t = \mu_0$ and 0 otherwise, i.e., instantaneous rates are all equal at all points in time along the considered edge (and thus equal to the average rate too). The joint probability density of the model parameters then becomes as follows:

$$\begin{aligned} p_{\Lambda_t,M_t,M_0,T}(z,z,z,t|U=u,V=v,I=i) \propto &[e^{-zt\mathbf{Q}}]_{v,u} \\ &\times p_T(t|I=i) \\ &\times \pi_v. \end{aligned} \tag{5}$$

When breaking the evolutionary path between times 0 and t into two time intervals, $[0,s]$ and $[s,t]$, the product, denoted as $\mathcal{A}$, of the two terms in Equation (2) describing the evolution of the rate of evolution, i.e., $\mathcal{A} := p_{\Lambda_t}(\lambda_t|\mu_t,\mu_0,t)\times p_{M_t}(\mu_t|\mu_0,t)$ is then defined as follows:

$$\begin{aligned} \mathcal{A} :=& p_{\Lambda_s,\Lambda_{t-s}}(\lambda_t,\lambda_{t-s}|M_t=\mu_t,M_s=\mu_s,M_0=\mu_0,S=s,T=t) \\ &\times p_{M_s}(\mu_s|M_0=\mu_0,S=s) \\ &\times p_{M_t}(\mu_t|M_s=\mu_s,S=s,T=t), \end{aligned} \tag{6}$$

and takes on the following value under the "standard" strict molecular clock model:

$$\mathcal{A} = \begin{cases} g_M(\mu) & \text{if } \lambda_t = \lambda_{t-s} = \mu_t = \mu_s = \mu_0 = \mu, \\ 0 & \text{otherwise,} \end{cases} \quad (7)$$

where M denotes the random variable giving the instantaneous rate of evolution everywhere in the tree under the standard strict clock model. In other words, the strict clock model assumes that average and instantaneous rates are the same everywhere in the tree, not just along individual edges as seen above.

2.1.1. The "Not-so-Strict" Clock Model

Thanks to the distinction between instantaneous and average substitution rates that is made explicit here, one may design a new clock model where instantaneous substitution rates fluctuate randomly during the course of evolution but the average rate stays the same along every edge in the tree. Under the so-called "not-so-strict" clock model, the first term in Equation (6) giving the joint density of average rates along the two successive time segments, may be defined as follows:

$$p_{\Lambda_s,\Lambda_{t-s}|\Lambda_s=\Lambda_{t-s}}(\lambda_s,\lambda_{t-s}|\mu_t,\mu_s,\mu_0,s,t) := \begin{cases} 0 & \text{if } \lambda_s \neq \lambda_{t-s}, \\ \frac{p_{\Lambda_s,\Lambda_{t-s}}(\lambda_s,\lambda_{t-s}|\mu_t,\mu_s,\mu_0,s,t)}{\int p_{\Lambda_s,\Lambda_{t-s}}(\lambda,\lambda|\mu_t,\mu_s,\mu_0,s,t)\mathrm{d}\lambda} & \text{otherwise.} \end{cases} \quad (8)$$

Genetic sequences combined with fossil data convey information about average, not instantaneous, rates. It is thus hopeless to try and fit the "not-so-strict" clock model to standard data sets used in molecular dating without any extra information. This model could nonetheless be relevant in particular circumstances. When considering intra-species data for instance, prior information about past variation of population sizes is sometimes available. These variations may serve as a basis to inform the part of the "not-so-strict" clock model describing the evolution of the instantaneous rates along the tree, even though the molecular clock hypothesis generally holds at that scale. More importantly, the "not-so-strict" model can be envisaged as an intermediate between relaxed and strict clock models. In case one has an informative prior about the frequency and the amplitude of the fluctuation of instantaneous rates of substitution, various molecular clock hypotheses could then be tested by comparing the strict to the "not-so-strict" models in the first place, and, depending on the outcome of that test, comparing the fit of the "not-so-strict" model to that of relaxed clock models (see next sections).

2.2. Uncorrelated Relaxed Clock Models

The strict and not-so-strict clock models can be expanded in two distinct ways. A first approach is to design a model whereby $\mu(l,t) = \mu(l)$ for all t along edge l, i.e., each lineage has its own rate of substitution, which may differ from that of other lineages, but the rate along a given lineage l is constant. The rate trajectory along a branch of the phylogeny is deterministic under this model, i.e., instantaneous rates do not fluctuate randomly along each lineage. These rates can change at internal nodes in the tree though, thereby authorizing deviations from the strict molecular clock hypothesis. This model has serious conceptual issues, however, since it is not *sampling-consistent*. Sampling-consistency is a concept different from the more standard concept of statistical consistency. It was used in Barton et al. (2010) to refer to the situation where the probabilistic distribution of the age of a particular node under a given tree-generating process depends on the total number of tips in the tree. In the present context, clock models lack sampling-consistency when the number of sampled individuals (or the number of internal nodes in the tree) influences the number of potential shifts in the substitution rate. Sampling-consistency is desirable since there is no sensible reason, based on biological evidence, to believe that the rate trajectory along a particular lineage should be influenced by the number of sampled cladogenesis events taking place along it (although it may be influenced by the *total*, i.e., sampled and unsampled, lineage splits).

A second approach to define a relaxed clock model is to enforce the following constraint: $\int_0^t \mu(l,x)\mathrm{d}x = \mu(l)t$ for all rate trajectories $\{\mu(l,x), 0 \leq x \leq t\}$. In this second model, rates can fluctuate randomly along edge l although all the rate trajectories have to average to $\mu(l)$. This model is thus the "not-so-strict" equivalent to the models introduced in the previous section, in a context where rates may change across lineages. That second interpretation of the relaxed clock model is certainly more satisfactory than the previous one from a biological point of view as changes of the rate of substitution can take place at any point in time during the course of evolution. Yet, random fluctuations of the instantaneous rates naturally imply random fluctuations of the average rates too. Hence, except in the special case where the periods of time considered are very long compared to the time scale at which the instantaneous rate varies noticeably (in which case the average rates taken over multiple trajectories should all converge to a fixed value) there is no strong reason to believe that the average rate along a given edge is not a randomly varying quantity. As will be seen below, alternative models exist that are more realistic than the relaxed clock ones, without compromising on the computational burden involved.

A first relaxed clock model assumes that the average rate of substitution along a branch is exponentially distributed (Drummond et al., 2006). This model is a popular choice as it is implemented in the BEAST 1 (Drummond and Rambaut, 2007) and BEAST 2 (Bouckaert et al., 2014) packages. According to it, $\mu(l)$ is the realization of a random variable exponentially distributed with parameter $1/\mu$. The exponential model is therefore a clock model as all lineages are governed by the same parametric distribution. It is *relaxed* since the average rate along each branch is taken as a new random draw from this distribution.

The exponential relaxed clock model considers that the average rate of substitution along a given branch has a mean equal to μ and a variance equal to μ^2. One thus relies here on a model in which the larger the average substitution rate, the larger its variance, which is consistent with the idea that large quantities vary more than smaller ones. Perhaps surprisingly, a more detailed analysis of this model reveals that the extent

of deviation from the strict clock constraint does *not* depend on how fast (or slow) substitutions accumulate. For instance, in the case where $\mu = 1$ substitution per unit of time, the probability for a lineage to evolve twice as fast as μ (or faster) is 0.14. When $\mu = 0.1$, this probability is also equal to 0.14. This property of the exponential distribution is reflected in the excess kurtosis, which measures how likely it is to observe extreme values in a probabilistic distribution. For the exponential family, the kurtosis is not a function of the mean or the variance (it is in fact simply equal to six). Hence, it is counter-intuitive that large deviations from the strict clock have the same probability to occur when molecules evolve slow or fast, even though the variance of the average rate is larger when sequences evolve quickly. More importantly, molecules that evolve slowly may be doing so as a consequence of stabilizing selection. In this case, natural selection may prohibit large deviations from the strict clock, and doubling the rate at which substitutions accumulate may be very unlikely indeed. In a symmetric manner, neutrally evolving sequences may have more latitude to double (or halve) their rates of evolution in the real world. In summary, the exponential distribution model may provide a reasonably good approximation of the true underlying distribution only for fast- *or* slow-evolving sequences, but not for both of them.

Beside the exponential family, the fluctuations of average rates of evolution across edges in the tree are often described by a lognormal or a gamma distribution (Lepage et al., 2007). Both models are available in BEAST 1 and 2 as well as in MrBayes 3.2 (Ronquist et al., 2012b). While the exponential distribution is fully specified with just one parameter, the lognormal and gamma distributions rely on two parameters instead. These two parameters set the mean and the variance of the corresponding distributions in a separate manner, i.e., without any "hard-coded" constraints as for the exponential family that impose a quadratic relationship between the mean and the variance. Moreover, the statistical properties of the lognormal and the gamma distributions are such that slow-evolving sequences are less likely than fast-evolving ones to deviate strongly from the strict clock constraint. The lognormal and gamma families thus appear superior to the exponential distribution from a biological point of view, although the increased realism comes at the cost of estimating an extra parameter.

Beside considerations regarding the properties of various relaxed clock models when focusing on a single edge of the tree, important observations are to be made when expanding our focus to whole evolutionary paths between the root and the tips of the phylogeny. Under the uncorrelated clock models, the average substitution rate along edges in the phylogeny are all iid random variables. The sum of edge lengths on a path between the root and a tip of the tree is thus a weighted sum of iid random variables, the weights corresponding to the times elapsed along every edge on that path. Assuming that these weights are all equal to one, the sum of k average rates along a path is given by $Z_k := \sum_{i=1}^{k} \mu(i, 1)$. Invoking the central limit theorem, the random variable Z_k is thus asymptotically normally distributed (i.e., when $k \to \infty$) with mean μk and variance $\sigma^2 k$, where μ and σ^2 are the expectation and the variance of $\mu(i, 1)$ respectively. Hence, in layman's terms, tip to root distances fluctuate more in large trees (large in terms of the number of tips, k) compared to small ones. This behavior is difficult to justify from a biological perspective as one would expect the patterns of rate variation along a single lineage to be unaffected by the total number of lineages included in the sample. All uncorrelated clock models have the same behavior in that regard. Moreover, this unwelcome relationship between variance and number of tips is likely to impact phylogenies in a differential manner depending on their shape, with comb-like phylogenies showing highly variable numbers of edges between the root and tip nodes, while more balanced tree shapes show less extreme variation in the length of these paths. In that regard, selecting subsets of taxa so as to make the phylogeny more balanced probably helps circumventing this issue. Yet, this practice may lead to increasing sampling errors due to the decreasing amounts of data available for the dating analysis.

I have assumed that the time elapsed along each branch was equal to one in the previous paragraph. I also focused on the distribution of a single tip-to-root distance. In practice, however, amounts of time elapsed along branches of actual phylogenies vary between edges. Moreover, one is usually interested in the variance of tip-to-root distances *within* a tree, not *across* trees. This "within-tree" variance is harder to characterize analytically due to the correlation between tip-to-root paths as defined by the tree. I thus performed simulations where trees were first generated according to a birth-death process with sampling. The `TreeSim` package (Stadler, 2019) available in the R programming language was used here in order to simulate trees conditioned on a given number of sampled tips. The birth and death rates were set to 1.0 and 0.5, respectively and 20,000 trees were generated. Each tree was obtained by simulation of a birth-death process with 100 species. The fraction of sampled taxa was chosen uniformly at random in $[0.1, 1]$. The length of each edge in these birth-death trees was then multiplied by a random draw taken from an exponential distribution with the rate set to 1.0. Tip-to-root distances were extracted from each tree using the `phytools` package (Revell, 2018) and their variance calculated.

Figure 1 gives the plot of the number of tips (on the x-axis) against the variance of the tip-to-root distances in each tree (y-axis). This plot confirms the positive correlation between the number of tips in the tree and the variance of the tip-to-root distances. Similar results were obtained with the gamma model (see **Figure S1**) while the Kishino et al. (2001) autocorrelated clock model does not display obvious signs of positive correlation (**Figure S2**). In the context of the Bayesian inference of divergence dates, uncorrelated clock models therefore define prior distributions on rates of evolution that entail stronger deviations from the strict clock in larger trees. This behavior may impact the inference of node ages. Indeed, an increased variability in rates along lineages in data sets with large numbers of tips may be responsible for an inflation of the variance in the age estimates themselves. Note also that this inflation may preferentially impact the ages of nodes along lineages that are composed of a large number of branches compared to those consisting of only a few branches. Although simulation studies generally focus on the accuracy with which ages are estimated, considering the impact of the various

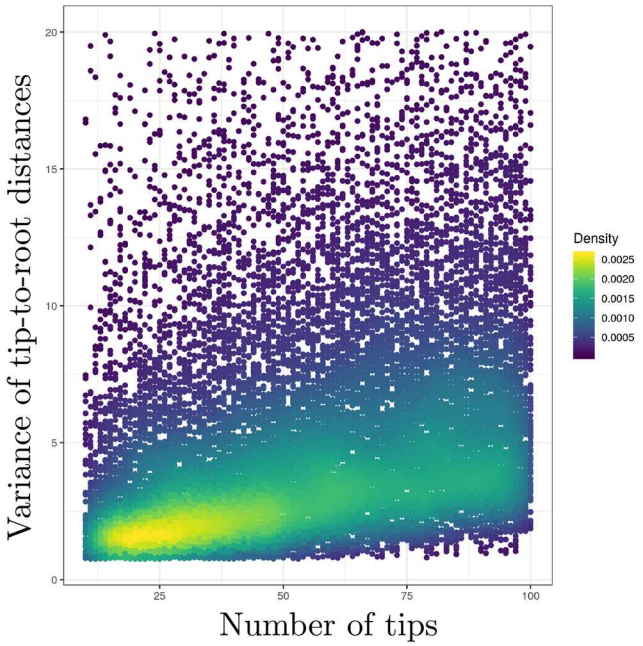


FIGURE 1 | Uncorrelated clock models produce stronger deviations from the strict clock constraint in larger trees. *y*-axis: variance of the tip-to-root distances. *x*-axis: number of tips. Deviation from the strict molecular clock hypothesis was generated using the uncorrelated exponential model. The scatterplot derives from the analysis of 20,000 observations, each observation being a value for the tip-to-root variance and the corresponding number of tips.

models of rate variation on the precision of these estimates should therefore be examined too.

Furthermore, Lepage et al. (2007) mention that defining a model where rate *trajectories* along each edge define gamma- or exponentially distributed *average* rates is not trivial (the same argument applies to the uncorrelated log-normal model). They state that these "trajectory models" would display rate autocorrelation within each edge, even though the trajectories across distinct edges would be truly independent. Hence, here again, the uncorrelated clock models appear to be sampling-inconsistent: the amount of (instantaneous) rate autocorrelation depends on the number of internal nodes (and thus the number of tips) in the tree. Heath et al. (2012) describe a more sophisticated uncorrelated clock model whereby the average substitution rate along each branch derives from a Dirichlet process prior (DPP). This model assumes a finite number of rate values. Each of these rates is considered as a random draw from a gamma distribution whose parameters are fixed *a priori*. The number of rate classes is estimated from the data through a so-called "concentration" parameter. The DPP model is thus conceptually very close to a slightly modified version of the standard uncorrelated clock model where rates follow a discretized gamma distribution. Further investigations would be required though in order to assess whether the DPP model suffers from the same shortcomings as that discussed above. It seems likely however that all uncorrelated models proposed so far, including DPP, lack sampling-consistency and implicitly rely on the questionable assumption that large trees (in terms of their number of tips) deviate more from the strict clock hypothesis than smaller ones.

2.3. Autocorrelated Clock Models

Equation (6) gives a generic expression characterizing the distribution of the instantaneous rates at three successive points in time along with that of the corresponding average rates. It is relatively straightforward to introduce correlation between substitution rates along the tree in this framework. The first model explicitly accommodating for autocorrelation was proposed by Thorne et al. (1998). Instead of modeling instantaneous rates at times s and t, corresponding to the end nodes of the first and second edge respectively, they focused on the rates at times $s' := s/2$ and $t' := (t+s)/2$, i.e., the "middle" of the corresponding two branches. Equation 6 therefore needs a slight re-writing to yield:

$$\begin{aligned}\mathcal{A} :=& p_{\Lambda_s,\Lambda_{t-s}}(\lambda_s,\lambda_{t-s}|M_{t'}=\mu_{t'},M_{s'}=\mu_{s'},S=s,T=t)\\ &\times p_{M_{t'}}(\mu_{t'}|M_{s'}=\mu_{s'},S=s,T=t)\\ &\times p_{M_{s'}}(\mu_{s'}|S=s).\end{aligned} \tag{9}$$

The instantaneous rate transition probability density (i.e., $p_{M_{t'}}(\mu_{t'}|M_{s'}=\mu_{s'},S=s,T=t)$) is given by a lognormal distribution with $\mathrm{E}[M_{t'}^n] = e^{n\log(\mu_{s'})+n^2\sigma^2(t'-s')/2}$. The first two moments fully specify the whole distribution. One thus assumes here that the logarithm of the instantaneous rate at time t' is a normally distributed random variable with a mean equal to the logarithm of the instantaneous rate at time s' and variance proportional to $t'-s'$. The mean of the lognormal distribution is thus equal to $\mu_{s'} \times e^{\sigma^2(t'-s')/2}$, which is larger than the ancestral rate $\mu_{s'}$ in general, thereby leading to an unwarranted increase of the substitution rate over time. Kishino et al. (2001) fixed this issue by replacing the original normal distribution with one with a mean of $\log(\mu_{s'}) - \sigma^2(t'-s')/2$ so that the mean of the lognormal distribution is now equal to $\mu_{s'}$, the ancestral rate.

The autocorrelated lognormal model assumes that the logarithm of the instantaneous rate follows a Brownian process. The rate itself thus evolves according to a geometric Brownian process. This model captures the idea that instantaneous rates of evolution vary little over short periods of time while longer time intervals may bear stronger fluctuations. The amplitude of these fluctuations is governed by the parameter σ which may be estimated from the data. This model also introduces correlation of rates between sister lineages. Indeed, when considering sister edges ending with nodes X_t and X_u in **Figure 2**, the random variables $M_{t'}$ and $M_{u'}$ are not conditionally independent given $M_{s'}$ since both of them share the evolutionary path between s' and s. In practice, the non-independence between sister edges seems to have been disregarded so that only an approximation of the joint density of instantaneous rates was used instead. Kishino et al. (2001) acknowledged this problem, later proposing a different model that corresponds to that defined by Equation (6), i.e., M_s and M_t, the rates at the end of the corresponding edges, replace the mid-point rates M'_s and M'_t.

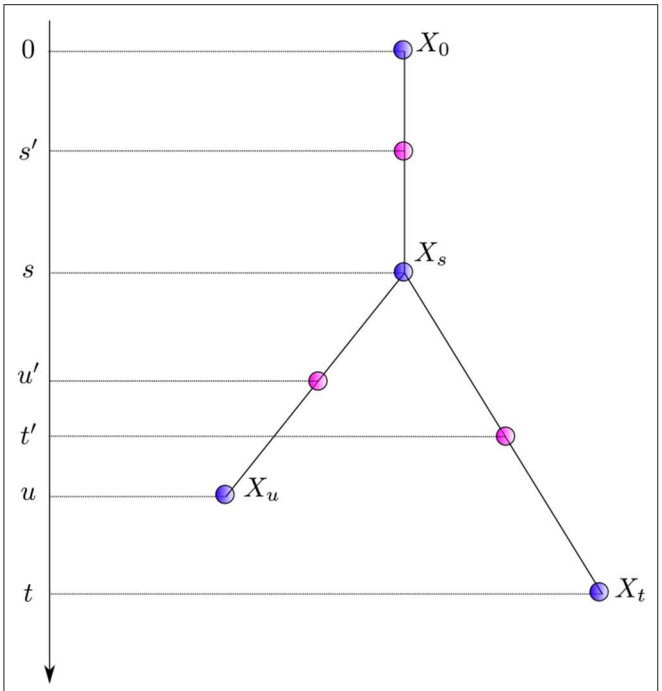


FIGURE 2 | Notation for nodes and times. X_i denotes a node with time t_i. Times of nodes (in purple) and that corresponding to the middle of each edge (in pink) are presented to the left.

Although the autocorrelated log-normal model provides what can be considered a reasonable description of the evolution of the instantaneous rates along a phylogeny, Lepage et al. (2006) point out that it does not agree with some of the tenets of evolutionary biology. Indeed, the theory of episodic evolution, where periods of adaptation are followed by evolutionary stasis, implies that high evolutionary rates are more likely to decrease than the contrary. The geometric Brownian assumes instead that, at a particular point in time, the substitution rate is as likely to double as it is to halve. Lepage et al. (2006) also convincingly argue that the distribution of the rate at time $t = +\infty$ should be unique, i.e., it should not depend on what the rate is at time $s < t$, which is not the case for the geometric Brownian process.

The Ornstein-Uhlenbeck (OU) model is a diffusion process that, unlike the geometric Brownian, satisfies this last property. However, it can take on negative values, which is not relevant when modeling rates of evolution. Aris-Brosou and Yang (2002) used the OU process in a Bayesian molecular dating approach nonetheless. It is not clear how the constraint of non-negativity of rates was implemented in this study, however. Lepage et al. (2006) proposed to use the Cox-Ingersoll-Ross process (Cox et al., 2005) instead. This process is a generalization of the squared OU model. It thus describes the random fluctuations of non-negative quantities. As for the OU model, the CIR process also has a unique limiting distribution. Two parameters specify the variance and the autocorrelation in rate trajectories in an independent manner. A third parameter defines the mean of the limiting distribution.

Beside the theoretical properties of the various clock models, practical aspects should also be considered carefully—the most important one being perhaps the relevance of the various models in the context of data analysis. Using simulations, Ho et al. (2015a) showed that detecting autocorrelation between rates is difficult so that uncorrelated and autocorrelated models often provide equally good fits to the data. Analyzing a large primate data set, dos Reis et al. (2018) observed however that the choice of rate model (autocorrelated vs. uncorrelated) has a substantial impact on the date estimates. An autocorrelated rate model provides here a significantly better fit than the uncorrelated model tested in their study. Even though autocorrelated rates do not always outperform uncorrelated ones, using autocorrelated rate models in cases rates are in fact not correlated should not, at least in principle, lead to poor date estimates. Hence, as long as the uncertainty around rate autocorrelation is taken into account in the inference, using autocorrelated clock models in practice seems preferable.

It is also essential to take into account what is known about the biology of substitution rate evolution when comparing the merits of various modeling approaches. Effective population size is one of the factors regulating the rate at which substitutions accumulate through its impact on the strength of genetic drift and selection. It is not clear however whether variations in population size during the course of evolution follow uncorrelated or correlated patterns (Lanfear et al., 2014). Life history traits such as body mass, body size and temperature, metabolic rate and generation time are also associated with the substitution rate (Levy Karin et al., 2017). Body size has been modeled as a diffusion process (see e.g., Clauset and Erwin, 2008), resulting in certain degree of autocorrelation. Yet, just because body size evolves in an autocorrelated fashion does not imply that substitution rates should follow the same patterns. It is thus safe to assume that population size and life history traits probably evolve in a seemingly uncorrelated manner when considering only distant species, so that uncorrelated models of substitution rate are appropriate at that scale. When the analysis focuses instead on closely related species and shorter evolutionary time scales, autocorrelated models are probably more relevant.

Furthermore, all models discussed here are clock models. They all assume that instantaneous substitution rates fluctuate around some average in an autocorrelated or uncorrelated manner. The fact that this average is shared by all lineages makes these models clock-like. It may be relevant to deviate from the clock assumption in particular circumstances though. Specifying multiple independent clock models may indeed be pertinent in cases where the biology of a subset of organisms is markedly distinct from that of the other taxa analyzed. For instance, comparative analyses including prokaryotes and organelles (Esser et al., 2004) may require two distinct clocks. In that respect, the "random local clock" model proposed by Drummond and Suchard (2010) addresses this particular need even though, strictly speaking, lineages still evolve under a clock model here.

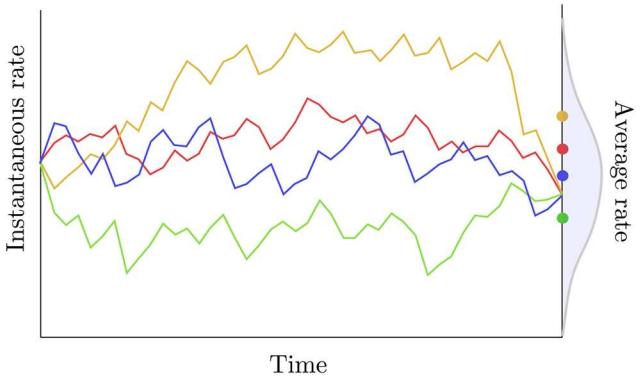


FIGURE 3 | Rate trajectories and average rates. The colored lines describe rate trajectories, i.e., values of the instantaneous rate of substitution at all points in time. The four trajectories have the same instantaneous rate at the beginning and at the end of the time period considered here. Traditional autocorrelated clock models would thus assign the same length to the corresponding edge in all four cases here. The dots to the right give the average rate for each trajectory taken over that same time period. These averages are iid random quantities whose distribution is given in gray.

2.3.1. Rate Trajectories vs. Averages

Going back to Equation (6), the question of defining the probability density of the mean rate along a given branch given the instantaneous rates at its extremities remains. Kishino et al. (2001) rely on a strong assumption about the corresponding distribution. Indeed, the corresponding probability density is defined as follows:

$$p_{\Lambda_s}(\lambda_s|M_s=\mu_s,M_0=\mu_0,S=s)\mathrm{d}\lambda_s := \begin{cases} 0 & \text{if } \lambda_s \neq (\mu_0+\mu_s)/2 \\ 1 & \text{otherwise,} \end{cases} \quad (10)$$

i.e., the distribution of the average rate has a point mass probability set at the average of the instantaneous rates at the two extremities of the edge considered. Two observations can be made regarding this definition. First, assuming that the rate trajectories are governed by a geometric Brownian process with a small variance parameter (σ) and/or considering a short period of time, then the trajectories are approximately linear and the average rate is indeed equal to the arithmetic mean of the instantaneous rates at the beginning and at the end of each trajectory. Kishino et al. (2001) also assume that the variance of the average rate is null. Here again, this assumption is only reasonable in the particular case where σ is small and/or a short time interval is considered. In general, however, given the instantaneous rates at both extremities of an edge and assuming geometric Brownian trajectories between these two values, the average rate along the branch should be treated as a random variable with potentially non-zero variance (**Figure 3**).

Lepage et al. (2006) were the first to clearly point out that assuming random trajectories for instantaneous rates implies that the average rate along each edge should also be considered random. Acknowledging the randomness of the average substitution rate poses the question of the derivation of transition probabilities between character states along edges of the phylogeny. Replacing the expression for the transition probability in Equation (2) by that given in Equation (4) yields:

$$\begin{aligned} p_{\Lambda_t,M_t,M_0,T}(\lambda_t,\mu_t,\mu_0,t|U=u,V=v,I=i) \propto\ & [e^{-\lambda_t t\mathbf{Q}}]_{v,u} \\ & \times p_{\Lambda_t}(\lambda_t|M_t=\mu_t,M_0=\mu_0,T=t) \\ & \times p_{M_t}(\mu_t|M_0=\mu_0,T=t) \\ & \times p_{M_0}(\mu_0) \\ & \times p_{T|I}(t|i) \\ & \times \pi_v. \end{aligned} \quad (11)$$

This last expression suggests that Bayesian inference of divergence dates should incorporate instantaneous rates at all nodes in the tree plus the corresponding average rate along each edge as latent variables, which could be effectively "integrated over" using standard Metropolis-Hasting operators for instance. Lepage et al. (2006), Guindon (2012), and Privault and Guindon (2015) went one step further by showing that it is in fact possible to drop the average rates from the (rather long) list of latent variables. The posterior distribution of interest then becomes:

$$\begin{aligned} p_{M_t,M_0,T}(\mu_t,\mu_0,t|U=u,V=v,I=i) \propto\ & \int [e^{-zt\mathbf{Q}}]_{v,u} \\ & \times p_{\Lambda_t}(\lambda_t|M_t=\mu_t,M_0=\mu_0,T=t)\mathrm{d}z \\ & \times p_{M_t}(\mu_t|M_0=\mu_0,T=t) \\ & \times p_{M_0}(\mu_0) \\ & \times p_{T|I}(t|i) \\ & \times \pi_v. \end{aligned} \quad (12)$$

The transition probability is thus derived here by integrating over all possible values that the average rate can take conditioned on the instantaneous rates at the branch extremities. The corresponding integral (i.e., the first term to the right of the equation above) can be solved analytically, or approximated, in some circumstances. I refer to this calculation as the "integrated average-rate approach," or IARA, in the following. Lepage et al. (2006) provide a closed-form formula for a IARA assuming that the instantaneous rates evolve under the CIR process. They used a simplified version of it in order to evaluate the likelihood on a three-taxon star-like tree. In a subsequent study, Lepage et al. (2007) used the mean of the distribution of the average rate under the CIR but assumed a null variance. In Guindon (2012), I focused instead on the geometric Brownian process, providing an approximation for the distribution of the average rate. The calculation of the transition probabilities under this IARA entails the same computational cost as that spent when considering that the average rate is not random. Privault and Guindon (2015) later examined this approximation further, confirming its validity for realistic ranges of parameters. They also provide a closed-form formula for the transition probabilities, although numerical precision issues may hamper the calculation in particular circumstances.

Another model describing the evolution of the rate of evolution is the compound Poisson process proposed by Huelsenbeck et al. (2000). This model properly accommodates

the randomness of the average rate along edges in the phylogeny. However, it relies on augmenting the data by assuming that the instantaneous rate is known at every point in time, along all lineages in the phylogeny, making this approach computationally expensive. The CIR and geometric Brownian IARAs thereby appear as the best options available so far to fully accommodate for the random fluctuations of the average rate of substitutions along edges due to the underlying stochastic process governing the instantaneous rate of evolution. Beyond their ability to describe rate trajectories in a sound mathematical framework, considering the variation of the average rate along each edge offers flexibility in accommodating for site-specific processes. Indeed, these approaches take into account the site-specific variation of the substitution rate along lineages in the very same way the covarion model (Fitch and Markowitz, 1970; Tuffley and Steel, 1998) does, i.e., by authorizing each site and each edge to evolve under its own rate of evolution. Despite these obvious advantages compared to alternative models, IARAs have not been used widely so far, most likely because of a lack of implementation in popular phylogenetics software (although the geometric Brownian model is implemented in PhyTime, a software program that is part of the PhyML package).

3. CALIBRATING THE CLOCK

Molecular dating goes beyond standard phylogenetic reconstruction by separating rates of molecular evolution from actual (i.e., calendar) times. Separating rates and times requires additional data in order to calibrate the estimated trees. Extra data may take three distinct forms: (1) information about the substitution rate, (2) information about paleogeography, or (3) fossil data.

3.1. Dating Without Fossils

Although data about the mutation rates in model organisms is available (see e.g., Drake et al., 1998), information about substitution rates is relatively scarce. The only noticeable exception concerns fast-evolving viruses. The pace at which some viruses (HIV or influenza for instance) evolve indeed makes it possible to obtain sequence data at different points in time such that non-negligible numbers of substitutions have accumulated between these time points (see Shankarappa et al., 1999; Biek et al., 2015, and the excellent internet resource https://nextstrain.org/, Hadfield et al., 2018). It then becomes feasible to infer a "global" substitution rate, and models of (relative) rates of evolution such as those introduced previously describe fluctuations around this trend. I will not elaborate further on serially sampled data for molecular dating here.

Information on the timing of past evolutionary events may also be informed by knowledge about geological events such as the appearance of land bridges or the emergence of islands (Heads, 2005). Indeed, assuming that geography drives speciation [through vicariance, geodispersal, or biological dispersal (Ho et al., 2015b)], the age of an island may, in some circumstances, provide a maximum (i.e., older) age for the birth of ancestral species that colonized this territory. In a symmetric manner, the appearance of land bridges is a necessary condition to explain speciation by vicariance for some species, here again potentially defining a maximum age for some internal nodes in the reconstructed phylogeny. The same land bridges can create barriers of dispersal (e.g., the Isthmus of Panama, that arose 3.5 Mya, is a barrier of migration between the Atlantic and Pacific oceans), thereby providing minimum rather than maximum ages for particular speciation events. Under this line of reasoning, one expects to observe a correlation between the splits corresponding to cladogeneses in a phylogeny and important geological events, mostly involving breakup sequences of Gondwana and Laurasia (Croizat, 1962). Yet, evidence for such correlation is difficult to ascertain (Heads, 2005) and there are examples where important geological events and cladogeneses appear to be disconnected. Hence, many terrestrial animals display strong capability for overseas dispersal so that the appearance of land bridges cannot always be used to establish a maximum age (see de Queiroz, 2005 for a review). Nonetheless, the rising and disappearance of physical connections between geographical regions on the globe is likely to influence the speciation process. Rigorous mathematical modeling, such as that proposed by Landis (2017) for instance, should thus be considered as an important step forward in molecular dating analyses and more generally in ecology.

3.2. Dating With Fossils

3.2.1. Pre-processing of Fossil Data

Fossil data is another source of information commonly used for molecular dating. It consists in a fairly intricate combination of time and morphological information. Time information is only indirect. It is derived from the estimated age of the sediments in which the extinct taxon was collected. The age of these sediments is itself often derived indirectly from that of rock bodies that "bracket" the sediments of interest (Sterli et al., 2013). Moreover, little information is available about the way the different stratigraphic ranges were sampled in general. The data produced by paleontologists generally consist in the combination of a fossil description and the stratigraphic layer in which this fossil was found. Additional information about the experimental design, including the number and types of geological layers surveyed or the excavation techniques that were used, is often difficult to access. That lack of information is problematic. For instance, from a mathematical modeling perspective, finding a particular fossil after searching a single stratigraphic layer is very different from finding the same fossil after examining multiple layers. Finally, the rock record itself is highly heterogeneous in space because of plate tectonics and net erosion, thereby complicating even further the task of finding and interpreting fossil data (Benton et al., 2009).

Morphological information is also difficult to deal with. First, the analysis of one or multiple specimens of a given fossil taxon by paleontologists leads to a selection of "informative characters." These morphological characters are selected based on the phylogenetic signal they convey and result from complex taphonomic processes. Characters showing apomorphies are selected. These characters show evidence of derived states in a subset of extant and extinct species while other species display ancestral states for the same character. The subset of

species showing apomorphies may vary from one morphological character to another. For instance, a given fossil may display a particular character state that is shared by species A and B but not by species C and D. This fossil may also display a second character with derived states shared by A, B, and C but not by D. Both morphological characters point to different phylogenetic placements for the fossilized species, thereby generating *de facto* uncertainty in the calibration of the time tree, even though the phylogenetic relationships between A, B, C, and D may be known with good precision. The fact that the selection of a subset of morphological characters is not random (i.e., characters are selected based on their variability across sampled species) may also be considered problematic (see section 3.2.3). Finally, it is commonplace to assemble morphological characters from multiple specimens of fossils that are deemed to belong to the same ancient species (Parham et al., 2011). Here again, this step relies on the interpretation of evidence by paleontologists. In this context, it is important to stress that the validity of molecular dating analyses relies heavily on the ease with which the whole scientific community can access raw fossil data along with detailed information about how this data was processed in order to define calibration constraints. Without open and systematic access to well-curated and extensive databases of raw fossil data, dating experiments will not be fully reproducible, thereby harming our field of research. Fortunately, rich sources of information about the way fossil data is prepared prior to dating analysis can be found online. For instance, the Fossil Calibration Database[1] (Ksepka et al., 2015) set out to use the rigorous set of guidelines defined in Parham et al. (2011). It provides useful, if partly outdated, information to generate calibration constraints for more than 200 clades and does so on a transparent forum that is open to the whole scientific community. Note however that knowledge about fossil data is constantly evolving (see e.g., Marjanović, 2019) and databases such as the Fossil Calibration Database require ongoing and constant efforts in order to remain relevant.

3.2.2. Expert-Based Analysis of Fossil Data

A fossil provides a minimum age for the smallest extant clade to which it belongs, i.e., the youngest node from which it as well as any two or more extant taxa are descended. The phylogenetic position of a fossil is determined either by a phylogenetic analysis of (a subset of) its parsimony-informative characters or by a manual comparison with a list of apomorphies derived from a prior phylogenetic analysis or from prephylogenetic taxonomic work. Both approaches leave varying amounts of uncertainty, depending in part on how fragmentary the fossil in question is. Every fossil-based calibration thus contains age uncertainty and phylogenetic uncertainty.

Fossils that branch near the tips in the calibrated clade define looser younger ages for that clade compared to older fossils that branch closer to the basal node. As a consequence, paleontologists are always eager to discover older fossils that still belong to the clade of interest. Unfortunately, older fossils did not have sufficient time to accumulate as many apomorphies as younger fossils did. Early members of a given taxon were also likely to be scarce and occupy a limited geographic area (Marshall, 2019). As a consequence, older fossils are also the most difficult to associate to well-defined clades. Considerable uncertainty into the placement of these fossils in the phylogenetic tree may then hamper further analysis. At the other extremity, young fossils are likely to sit at the end of long branches, along which numerous morphological changes accumulated that took place along this branch only, potentially leading to a saturation of the signal.

Linking a given fossil to a group of taxa as is done here involves a considerable amount of work and discussions among multiple experts, as demonstrated by the wealth of information provided in the journal Palaeontologia Electronica for animals, for instance (see https://palaeo-electronica.org/content/fc-1). Note however that only taxa considered as important receive high levels of scrutiny. Hence, existing databases can be used to find well-curated calibration information for "popular" taxa, in which cases researchers rely on previous rigorous work by paleontologists in order to calibrate their own analysis. Calibrating an analysis of less well-characterized taxa generally involves a thorough search of the primary literature followed by an in-depth comparative analysis of raw morphological data as briefly explained in the previous paragraphs (see also section 3.2.1 and Saladin et al., 2017 for an example).

3.2.2.1. The fossilized-birth-death model and other model-based approaches to calibration

As seen above, defining younger bounds for calibration intervals is difficult, although the fossil record provides rich information to conduct this task in some cases. Deriving older bounds is even more challenging. Indeed, the younger bound for a given clade does not put any constraint on the older bound for a nested clade. In other words, the younger bound for the MRCA of a clade made of species A, B and C does not convey any relevant information about the older bound for the MRCA of the subclade made of A and B. Beside the intervention of paleontologists, probabilistic modeling can also be used to define older calibration bounds. The two techniques described below rely on mathematical models depicting the processes governing fossilization in order to define older bounds of calibration constraints whose position in the tree is defined as outlined above.

Tavaré et al. (2002) developed a probabilistic model and an inference technique to estimate the length of the temporal gap between the oldest fossil and the time of the MRCA of a given clade, along with a confidence interval for that length. This method uses as input the number of extant species in the clade whose age is to be calibrated, the ages of the relevant stratigraphic layers, the number of fossils found in them and the (relative) fossil sampling intensities in these layers. Although this approach relies on a sound mathematical model of species diversification and properly accommodates for the specifics of the collection of fossil data, it has not been used widely so far, most likely because of the lack of software implementing it and, perhaps, the lack of information regarding fossil sampling intensities. In a very

[1] https://fossilcalibrations.org/

thorough study about the treatment of fossil data to calibrate the molecular clock, Marshall (2019) also described a method similar to that of Tavaré et al. (2002), where the time elapsed between the divergence date to be calibrated and the age of the oldest fossil of a given focal clade has a probabilistic distribution whose parameters depend on the number of fossil localities for that clade.

Another attempt to tackle the same issue was proposed by Stadler (2010) and Didier et al. (2012). The proposed approach is based on a probabilistic model describing the tree-generating process. Yet, it still relies on expert knowledge for placing calibration constraints in a tree. The model put forward in these two studies lies now at the core of the so-called "tip-dating" methods whereby calibrating the molecular clock derives from time information available at the tips of the phylogeny corresponding to both extant and extinct (and fossilized) species. Both approaches model the stochastic process generating a tree including sampled extant and fossil species. The so-called fossilized birth-death process (FBD) assumes that lineages give birth to new species or die at given per capita rates, which are deemed to be constant during the course of evolution. An ancestral lineage may also fossilize, an event that takes place at a particular rate, which is to be estimated from the data. After sampling, only a subset of extant and extinct species are available for the analysis (**Figure 4**). The joint probability density of the age of "observed" lineage splits in a phylogeny given the time of sampling of extant and extinct taxa, along with the birth, death, fossilization rates, and sampling fraction, can be evaluated analytically. When considering the calibration of one particular node using the FBD model, one has to truncate the joint probability density of node ages such that its value is null whenever the node is younger than the oldest fossil used for this particular calibration. This truncation is fairly straightforward to deal with in the context of Bayesian molecular dating using Markov Chain Monte Carlo techniques, and Heath et al. (2014) provide an analytical solution to this problem. Younger bounds for calibration intervals thus derive directly from the analysis of the fossil data. Older bounds are defined indirectly and depend on the values of the FBD model parameters. More specifically, information about the relative node heights derives mainly from the analysis of genetic sequences only (i.e., calibration data is less informative about the phylogeny itself). These relative heights then serve as a basis to infer the parameters of the FBD model which, in turn, help defining older bounds for the ages of the calibrated nodes. One thus relies here on a hierarchical model whereby the phylogeny is treated alternatively as a parameter and as data, depending on the level in the hierarchy that is considered.

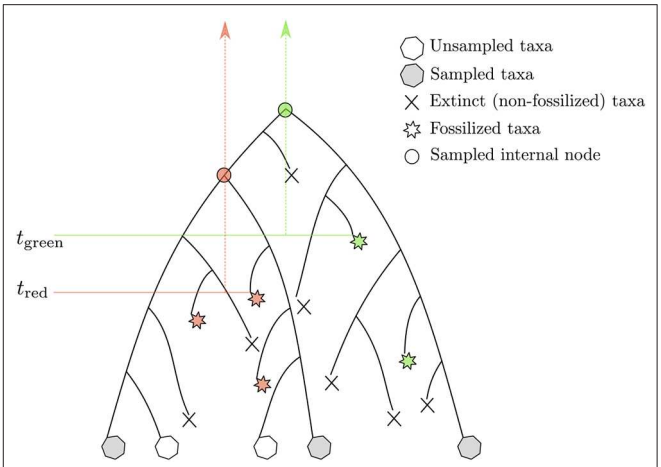


FIGURE 4 | Rationale underlying the calibration of the molecular clock with the FBD model. The stars point to fossils. The fossils in red (green) calibrate the internal nodes presented as disks of the same color. The vertical arrows define the corresponding calibration intervals. t_{red} and t_{green} refer to the lower bounds associated to these two intervals. The corresponding upper bounds are determined by the birth, death, and fossilization parameters of the FBD model.

3.2.2.2. Performance of model-based approaches for defining calibration constraints

The two techniques aforementioned are not the only ones that can be used to define calibration constraints. In fact, any tree-generating model may extract some information from the available data about the marginal age of nodes used for calibrating the clock. In the following, I assess the relevance of tree-generating models for specifying calibration constraints by focusing on the particular case where genetic sequences are of infinite length and the correct models of sequence evolution and rate variation across edges are used for molecular dating. The relative node ages are then known with maximum precision. We observe a single fossil that helps us determine the scaling factor of all node ages, thereby allowing us to infer absolute (rather than relative) node ages. I will here consider the simple case where only three taxa are examined. t_1 denotes the relative age of node n_1, the MRCA of these three species, and t_2 is the relative age of the calibrated node, n_2 (see insert in **Figure 5**). Time grows backward with the present time set to zero. Finally, u is the age of the fossil. We therefore have $0 \leq u \leq \alpha t_2 \leq \alpha t_1$, where α is a scaling factor that defines the correspondence between relative and absolute node ages. Equation (5) in Stadler (2012) gives the expression for $p_{\text{BD}}(\alpha t_2|\alpha t_1, \lambda, \mu)$, the probability density of αt_2 given αt_1 (and λ plus μ, the birth and death rates respectively, as well as the sampling fraction which we assume to be equal to 1.0 here) under the birth-death process with sampling, conditioned on the time of the MRCA of the three sampled species αt_1. This expression serves as a basis to derive that of $p_{\text{BD}}(\alpha t_2|\alpha t_1, t_2 \geq u, \lambda, \mu) = p_{\text{BD}}(\alpha t_2|\alpha t_1, \lambda, \mu)/\int_u^{\alpha t_1} p_{\text{BD}}(x|\alpha t_1, \lambda, \mu)\text{d}x$, i.e., the probability density of the age of node n_2 conditioned on its value being greater than the age of the fossil used for the calibration, u.

Figure 5 focuses on the value of $p_{\text{BD}}(\alpha t_2|\alpha t_1, \alpha t_2 \geq u, \frac{\lambda}{\alpha}, \frac{\mu}{\alpha})\text{d}(\alpha t_2)$ (y-axis) as a function of α (x-axis). This expression corresponds to the conditional probability density of the age of n_2 being equal to αt_2, given that the age of n_1 is equal to αt_1 while the birth and death rates are equal to $\frac{\lambda}{\alpha}$ and $\frac{\mu}{\alpha}$, respectively. α is thus used here to control the pace at which the tree process unfolds. Values of this parameter greater than one therefore lead to a decrease in the rate at which birth and death of lineages take place, thereby pushing nodes in the phylogeny to be older. Values of the parameter smaller than one have the opposite effect. Note however that precise characterization of the relationship

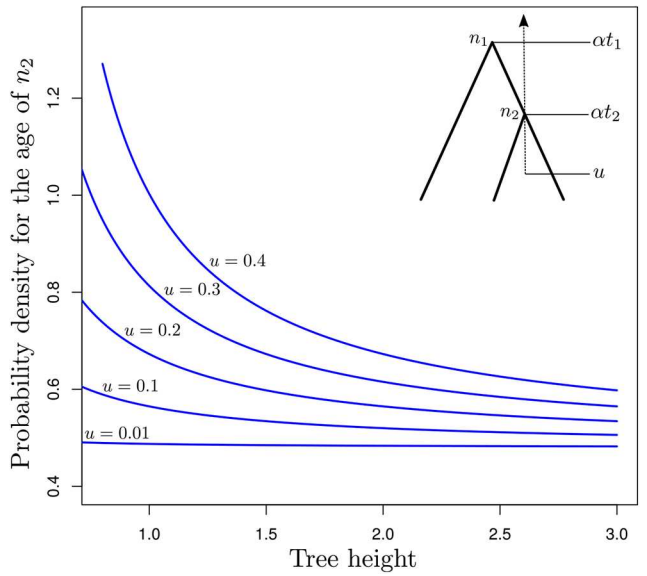


FIGURE 5 | Probability density of the age of n_2 as a function of the tree length for various calibration constraints. The y-axis gives $p_{\text{BD}}(\alpha t_2 | \alpha t_1, \alpha t_2 \geq u, \frac{\lambda}{\alpha}, \frac{\mu}{\alpha})\text{d}(\alpha t_2)$ as a function of α (x-axis) for a three-taxon rooted tree with n_2 being the youngest internal node and n_1, the root node. This tree has fixed relative node ages $t_2 = 0.5$ (age of n_2) and $t_1 = 1.0$ (age of n_1). The corresponding absolute node ages are αt_2 and αt_1, where α is a scalar which takes values in $[u/t_2, 3]$. u is the age of the fossil calibrating αt_2, defining the younger bound for the corresponding node.

between α and the ages of n_1 and n_2 would deserve a more thorough examination.

When the fossil is found very near the tips of the tree (i.e., $u = 0.01$), the age distribution of node n_2 is not influenced by the value of α, i.e., older ages and lower birth and death rates have the same probability as younger ages and higher rates. This observation is not surprising: parameters of the birth-death model are not identifiable from the relative node ages only, and multiplying the node heights by a given scaling factor while dividing the birth and death rates by the same factor leads to absolute node ages with the same probabilities. Increasing the value of u, i.e., considering older ages for the fossil, sees increasing amounts of information about absolute node ages. Indeed, there are substantial differences between probabilities of trees with young node ages (high probabilities, small values of α) compared to that of trees with older ages (lower probabilities, large values of α). Therefore, with older fossils, the probabilistic distribution of the absolute age of n_2 is no longer flat as it is when $u \simeq 0.0$ and it becomes feasible to use the tree-generating process to define meaningful older bounds for calibration intervals. Results in **Figure 5** therefore suggest that the calibration interval will be tighter for analyses that rely on older fossils compared to younger ones. A better understanding of this result may be obtained by considering how trees could be simulated under various constraints on the age of n_2. The simplest approach would be to simulate under the standard birth-death process and then discard the generated trees where n_2 is younger than u. It then becomes clear that the older u, the larger the proportion of discarded trees. The valid trees therefore represent a smaller proportion of all possible trees. In layman's terms, these trees therefore convey more information about the distribution of node ages compared to the case where no constraints apply.

The previous analysis focuses on the relationship between the age of the young bound for a calibration interval and the probabilistic distribution of the corresponding older bound. Yet, the applicability of these results depends on how accurate estimates of birth, death and, in the case of the FBD model, fossilization rates are (assuming, here again, complete sampling). In particular, the fossilization rate might be very difficult to estimate as it is influenced by a variety of factors. First, taphonomic phenomena are the source of major biases in the fossil record since organisms with hard body parts have a greater chance of being represented in this record. Moreover, the majority of fossils result from material deposited on the bottom of water bodies, thereby adding another source of bias shaping the fossil record. This heterogeneity in the fossil record is expected to show at large time scales. Smaller time scales, however, are expected to be less impacted by this phenomenon and the Poisson process arguably provides here a better description of the fossilization process.

Inaccurate estimates of the fossilization rates may impact the inference of node ages substantially (Matschiner, 2019). **Figure 6** shows the influence of the rate of fossilization on the probabilistic distribution of the age of a node that may then serve as a basis to calibrate the dating analysis. Didier and Laurin (2020) describe a method to derive the marginal distribution for the age of an internal node given the ages of all tips in a set of phylogenetic tree topologies. Using the same data set as in their article, comprising 109 dated extinct taxa covering Amniota and Diadectomorpha, I used the software DateFBD (available from https://github.com/gilles-didier/DateFBD) to infer the age of the Amniota clade. The estimation is conditioned here on a particular tree topology, in which Diadectomorpha is placed as outgroup. This topology is one of thousands of equally-most-parsimonious trees obtained from the analysis of a matrix of morphological characters (Didier and Laurin, 2020).

Although the three distributions are not conditioned here on having the same younger bound (i.e., by forcing the three densities to have the same 95% upper quantile for instance), substantial differences in their modes and the variances are observed depending on the value of the fossilization rate. Note however that the uncertainty around the fossilization rate estimated by Didier and Laurin (2020) is much smaller than the two orders of magnitude considered here. Their analysis focused on 109 data points corresponding to a rich collection of fossils. In situations where the fossil record is not as extensive as it is here, our results suggest that the conversion of node heights from molecular into calendar time units is highly sensitive to the fossilization rate estimate. Ages are indeed shifted toward older values for large values of that rate and vice versa. Accurate estimation of that rate is therefore paramount to the accurate inference of node ages. In practice, the sensitivity of the age estimates to the prior distribution on fossilization rate should thus be assessed on a systematic basis when using the FBD model, or any other tree-generating process, for molecular dating.

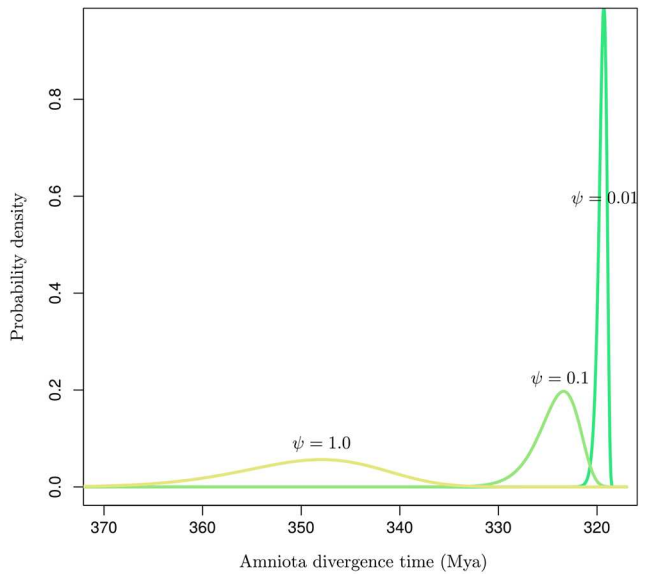


FIGURE 6 | Distribution of the time of divergence of Amniota under the FBD model for various values of the fossilization rate (ψ). The birth and death rates were set to those given in Didier and Laurin (2020).

Stadler et al. (2018) recently proposed an extension of the original FBD model that accounts for multiple fossils per species and different modes of speciation. Although this work inches toward a more realistic modeling of the actual mechanisms generating the observations, more efforts are still needed in order to assess the robustness of the FBD model to deviations from the current hypotheses regarding the mechanisms of fossilization and the specifics of the data collection process. In particular, calibration information is generally provided by the oldest fossils in their respective clades instead of collections of randomly sampled fossils. In its current form, the FBD model does not take this information into account, which may result in substantial node age overestimation (Matschiner, 2019). The FBD also neglects information related to the absence of fossils in particular geological strata. In other words, this approach expects to find fossils in geological strata in which there is strong evidence *against* their presence. The absence of fossils is valuable information and node dating techniques provide a framework that helps accommodating for it in a simplified manner (see Marjanović and Laurin, 2007 for an example). Dealing with the absence of fossils in certain geological strata is not straightforward though. "Absence of evidence" should be distinguished from "evidence of absence." Indeed, in the context of interest here, "absence of evidence" has to do with the way the sampling of geological strata was conducted. Models such as the birth-death skyline plot (Stadler et al., 2013) provide a relevant framework to accommodate for the variation in time of sampling intensity. "Evidence of absence" corresponds instead to the outcome of the data-generating process. Uneven preservation of fossils is one of the phenomena involved here. These processes need to be dealt with through adequate probabilistic modeling.

3.2.2.3. Calibrating using marginal distributions

Molecular dating based on the FBD model can be considered a mechanistic approach as it relies on a model that mimics the actual process underlying splitting or extinction of lineages and fossilization events. The use of univariate probabilistic distributions to describe the age of certain clades has more to do with a phenomenological approach instead. Drummond et al. (2006) first introduced probabilistic distributions for node age calibration in the context of molecular dating using a fully Bayesian approach (but see Hedges and Kumar, 2004; Yang and Rannala, 2005 for an earlier introduction of this idea). They used normal and log-normal distributions but did not provide detailed explanations about the way the parameters of these distributions were selected from the analysis of the fossil record. Current approaches, available in BEAST (1 and 2) and MrBayes, implement standard statistical distributions, such as exponential, lognormal, or normal densities, with offset values set to the younger (i.e., minimum) age of the corresponding clade. mcmctree offers a selection of more sophisticated probabilistic distributions (see Yang and Rannala, 2005) but still relies on the same rationale. The older bound for each calibration is then defined by a variance parameter that controls for the probability that the age of the clade of interest is older than a given value, corresponding typically to the 95% upper quantile of the distribution. Hence, as opposed to the previous approach based on the FBD model, it is relatively straightforward here to define older bounds for each calibration interval. However, in its classical formulation, the "marginal distribution" approach does not account for uncertainty in the placement of the fossils in the phylogeny, which constitutes a serious limitation of that technique (but see Guindon, 2018 and below).

Drummond et al. (2006) account for the interaction between marginal priors[2] and the joint distribution of the other node ages using an approach that does not respect fundamental rules of calculus with probabilities. Their approach results in multiple distributions (one defined by the marginal prior plus one derived from the tree-generating process) applying to the same node ages (Heled and Drummond, 2011; Warnock et al., 2015; Rannala, 2016). In Yang and Rannala (2005), calibrated and non-calibrated nodes are clearly separated in the derivation of a joint prior density of node ages that accommodates for fossil information. More specifically, let $T = \{T_1, \ldots, T_{n-1}\}$ denote the vector of all internal node ages $T_1 \geq \ldots \geq T_{n-1}$ and Ψ the ranked tree topology with n tips. Both T and Ψ are random variables here. t and τ denote realizations of these random variables. e and i denote subsets of taxa and the corresponding time constraints respectively (R.V.: E and I). Each subset of taxa in e defines a clade and the corresponding element in i defines the time interval for the age of that clade. The joint density of the vector of node

[2]These marginal distributions are in fact derived from fossil data, and should thus not be considered prior densities *per se*, although most studies refer to them using this term based on the fact that marginal distributions are instantiated prior to observing genetic sequences.

ages t and the tree topology τ, given calibration constraints e, i, is as follows:

$$p_{T,\Psi}(t,\tau|E=e,I=i) = p_{T_{-c}}(t_{-c}|T_c=t_c,\Psi=\tau)\times p_{T_c}(t_c|\Psi=\tau,E=e,I=i) \times \Pr(\Psi=\tau|E=e,I=i), \quad (13)$$

where T_{-c} is the vector of node ages that are not subject to any calibration constraint. Yang and Rannala (2005) give an expression for the conditional density $p_{T_{-c}}(t_{-c}|t_c,\tau)$ under the birth-death model with sampling. $p_{T_c}(t_c|\tau,e,i)$ is the joint density of the ages of all calibrated nodes. Yang and Rannala (2005) define this joint density as the product of the marginal "prior" densities used for calibration purposes. This definition is problematic since the joint density of calibrated ages is conditioned on the ranked tree topology. As a consequence, the calibrated ages cannot be considered independent from one another. In fact, when conditioning on a particular ranked tree, some combinations of node ages are not observable and the corresponding joint density should in fact be equal to zero. The probability of such "non-observable" outcomes depends on the parameters of the tree-generating process and cannot be simply ignored by the MCMC analyses used for Bayesian molecular dating. The discrepancy between the product of marginal calibration densities and the actual joint density that is used by these inference techniques is arguably the most obvious manifestation of the same issue (Heled and Drummond, 2011; Warnock et al., 2015). Rannala (2016) acknowledges this conundrum, only to reach the conclusion that *"the objective of preserving marginal calibrations is impossible to attain."*

In Guindon (2018), I describe a new approach to node dating and provide a solution to the problem of uncertainty in the placement of fossil lineages in the tree. In this work, calibration constraints, e and i, are no longer considered as data. Instead, one acknowledges here that the actual data correspond to the fossils, noted as α (R.V.: F) and the calibration constraints then become parameters of the model, with inherent uncertainty. More specifically, the joint density of the time-tree and the calibration parameters given fossil data is now expressed as follows:

$$p_{T,\Psi,E,I}(t,\tau,e,i|F=\alpha) = p_{T,\Psi}(t,\tau|E=e,I=i) \times p_{E,I}(e,i|F=\alpha), \quad (14)$$

where E and I denote the random variables corresponding to subsets of taxa and the corresponding time intervals that, altogether, define calibration constraints. The term $p_{E,I}(e,i|F=\alpha)$ serves as a basis to incorporate uncertainty in the calibration constraint due to ambiguity in interpreting the fossil data. In practice, experts may decide that a given fossil calibrates the age of the MRCA of species A and B with probability p, while the clade defined by species A, B, and C is calibrated by the same fossil with probability $1-p$, thereby effectively accommodating for uncertainty related to fossil data. Moreover, the probabilistic distribution of the calibration constraints is not conditioned here on the tree topology. The combination of multiple calibration constraints therefore does not suffer from the issues mentioned above that are responsible for the discrepancy between "realized" distributions of calibrated node ages and the corresponding marginal priors. In other words, the marginal "priors" agree with their joint density as given by $p_{E,I}(e,i|\alpha)$ (yet, the marginal distributions of calibrated node ages derived from $p_{T,\Psi}(t,\tau,e,i|\alpha)$ still disagree with these marginal priors, as expected from Equation 14).

Finally, the same approach may be extended in order to incorporate knowledge about the way fossils were collected. Equation (14) would then yield:

$$p_{T,\Psi,E,I}(t,\tau,e,i|F=\alpha,S=s) = p_{T,\Psi}(t,\tau|E=e,I=i) \times p_{E,I}(e,i|F=\alpha)\times p_F(\alpha|S=s), \quad (15)$$

where the random variable S conveys information about the way sampling was conducted when collecting fossil data. For instance, the probability of observing a fossil of an ancient species that lived X million years ago would be equal to zero if sampling was conducted in geological strata corresponding to time intervals that do not include X. This term could also serve as a basis to incorporate geographical information in the analysis, translating the fact that some fossils are more likely to be found in particular regions and less in other areas.

3.2.3. Model-Based Analysis of Fossil Data

The molecular dating methods presented above all rely on expert knowledge to determine, even approximately, where in the phylogeny fossil lineages should be placed. Yet, computational approaches can replace expert judgment here. Phylogenetic analyses of morphological data, including fossils, are most often conducted using parsimony. However, parametric methods like Bayesian inference, which use probabilistic models to describe the evolution of the morphological characters of which fossil data consist, can also be used. Such approaches offer the advantage that molecular data can be included in the same analysis (the "total-evidence" approach). Further, they can be combined with a tip-dating analysis to derive the joint posterior distribution of internal node ages from the combined probabilistic analysis of molecular and morphological data.

Pyron (2011) and Ronquist et al. (2012a) implemented this approach and analyzed concatenated alignments of molecular and morphological data. In this large matrix, molecular data is observed only for present-day sampled species (although ancient DNA is also a source of molecular data) while morphological data along with time information are available for both the sampled fossils and modern species. A phylogeny that incorporates fossils as *bona fide* taxa is then built from the analysis of this data. In this context, it is thus relatively straightforward to account for uncertainty in the placement of the fossil lineages. All the internal nodes on the path between a fossil tip and the root are constrained to be older than the age of the fossil. In theory, the time elapsed between fossil tips and the present could help define the rate of morphological evolution which would then serve as a basis to express all node ages in calendar time units.

Ronquist et al. (2012a) opted for an *ad hoc* approach where the rate of molecular evolution was first estimated using a node dating approach and then used as prior information in a subsequent dating analysis based on a total-evidence approach. Pyron (2011) also relied on a classical node dating approach to

specify the distribution of the age of the root node. More recently, Gavryushkina et al. (2017) relied on a node dating approach too to calibrate the origin of the tree-generating process (thereby indirectly defining a marginal prior for the age of the root node), combined with a prior distribution on the rate of morphological evolution. Moreover, their approach rests on the FDB model as the tree-generating process, thereby using information from rates of birth, death and fossilization to further inform the absolute age of internal nodes (see section 3.2.2).

The three studies cited above are tip-dating analyses based on total evidence. They all relied on a classical node dating approach to derive absolute node ages from the combined analysis of molecular and fossil data. Although there is nothing wrong with mixing various approaches, the systematic reliance on node dating suggests that, at least in these three cases, morphological data alone may not have conveyed enough signal to infer reliable rates of morphological character evolution in practice, a necessary step for inferring absolute node ages in the absence of additional information for calibrating the clock. In fact, one may even wonder whether such a rate exists at all. Each morphological character having its own state space, one may indeed question whether it is meaningful to refer to an expected number of character changes per unit of calendar time (see Goloboff et al., 2018 for a discussion of this issue along with Goloboff et al., 2019). Furthermore, unlike for nucleotide or amino-acid characters, it is not always straightforward to define the alphabet of states for each morphological character (see e.g., Gavryushkina et al., 2017). dos Reis et al. (2016) also indicate that ascertainment biases due to the selection of parsimony-informative morphological characters from raw fossil data is difficult to deal with from a computational perspective and a proper correction, able to handle ambiguous alphabets of character states, is not implemented in any current software program for molecular dating.

4. CONCLUSIONS

Modeling the evolution of the rate of molecular evolution and accounting for fossil data are two challenging tasks that lie at the core of molecular dating techniques. Although much work has been done on these different aspects, in-depth exposition of the simplifications, the approximations, and the assumptions behind the proposed approaches helps gain a better understanding of their inherent strengths and limitations.

For instance, clearly separating the substitution rate trajectories that depict the fluctuations of the instantaneous substitution rates, from the average rates along edges of the phylogeny, leads to interesting observations. In particular, the "not-so-strict" clock model in which instantaneous rates vary while averages do not, could serve as a basis to revisit the clock hypothesis. At the very least, it constitutes an intermediate model between the strict and relaxed clock models that is worth considering. Furthermore, close examination of uncorrelated clock models reveals some of their shortcomings. The exponential clock model, in particular, has statistical properties that are not realistic from a biological perspective. More generally, uncorrelated clock models lead to stronger deviations from the strict clock constraint in trees with large numbers of tips compared to smaller trees, thereby revealing sampling-consistency issues that should be of concern. Autocorrelated clock models behave more sensibly altogether. Some of these models explicitly accommodate the variation of both instantaneous and average substitution rates without extra computational cost, making them superior to uncorrelated models from that point of view.

Taking into account fossil data in molecular dating experiments is another challenging statistical problem. The most recent techniques bet on an "all-modeling" approach that is hindered by a number of important limitations. In particular, unrealistic assumptions underlying the probabilistic models describing the evolution of selected morphological features should be of serious concern to total-evidence approaches. Assuming that fossils are "presence-only" data is also problematic. However, valuable information about the absence of some fossils in older geological strata is often available. The most recent inference techniques, including tip-dating and all approaches based on the fossilized-birth-death model, ignore this information, thereby enabling node age estimates that potentially contradict what is known from the fossil record. Node dating techniques rely on expert knowledge to define the position of fossils in the phylogeny plus the younger (and, oftentimes, the older) age bound(s) for the calibrated clades. Although expert knowledge involves subjectivity, which can be perceived as a weakness, one could argue that these approaches make better use of the available data for now. The future of molecular dating probably lies at the frontier between "all-expert" and "all-model" approaches whereby experts will provide prior information to plug into relevant statistical models for describing curated fossil data.

In any case, molecular dating will undoubtedly keep playing a crucial role in biology in the future. Our understanding of important phenomena such as species diversification or dispersal, population migration and demography, or the molecular signature resulting from environmental changes, depends on our ability to date past evolutionary events. The wealth of available techniques to perform this task provides a powerful set of tools to make progress in this direction. Yet, in-depth analysis of the mathematical and biological properties of the proposed new techniques, combined with rigorous and extensive assessments of their implementations, will be decisive to ensuring the soundness of our findings.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation, to any qualified researcher.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

This work was supported by the ANR project GENOSPACE.

ACKNOWLEDGMENTS

I would like to thank the associate editor, MA, plus three reviewers, and DM for their comments and suggestions that helped me to improve this article.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.00526/full#supplementary-material

REFERENCES

Aris-Brosou, S., and Yang, Z. (2002). Effects of models of rate evolution on estimation of divergence dates with special reference to the metazoan 18S ribosomal RNA phylogeny. *Syst. Biol.* 51, 703–714. doi: 10.1080/10635150290102375

Baer, C. F., Miyamoto, M. M., and Denver, D. R. (2007). Mutation rate variation in multicellular eukaryotes: causes and consequences. *Nat. Rev. Genet*. 8:619. doi: 10.1038/nrg2158

Barton, N., Etheridge, A., and Véber, A. (2010). A new model for evolution in a spatial continuum. *Electron. J. Probabil.* 15, 162–216 doi: 10.1214/EJP.v15-741

Benton, M., Donoghue, P., and Asher, R. (2009). Calibrating and constraining molecular clocks. *Timetree Life* 35:86.

Biek, R., Pybus, O. G., Lloyd-Smith, J. O., and Didelot, X. (2015). Measurably evolving pathogens in the genomic era. *Trends Ecol. Evol.* 30, 306–313. doi: 10.1016/j.tree.2015.03.009

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., et al. (2014). BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* 10:e1003537. doi: 10.1371/journal.pcbi.1003537

Bromham, L., Duchêne, S., Hua, X., Ritchie, A. M., Duchêne, D. A., and Ho, S. Y. (2018). Bayesian molecular dating: opening up the black box. *Biol. Rev.* 93, 1165–1191. doi: 10.1111/brv.12390

Cartwright, R. A., Lartillot, N., and Thorne, J. L. (2011). History can matter: non-Markovian behavior of ancestral lineages. *Syst. Biol.* 60, 276–290. doi: 10.1093/sysbio/syr012

Clauset, A., and Erwin, D. H. (2008). The evolution and distribution of species body size. *Science* 321, 399–401. doi: 10.1126/science.1157534

Condamine, F. L., Nagalingum, N. S., Marshall, C. R., and Morlon, H. (2015). Origin and diversification of living cycads: a cautionary tale on the impact of the branching process prior in Bayesian molecular dating. *BMC Evol. Biol.* 15:65. doi: 10.1186/s12862-015-0347-8

Cox, J. C., Ingersoll, J. E. J., and Ross, S. A. (2005). "A theory of the term structure of interest rates," in *Theory of Valuation*, eds S. Bhattacharya and G. M. Constantinides (World Scientific), 129–164. doi: 10.1142/9789812701022_0005

Croizat, L. (1962). *Space, Time, Form: the Biological Synthesis*. Caracas: L. Croizat.

Currat, M., Arenas, M., Quilodrán, C. S., Excoffier, L., and Ray, N. (2019). SPLATCHE3: simulation of serial genetic data under spatially explicit evolutionary scenarios including long-distance dispersal. *Bioinformatics* 35, 4480–4483. doi: 10.1093/bioinformatics/btz311

de Queiroz, A. (2005). The resurrection of oceanic dispersal in historical biogeography. *Trends Ecol. Evol.* 20, 68–73. doi: 10.1016/j.tree.2004.11.006

Didier, G., and Laurin, M. (2020). Exact distribution of divergence times from fossil ages and tree topologies. *bioRxiv [preprint]*. doi: 10.1093/sysbio/syaa021

Didier, G., Royer-Carenzi, M., and Laurin, M. (2012). The reconstructed evolutionary process with the fossil record. *J. Theor. Biol.* 315, 26–37. doi: 10.1016/j.jtbi.2012.08.046

dos Reis, M., Donoghue, P. C., and Yang, Z. (2016). Bayesian molecular clock dating of species divergences in the genomics era. *Nat. Rev. Genet*. 17:71. doi: 10.1038/nrg.2015.8

dos Reis, M., Gunnell, G. F., Barba-Montoya, J., Wilkins, A., Yang, Z., and Yoder, A. D. (2018). Using phylogenomic data to explore the effects of relaxed clocks and calibration strategies on divergence time estimation: primates as a test case. *Syst. Biol.* 67, 594–615. doi: 10.1093/sysbio/syy001

Drake, J. W., Charlesworth, B., Charlesworth, D., and Crow, J. F. (1998). Rates of spontaneous mutation. *Genetics* 148, 1667–1686.

Drummond, A. J., Ho, S. Y., Phillips, M. J., and Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4:e88. doi: 10.1371/journal.pbio.0040088

Drummond, A. J., and Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7:214. doi: 10.1186/1471-2148-7-214

Drummond, A. J., and Suchard, M. A. (2010). Bayesian random local clocks, or one rate to rule them all. *BMC Biol.* 8:114. doi: 10.1186/1741-7007-8-114

Duret, L., and Galtier, N. (2009). Biased gene conversion and the evolution of mammalian genomic landscapes. *Annu. Rev. Genom. Hum. Genet*. 10, 285–311. doi: 10.1146/annurev-genom-082908-150001

Esser, C., Ahmadinejad, N., Wiegand, C., Rotte, C., Sebastiani, F., Gelius-Dietrich, G., et al. (2004). A genome phylogeny for mitochondria among α-proteobacteria and a predominantly eubacterial ancestry of yeast nuclear genes. *Mol. Biol. Evol.* 21, 1643–1660. doi: 10.1093/molbev/msh160

Fitch, W. M., and Markowitz, E. (1970). An improved method for determining codon variability in a gene and its application to the rate of fixation of mutations in evolution. *Biochem. Genet*. 4, 579–593. doi: 10.1007/BF00486096

Fletcher, W., and Yang, Z. (2009). INDELible: a flexible simulator of biological sequence evolution. *Mol. Biol. Evol.* 26, 1879–1888. doi: 10.1093/molbev/msp098

Gavryushkina, A., Heath, T. A., Ksepka, D. T., Stadler, T., Welch, D., and Drummond, A. J. (2017). Bayesian total-evidence dating reveals the recent crown radiation of penguins. *Syst. Biol.* 66, 57–73. doi: 10.1093/sysbio/syw060

Gillespie, J. H. (1994). *The Causes of Molecular Evolution, Vol. 2*. Oxford University Press.

Gojobori, T., Li, W.-H., and Graur, D. (1982). Patterns of nucleotide substitution in pseudogenes and functional genes. *J. Mol. Evol.* 18, 360–369. doi: 10.1007/BF01733904

Goloboff, P. A., Pittman, M., Pol, D., and Xu, X. (2019). Morphological data sets fit a common mechanism much more poorly than DNA sequences and call into question the MKV model. *Syst. Biol.* 68, 494–504. doi: 10.1093/sysbio/syy077

Goloboff, P. A., Torres, A., and Arias, J. S. (2018). Weighted parsimony outperforms other methods of phylogenetic inference under models appropriate for morphology. *Cladistics* 34, 407–437. doi: 10.1111/cla.12205

Guindon, S. (2012). From trajectories to averages: an improved description of the heterogeneity of substitution rates along lineages. *Syst. Biol.* 62, 22–34. doi: 10.1093/sysbio/sys063

Guindon, S. (2018). Accounting for calibration uncertainty: Bayesian molecular dating as a "doubly intractable" problem. *Syst. Biol.* 67, 651–661. doi: 10.1093/sysbio/syy003

Hadfield, J., Megill, C., Bell, S. M., Huddleston, J., Potter, B., Callender, C., et al. (2018). Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* 34, 4121–4123. doi: 10.1093/bioinformatics/bty407

Halpern, A. L., and Bruno, W. J. (1998). Evolutionary distances for protein-coding sequences: modeling site-specific residue frequencies. *Mol. Biol. Evol.* 15, 910–917. doi: 10.1093/oxfordjournals.molbev.a025995

Heads, M. (2005). Dating nodes on molecular phylogenies: a critique of molecular biogeography. *Cladistics* 21, 62–78. doi: 10.1111/j.1096-0031.2005.00052.x

Heath, T. A., Holder, M. T., and Huelsenbeck, J. P. (2012). A Dirichlet process prior for estimating lineage-specific substitution rates. *Mol. Biol. Evol.* 29, 939–955. doi: 10.1093/molbev/msr255

Heath, T. A., Huelsenbeck, J. P., and Stadler, T. (2014). The fossilized birth-death process for coherent calibration of divergence-time estimates. *Proc. Natl. Acad. Sci. U.S.A.* 111, E2957–E2966. doi: 10.1073/pnas.1319091111

Hedges, S. B., and Kumar, S. (2004). Precision of molecular time estimates. *Trends Genet.* 20, 242–247. doi: 10.1016/j.tig.2004.03.004

Heled, J., and Drummond, A. J. (2011). Calibrated tree priors for relaxed phylogenetics and divergence time estimation. *Syst. Biol.* 61, 138–149. doi: 10.1093/sysbio/syr087

Ho, S. Y., Duchêne, S., and Duchêne, D. (2015a). Simulating and detecting autocorrelation of molecular evolutionary rates among lineages. *Mol. Ecol. Resour.* 15, 688–696. doi: 10.1111/1755-0998.12320

Ho, S. Y., Tong, K. J., Foster, C. S., Ritchie, A. M., Lo, N., and Crisp, M. D. (2015b). Biogeographic calibrations for the molecular clock. *Biol. Lett.* 11:20150194. doi: 10.1098/rsbl.2015.0194

Huelsenbeck, J. P., Larget, B., and Swofford, D. (2000). A compound Poisson process for relaxing the molecular clock. *Genetics* 154, 1879–1892.

Kishino, H., Thorne, J., and Bruno, W. (2001). Performance of a divergence time estimation method under a probabilistic model of rate evolution. *Mol. Biol. Evol.* 18, 352–361. doi: 10.1093/oxfordjournals.molbev.a003811

Ksepka, D. T., Parham, J. F., Allman, J. F., Benton, M. J., Carrano, M. T., Cranston, K. A., et al. (2015). The Fossil Calibration Database-a new resource for divergence dating. *Syst. Biol.* 64, 853–859. doi: 10.1093/sysbio/syv025

Landis, M. J. (2017). Biogeographic dating of speciation times using paleogeographically informed processes. *Syst. Biol.* 66, 128–144. doi: 10.1093/sysbio/syw040

Lanfear, R., Kokko, H., and Eyre-Walker, A. (2014). Population size and the rate of evolution. *Trends Ecol. Evol.* 29, 33–41. doi: 10.1016/j.tree.2013.09.009

Lepage, T., Bryant, D., Philippe, H., and Lartillot, N. (2007). A general comparison of relaxed molecular clock models. *Mol. Biol. Evol.* 24, 2669–2680. doi: 10.1093/molbev/msm193

Lepage, T., Lawi, S., Tupper, P., and Bryant, D. (2006). Continuous and tractable models for the variation of evolutionary rates. *Math. Biosci.* 199, 216–233. doi: 10.1016/j.mbs.2005.11.002

Levy Karin, E., Wicke, S., Pupko, T., and Mayrose, I. (2017). An integrated model of phenotypic trait changes and site-specific sequence evolution. *Syst. Biol.* 66, 917–933. doi: 10.1093/sysbio/syx032

Marjanović, D. (2019). Recalibrating the transcriptomic timetree of jawed vertebrates. *bioRxiv [preprint]*. doi: 10.1101/2019.12.19.882829

Marjanović, D., and Laurin, M. (2007). Fossils, molecules, divergence times, and the origin of lissamphibians. *Syst. Biol.* 56, 369–388. doi: 10.1080/10635150701397635

Marshall, C. R. (2019). Using the fossil record to evaluate timetree timescales. *Front. Genet.* 10:1049. doi: 10.3389/fgene.2019.01049

Matschiner, M. (2019). Selective sampling of species and fossils influences age estimates under the fossilized birth-death model. *Front. Genet.* 10:1064. doi: 10.3389/fgene.2019.01064

Nielsen, R., and Yang, Z. (2003). Estimating the distribution of selection coefficients from phylogenetic data with applications to mitochondrial and viral DNA. *Mol. Biol. Evol.* 20, 1231–1239. doi: 10.1093/molbev/msg147

Parham, J. F., Donoghue, P. C., Bell, C. J., Calway, T. D., Head, J. J., Holroyd, P. A., et al. (2011). Best practices for justifying fossil calibrations. *Syst. Biol.* 61, 346–359. doi: 10.1093/sysbio/syr107

Pascual-García, A., Arenas, M., and Bastolla, U. (2019). The molecular clock in the evolution of protein structures. *Syst. Biol.* 68, 987–1002. doi: 10.1093/sysbio/syz022

Privault, N., and Guindon, S. (2015). Closed form modeling of evolutionary rates by exponential Brownian functionals. *J. Math. Biol.* 71, 1387–1409. doi: 10.1007/s00285-015-0863-6

Pyron, R. A. (2011). Divergence time estimation using fossils as terminal taxa and the origins of Lissamphibia. *Syst. Biol.* 60, 466–481. doi: 10.1093/sysbio/syr047

Rannala, B. (2016). Conceptual issues in Bayesian divergence time estimation. *Philos. Trans. R. Soc. B* 371:20150134. doi: 10.1098/rstb.2015.0134

Revell, L. (2018). *Phytools in R-Phylogenetic Tools for Comparative Biology (and Other Things)*. R package version 3.6.2. Available online at: https://cran.r-project.org/web/packages/phytools/index.html. doi: 10.1111/j.2041-210X.2011.00169.x

Ronquist, F., Klopfstein, S., Vilhelmsen, L., Schulmeister, S., Murray, D. L., and Rasnitsyn, A. P. (2012a). A total-evidence approach to dating with fossils, applied to the early radiation of the Hymenoptera. *Syst. Biol.* 61, 973–999. doi: 10.1093/sysbio/sys058

Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D. L., Darling, A., Höhna, S., et al. (2012b). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61, 539–542. doi: 10.1093/sysbio/sys029

Saladin, B., Leslie, A. B., Wüest, R. O., Litsios, G., Conti, E., Salamin, N., et al. (2017). Fossils matter: improved estimates of divergence times in Pinus reveal older diversification. *BMC Evol. Biol.* 17:95. doi: 10.1186/s12862-017-0941-z

Sarich, V., and Wilson, A. (1967). Immunological time scale for hominid evolution. *Science* 158, 1200–1203. doi: 10.1126/science.158.3805.1200

Shankarappa, R., Margolick, J., Gange, S., Rodrigo, A., Upchurch, D., Farzadegan, H., et al. (1999). Consistent viral evolutionary changes associated with the progression of human immunodeficiency virus type 1 infection. *J. Virol.* 73, 10489–10502. doi: 10.1128/JVI.73.12.10489-10502.1999

Soubrier, J., Steel, M., Lee, M. S., Der Sarkissian, C., Guindon, S., Ho, S. Y., et al. (2012). The influence of rate heterogeneity among sites on the time dependence of molecular rates. *Mol. Biol. Evol.* 29, 3345–3358. doi: 10.1093/molbev/mss140

Stadler, T. (2010). Sampling-through-time in birth-death trees. *J. Theor. Biol.* 267, 396–404. doi: 10.1016/j.jtbi.2010.09.010

Stadler, T. (2012). How can we improve accuracy of macroevolutionary rate estimates? *Syst. Biol.* 62, 321–329. doi: 10.1093/sysbio/sys073

Stadler, T. (2019). *TreeSim in R-Simulating Trees Under the Birth-Death Model*. R package. Available online at: https://cran.r-project.org/web/packages/TreeSim/index.html

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

Stadler, T., Kühnert, D., Bonhoeffer, S., and Drummond, A. J. (2013). Birth-death skyline plot reveals temporal changes of epidemic spread in HIV and hepatitis C virus (HCV). *Proc. Natl. Acad. Sci. U.S.A.* 110, 228–233. doi: 10.1073/pnas.1207965110

Sterli, J., Pol, D., and Laurin, M. (2013). Incorporating phylogenetic uncertainty on phylogeny-based palaeontological dating and the timing of turtle diversification. *Cladistics* 29, 233–246. doi: 10.1111/j.1096-0031.2012.00425.x

Tavaré, S., Marshall, C. R., Will, O., Soligo, C., and Martin, R. D. (2002). Using the fossil record to estimate the age of the last common ancestor of extant primates. *Nature* 416:726. doi: 10.1038/416726a

Thorne, J., Kishino, H., and I, P. (1998). Estimating the rate of evolution of the rate of molecular evolution. *Mol. Biol. Evol.* 15, 1647–1657. doi: 10.1093/oxfordjournals.molbev.a025892

Thorne, J. L., Choi, S. C., Yu, J., Higgs, P. G., and Kishino, H. (2007). Population genetics without intraspecific data. *Mol. Biol. Evol.* 24, 1667–1677. doi: 10.1093/molbev/msm085

Tuffley, C., and Steel, M. (1998). Modeling the covarion hypothesis of nucleotide substitution. *Math. Biosci.* 147, 63–91. doi: 10.1016/S0025-5564(97)00081-3

Warnock, R. C., Parham, J. F., Joyce, W. G., Lyson, T. R., and Donoghue, P. C. (2015). Calibration uncertainty in molecular dating analyses: there is no substitute for the prior evaluation of time priors. *Proc. R. Soc. B Biol. Sci.* 282:20141013. doi: 10.1098/rspb.2014.1013

Yang, Z. (1994). Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *J. Mol. Evol.* 39, 306–314. doi: 10.1007/BF00160154

Yang, Z., and Rannala, B. (2005). Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Mol. Biol. Evol.* 23, 212–226. doi: 10.1093/molbev/msj024

Zuckerkandl, E., and Pauling, L. (1965). Molecules as documents of evolutionary history. *J. Theor. Biol.* 8, 357–366. doi: 10.1016/0022-5193(65)90083-4

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Guindon. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 26 June 2020
doi: 10.3389/fevo.2020.00183

Ignoring Fossil Age Uncertainty Leads to Inaccurate Topology and Divergence Time Estimates in Time Calibrated Tree Inference

Joëlle Barido-Sottani[1,2,3†], Nina M. A. van Tiel[1,2†], Melanie J. Hopkins[4], David F. Wright[4,5], Tanja Stadler[1,2] and Rachel C. M. Warnock[6]*

[1] Department of Biosystems Science and Engineering, ETH Zürich, Basel, Switzerland, [2] Swiss Institute of Bioinformatics (SIB), Lausanne, Switzerland, [3] Department of Ecology, Evolution and Organismal Biology, Iowa State University, Ames, IA, United States, [4] Division of Paleontology, American Museum of Natural History, New York, NY, United States, [5] Department of Paleobiology, National Museum of Natural History, Smithsonian Institution, Washington, DC, United States, [6] GeoZentrum Nordbayern, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

OPEN ACCESS

Edited by:
Jeffrey Peter Townsend, Yale University, United States

Reviewed by:
Guillaume Guinot, UMR5554 Institut des Sciences de l'Evolution de Montpellier (ISEM), France
Nicolas Mongiardino Koch, Yale University, United States

***Correspondence:**
Rachel C. M. Warnock
rachel.warnock@fau.de

[†] *These authors have contributed equally to this work*

Specialty section:
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Ecology and Evolution

Received: *14 January 2020*
Accepted: *25 May 2020*
Published: *26 June 2020*

Citation:
Barido-Sottani J, van Tiel NMA, Hopkins MJ, Wright DF, Stadler T and Warnock RCM (2020) Ignoring Fossil Age Uncertainty Leads to Inaccurate Topology and Divergence Time Estimates in Time Calibrated Tree Inference. Front. Ecol. Evol. 8:183. doi: 10.3389/fevo.2020.00183

Time calibrated trees are challenging to estimate for many extinct groups of species due to the incompleteness of the rock and fossil records. Additionally, the precise age of a sample is typically not known as it may have occurred at any time during the time interval spanned by the rock layer. Bayesian phylogenetic approaches provide a coherent framework for incorporating multiple sources of evidence and uncertainty. In this study, we simulate datasets with characteristics typical of Palaeozoic marine invertebrates, in terms of character and taxon sampling. We use these datasets to examine the impact of different age handling methods on estimated topologies and divergence times obtained using the fossilized birth-death process. Our results reiterate the importance of modeling fossil age uncertainty, although we find that the relative impact of fossil age uncertainty depends on both fossil taxon sampling and character sampling. Sampling the fossil ages as part of the inference gives topology and divergence time estimates that are as good as those obtained by fixing ages to the truth, whereas fixing fossil ages to incorrect values results in higher error and lower coverage. The relative effect increases with increased fossil and character sampling. Modeling fossil age uncertainty is thus critical, as fixing incorrect fossil ages will negate the benefits of improved fossil and character sampling.

Keywords: time calibrated phylogeny, divergence time estimates, Bayesian phylogenetic analysis, fossil age uncertainty, fossilized birth death model

1. INTRODUCTION

Estimating phylogenetic relationships and divergence times among species are key components of piecing together evolutionary and geological history. Approaches to building time trees in paleobiology have traditionally involved estimating the topology and branch lengths scaled to time in separate, sequential analyses (Bapst and Hopkins, 2017). Bayesian phylogenetic models make it possible to estimate these parameters in combination. An advantage of this joint inference is that temporal evidence can be used to inform the tree topology, in combination with character data, and the posterior output will better reflect the uncertainty associated with the results (Ronquist et al., 2012).

Statistically coherent models for incorporating extinct species into time calibrated tree inference only recently became available. In particular, the fossilized birth-death (FBD) process provides a joint description of the diversification and fossil sampling processes (Stadler, 2010; Heath et al., 2014). Under this model, extinct dated samples are considered as part of the tree, therefore contributing temporal information, and their phylogenetic position can be recovered, either as terminal branches (tips) or ancestral to other samples (sampled ancestors). This modeling framework has created enormous potential for incorporating more paleontological data into divergence time analyses and we are only just beginning to explore the impact and existing limitations of this approach.

Analyses using the FBD process can be divided into two categories depending on the amount of data available. The first category uses topological constraints which assign fossils to specific clades (Gavryushkina et al., 2014; Heath et al., 2014). In these analyses the position of the fossils in the tree is thus not part of the inference. The second category are so-called "total-evidence" approaches, which use morphological data to place the fossils on the tree as part of the inference (Ronquist et al., 2012; Zhang et al., 2015; Gavryushkina et al., 2017). Total-evidence analyses better reflect the uncertainty associated with fossil placement than analyses that fix the position of fossils and thus may lead to more accurate results, particularly in clades where the fossil taxonomy is contested. This approach can also be applied to entirely extinct groups, for which only morphological and no molecular data are available (Lee et al., 2014; Slater, 2015; Wright, 2017b; Wright and Toom, 2017; Paterson et al., 2019).

Simulations play an important role in testing the limits of tree inference methods. Different taxonomic groups and time periods are associated with different issues that contribute to challenges inferring topology and time, and a growing number of studies have sought to explore the performance of phylogenetic inference under the FBD model in different scenarios. Several studies have focused on specific model violations, including the impact of non-uniform sampling among living taxa (Zhang et al., 2015; Matschiner, 2019), non-uniform sampling of fossil taxa over time (Heath et al., 2014; Zhang et al., 2015; O'Reilly and Donoghue, 2019) or across lineages (Heath et al., 2014; Matschiner et al., 2017), as well as the effect of ignoring sampled ancestors (Gavryushkina et al., 2014). A clear consensus that emerges from this work is that higher sampling rates of taxa and characters result in better estimates of time and (when co-estimated) topology, provided model violation is not extreme. In an extensive set of simulations, Luo et al. (2019) examined the performance of total-evidence inference under the FBD model. This work indicated that a large degree of uncertainty is anticipated to be associated with the placement of extinct samples, for which only morphology is available, and that fossil sampling may ultimately outweigh the significance of other issues encountered in dating analyses, including character sampling and among-lineage rate variation.

Barido-Sottani et al. (2019a) focused on one particular aspect of the fossil record, namely the uncertainty associated with the age assigned to each fossil sample. As the age of fossils is established in reference to the geological record, fossil samples are not dated to a single numerical value but rather to an interval of time; this is referred to hereafter as the "age range" of the sample. This uncertainty can be handled in FBD analyses by sampling fossil ages as part of the inference (Drummond and Stadler, 2016), but many studies in the existing literature chose instead to fix fossil ages to a single value, usually the midpoint of the age range (e.g., Larabee et al., 2016) or an age sampled uniformly at random inside the range (e.g., Grimm et al., 2015). Barido-Sottani et al. (2019a) tested these different approaches of handling fossil age uncertainty in analyses using topological constraints to place fossils and found that fixing the fossil ages to incorrect values led to important errors in divergence times estimates. Here, we extend the Barido-Sottani et al. (2019a) study to time calibrated tree inference using morphological data only. Using simulated datasets, we explore the impact of character and taxon sampling, approaches to handling fossil age uncertainty, and clock model priors on estimates of topology and divergence times. We also compare our results to those obtained using temporally unconstrained (i.e., non-time calibrated) Bayesian tree inference. Finally, we apply the FBD model and several different methods for handing fossil age uncertainty to an empirical dataset that is typical of those available for Paleozoic marine invertebrates.

2. METHODS

2.1. Simulated Datasets

The design of our simulation study is broadly based on features that are typical for datasets of Paleozoic marine invertebrates. In order to select parameter values that would reflect the size and scale of these datasets (in terms of taxon and character sampling) we first tallied 81 studies of Paleozoic invertebrate groups, which included trilobites (67%), brachiopods (18%), and crinoids (15%). The majority of these studies used species as operational taxonomic units (OTUs) (77%), while the rest coded genera (23%). The results are summarized in **Figure S1** and **Appendix Table 1**. The typical size of these datasets was 25–35 taxa, at both taxonomic levels (mean for species OTUs = 25, mean for genus OTUs = 32), with a maximum of 85. For 63 of these studies we were able to estimate approximate time spans in millions of years (Myr). Across all studies, the typical time span was 50 Myr, with 85% <75 Myr. However, there was a large difference between taxonomic scales: the mean total time span for studies using species OTUs was 37 Myr, while the mean total time span for studies using genus OTU studies was 88 Myr. Interestingly, no relationship was observed between the number of taxa in the study and the total time span. Intuitively, we might expect sampling taxa over longer intervals to lead to datasets containing larger numbers of taxa, i.e., because the number of opportunities for sampling increases. However, it is not clear whether this observation reflects a genuine lack of correlation between time and the number of taxa sampled, or the fact that studies chose not to include all available taxa in phylogenetic studies for practical reasons, e.g., due to the intense effort required to collect morphological characters. The average number of characters was 35 for both species and genus levels, with an average of 60% binary characters.

TABLE 1 | Parameters values for low and high fossil sampling settings.

Parameter	Low sampling	High sampling
Fossilization rate ψ	0.03/Myr	0.1/Myr
Minimum number of fossils n_{min}	20	80
Maximum number of fossils n_{max}	40	120
Minimum span t_{span}	30 Myr	50 Myr

Based on these empirical data characteristics, we established two main parameter settings that determined the number of fossils sampled during simulation, one based on the average size of empirical datasets (referred to as *low sampling*) and the other based on a more optimistic sampling scenario (referred to as *high sampling*). We also explored the effect of morphological matrix length (30, 300, or 3,000 characters), where the lowest value was based on our sample of empirical studies and the higher values represented more optimistic scenarios. The optimistic scenarios are more similar to previous simulation studies that have focused on morphology based tree inference (Wright and Hillis, 2014; O'Reilly and Donoghue, 2017; Puttick et al., 2017). Note that *a priori* we do not expect to recover good results under the low sampling scenario and with the small number of characters typical of empirical datasets. The more optimistic scenarios were necessarily included to gain robust insights into the behavior of our inference framework. For each set of parameter values we simulated 50 replicates.

2.1.1. Simulation of Phylogenies and Fossil Samples

Trees were simulated under a constant rate birth-death process with speciation rate $\lambda = 0.06$/Myr and extinction rate $\mu = 0.045$/Myr, using the R package `TreeSim` (Stadler, 2011). These estimates were taken from an empirical study of Paleozoic crinoids (Wright, 2017a), which is the only one of the 81 datasets evaluated that has been the subject of an analysis using the FBD model, and for which empirical estimates of these parameters were readily available. The birth-death simulation was allowed to run for 130 Myr, which approximates the temporal duration (Ordovician to Devonian) of the crinoid clade of Wright (2017a).

Fossils were sampled on the complete phylogeny following a Poisson process with a constant fossilization rate ψ, using the R package `FossilSim` (Barido-Sottani et al., 2019b). We rejected phylogenies with less than the minimum number n_{min} or more than the maximum number n_{max} of sampled fossils, and phylogenies for which the fossils spanned less than the minimum time span t_{span} Myr. Values for ψ, n_{min}, n_{max}, and t_{span} depended on the fossil sampling setting (high vs. low sampling) and are detailed in **Table 1**. The average time span of simulated datasets was 111 Myr in the low sampling scenario, and 123 Myr in the high sampling scenario, consistent with our assumption of constant fossilization rate over the entire simulated period.

2.1.2. Simulation of Fossil Age Uncertainty

Fossil age uncertainty was simulated using the procedure described in Barido-Sottani et al. (2019a). Realistic age ranges for simulated data are based on empirical ranges of fossil crinoids obtained from the Paleobiology Database (PBDB) using the following parameters: time intervals = from Ordovician to Devonian, scientific name = Crinoidea (download date: 07/03/2018). Each simulated fossil sample was assigned to an interval based on its true age. If a simulated fossil age could be assigned to multiple intervals, a single interval was selected at random by weighting all possible intervals by their frequency of appearance in the PBDB data. If no intervals appeared in the PBDB data for a simulated fossil age, a random interval containing the true age was drawn, with a length equal to the average length of all intervals in the PBDB data, i.e., 12 Myr. Thus, the simulated interval for each fossil always included the correct age of the fossil.

2.1.3. Simulation of Morphological Data

As the majority of the characters used in our sample of empirical studies were binary and the number of character states was not the focus of our study, we chose to simulate binary characters only. These characters were simulated for each fossil using the function `sim.char` from the R package `geiger` (Pennell et al., 2014). A strict clock model was used and the rate of character state change was set to 0.033/Myr, based on the rate obtained by Wright (2017a). For both fossil sampling settings, character matrices of length 30, 300, and 3,000 were simulated. We did not filter the resulting matrices to remove uninformative characters. However, the proportion of uninformative characters was low: 0% for the matrices with 30 characters, 0.07% for the matrices with 300 characters, and 0.04% for the matrices with 3,000 characters.

2.1.4. Bayesian Inference

Markov Chain Monte-Carlo (MCMC) inference using the FBD process is implemented in the Sampled Ancestors package (Gavryushkina et al., 2014) for the software BEAST2 (Bouckaert et al., 2014). We extended this package to be able to use a tree with no extant samples. This extension made no changes to the FBD model or to the likelihood function, and was done simply to allow for sampling fossil ages on a fully extinct tree. This package was used to perform Bayesian phylogenetic inference on the simulated datasets. The fossil ages were handled using five different methods, detailed here and illustrated in **Figure 1**.

- Correct ages: the fossil ages are fixed to the true ages as simulated.
- Interval ages: the fossil ages are not fixed, but are sampled along with the other parameters within the simulated age range.
- Median ages: the fossil ages are fixed to the midpoint of their simulated age range.
- Random ages: the fossil ages are fixed to an age sampled uniformly at random inside of their simulated age range.
- Symmetric interval ages: the fossil ages are not fixed, but are sampled along with the other parameters. Each fossil age is sampled within a symmetric interval of length 12 Myr (i.e., equal to the average length of all intervals in the PBDB data) around the true age of the fossil. The purpose of this setting

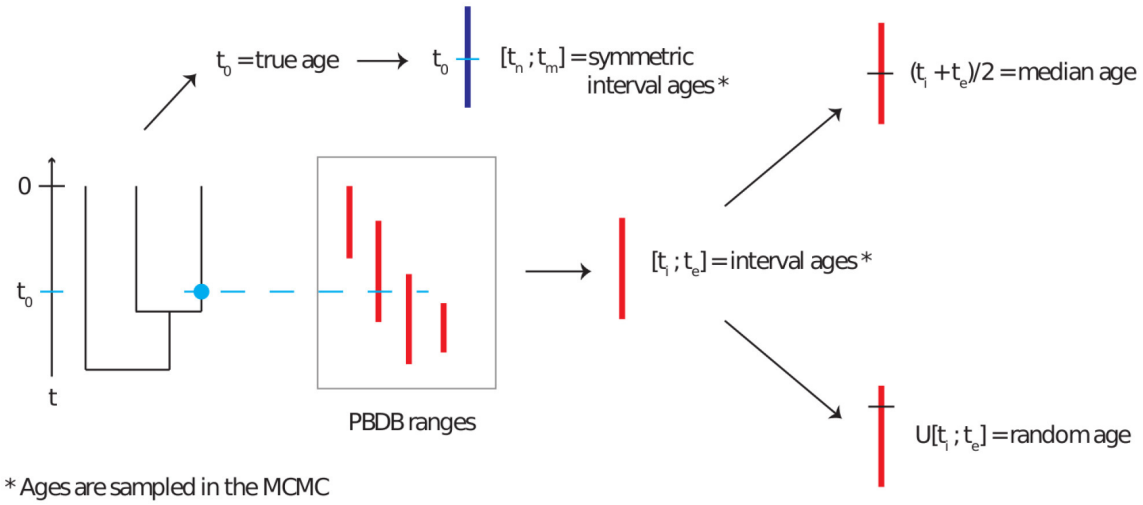


FIGURE 1 | Representation of the age uncertainty simulation process, reproduced from Barido-Sottani et al. (2019a). Phylogenies with fossils are simulated according to a birth-death-fossilization process. The correct age of each fossil is used to draw an age interval for that fossil from the set obtained from PBDB. This age interval is then used as the basis for the median and random age assignment. A symmetric age interval is also drawn from the correct age.

was to evaluate whether the position of the interval relative to the true age affected the resulting estimates.

Note that for the interval age methods, we sample trees as in Drummond and Stadler (2016), i.e., we set the probability density of the proposed tree to the FBD probability density if all fossil ages are within their intervals, and 0 otherwise. The effective prior on fossil ages, i.e., the fossil age distribution when using all information excluding sequence data, is thus not a uniform prior, as the FBD model already induces a distribution on fossil ages.

The Lewis Mk model of morphological character evolution was used (Lewis, 2001). The strict clock model was used with three different priors on the clock rate: an unbounded uniform prior, a lognormal prior with median = 0.033/Myr, equal to the true rate [i.e., LogNormal(−3.4, 0.3)] and a lognormal prior with median = 1.220/Myr, different from the true rate [i.e., LogNormal(0.2, 1.25)]. The inference was run for at least 100,000,000 iterations, or until convergence was considered satisfactory, and sampled every 10,000 steps. Convergence was assessed in the software Tracer v. 1.7 (Rambaut et al., 2018) and considered satisfactory if the effective sample sizes were more than 200. Datasets that did not converge after 120 h were excluded (<5% of runs).

Additionally, unconstrained Bayesian phylogenetic inferences were performed on the simulated datasets using the RevBayes framework (Höhna et al., 2016) without including any fossil age information. These inferences were performed on all simulated datasets for both low and high fossil sampling and, with character data simulated under the strict clock model. We used the Lewis Mk model, a uniform prior on the tree topology and an exponential prior on the branch lengths. The mean of the distribution on the branch lengths, which is determined by the rate parameter λ, was estimated using an exponential hyperprior with mean = 1. The use of alternative branch length priors did not impact estimates of tree accuracy. See **Supplementary Material Section 3** and **Figure S4** for more details. Convergence and MCMC diagnostics were assessed using identical guidelines as those described above.

2.1.5. Assessing Inference Results

We assessed the accuracy of the FBD model parameters by measuring the relative error of the median posterior estimates, where the relative error was defined as the difference between the true value and the estimated value, divided by the true value. The relative error was averaged over all replicates. We also calculated the coverage, i.e., the proportion of analyses in which the true parameter value was included in the 95% highest posterior density (HPD) interval. In order to evaluate the accuracy of the divergence time estimates, we considered nodes defined as the most recent common ancestor (MRCA) of samples t_1 and t_2, for all pairs of samples. This definition allowed us to obtain nodes which were always present in the inferred tree, regardless of the accuracy of the topology. Similarly to the FBD model parameters, we calculated the relative error of the median posterior estimates and the coverage of the divergence times, averaged across all nodes and replicates.

To assess the accuracy of inferred topologies we calculated the mean normalized Robinson-Foulds (RF) distance (Robinson and Foulds, 1981) between simulated trees and tree samples from the posterior distribution. The RF distance only depends on the topology of the trees. The normalized RF distance between two trees with n tips is computed by dividing the RF distance between these trees by the maximum possible RF distance between two trees with n tips, thus scaling the distances between 0 and 1.

The normalized RF distances were calculated using the `RF.dist` function from the R package `phangorn` (Schliep, 2010), and averaged over all the trees sampled during the MCMC

(ignoring the first 10% of the samples as burn-in), and averaged over all replicates for each parameter combination.

All trees were unrooted prior to calculating the RF distance to facilitate comparison between the time constrained and unconstrained analyses.

2.2. Empirical Dataset

To explore the impact of different approaches to handling stratigraphic age uncertainty on empirical estimates of divergence times, Bayesian phylogenetic inference was performed on a dataset of North American Devonian brachiopod species (Stigall Rode, 2005). This dataset was chosen because, with 18 taxa and 36 characters, it represents the average size of the 81 studies we evaluated (**Figure S1**). Among datasets of similar size, it also comprised OTUs sampled across geologic stages. The latter criterion is important because at global scales, the geologic time scale is generally coarse enough that closely related species occur within the same geologic stage, and obtaining a finer resolution time scale is not straightforward, or even possible, in many instances (Hopkins et al., 2018).

Fossil occurrences were assigned to geologic stages based on vetted occurrences in the Paleobiology Database and additional literature (Stigall Rode, 2005; Menning et al., 2006). Minimum and maximum ages for stage boundaries were assigned following the International Commission on Stratigraphy 2018 chart (www.stratigraphy.org). Species for which all specimens were recovered from the same geological stage were treated as a single OTU. Three species had specimens which were sampled from more than one stage and were treated as multiple OTUs, corresponding to one OTU for each stage. For these species, the species was constrained to be monophyletic and morphology was included for the oldest specimen only. This approach was taken in order to avoid having multiple specimens associated with the same morphology over long intervals of time, which would represent a strong violation of the Mk model. The analysis used the same model parameterization and priors as the simulated data. The clock rate prior was set to a lognormal distribution [LogNormal(−3.4, 0.3)]. As the true ages of the fossils in this dataset are unknown, we limited our comparison to the median, random, and interval ages in BEAST2, and the unconstrained analysis in RevBayes. To facilitate comparison between constrained and unconstrained topologies, unconstrained trees were rooted using *Xystostrophia umbraculum* as the outgroup taxon (Stigall Rode, 2005). The FBD analyses were run both excluding and including the outgroup. This choice did not impact the overall results but we note that the inclusion of an outgroup taxon represents a violation of the assumption of uniform taxon sampling throughout the tree.

3. RESULTS

3.1. Simulated Datasets

3.1.1. Impact of the Clock Rate Prior

The use of different clock rate priors had a minor impact on the results. For most parameters, including divergence times, estimates of coverage and relative error were similar or even identical for different clock priors, particularly when character and fossil sampling were both high (**Figures S2, S3**). The use of different priors on the clock rate also had little impact on the topological accuracy based on RF distances (**Figure 2**). The largest impact was observed on the clock rate itself. When character and fossil sampling were low, there was less signal in the data to inform this parameter. Running the analysis with an unbounded uniform prior (i.e., from 0 to ∞) on the clock rate in the low sampling scenario produced rate estimates of $\approx 10^{307}$ (i.e., the numerical upper limit of the software) in approximately 20% of replicates, as the posterior followed the prior. Thus, we excluded this condition from **Figure S2** (lower panel).

As the clock rate prior exerted a negligible impact on parameter estimates, for the remainder of the results we focus on describing the output obtained using a lognormal prior on the clock rate with a median that differs from the true rate. This setting best matches a plausible scenario for empirical studies and avoids the issues encountered using the unbounded prior.

3.1.2. The Combined Effects of Stratigraphic Age Uncertainty, Fossil Sampling, and Character Sampling

Figures 3–5 present the results obtained under different character and fossil sampling settings when running the analysis using the lognormal clock rate prior with a median that differs from the true rate. Using symmetric interval ages results in very similar estimates compared to interval ages, showing that the accuracy of the estimates is not affected by the position of the interval relative to the true age of the fossil. Thus, in the following we will refer to these two conditions together.

The accuracy of inferred divergence times, in terms of coverage and relative error, show similar behavior across fossil and character sampling settings (**Figure 3**). In particular, we obtained high accuracy (i.e., high coverage and low relative error) when the fossil ages were fixed to the correct ages or sampled from within the known interval of uncertainty as part of the MCMC, irrespective of fossil or character sampling. In contrast, we obtained low accuracy when the ages were fixed to incorrect (median or random) ages, but the extent to which the results were worse depended on both fossil and character sampling. In the case of fixed incorrect ages, increased fossil, and character sampling decreased the accuracy of divergence time estimates. A similar trend is observed for the diversification and turnover parameters (**Figure 4**). The clock rate parameter showed the same trends for coverage (i.e., higher fossil and character sampling lead to lower coverage with median or random fossil ages), but a different trend was recovered for relative error (**Figure 3**). Specifically, when fossil and character sampling were low, relative error was higher when fossil ages were co-estimated compared to when the ages were fixed to either the correct or incorrect ages. However, coverage was consistently lower with incorrect fossil ages.

The accuracy of inferred trees follow a pattern which is similar overall to the divergence times estimates, across fossil age handling approaches and fossil sampling settings. However, character sampling had a large impact on the magnitude of the differences observed under different age handling and fossil sampling scenarios (**Figure 5**). In particular, when character

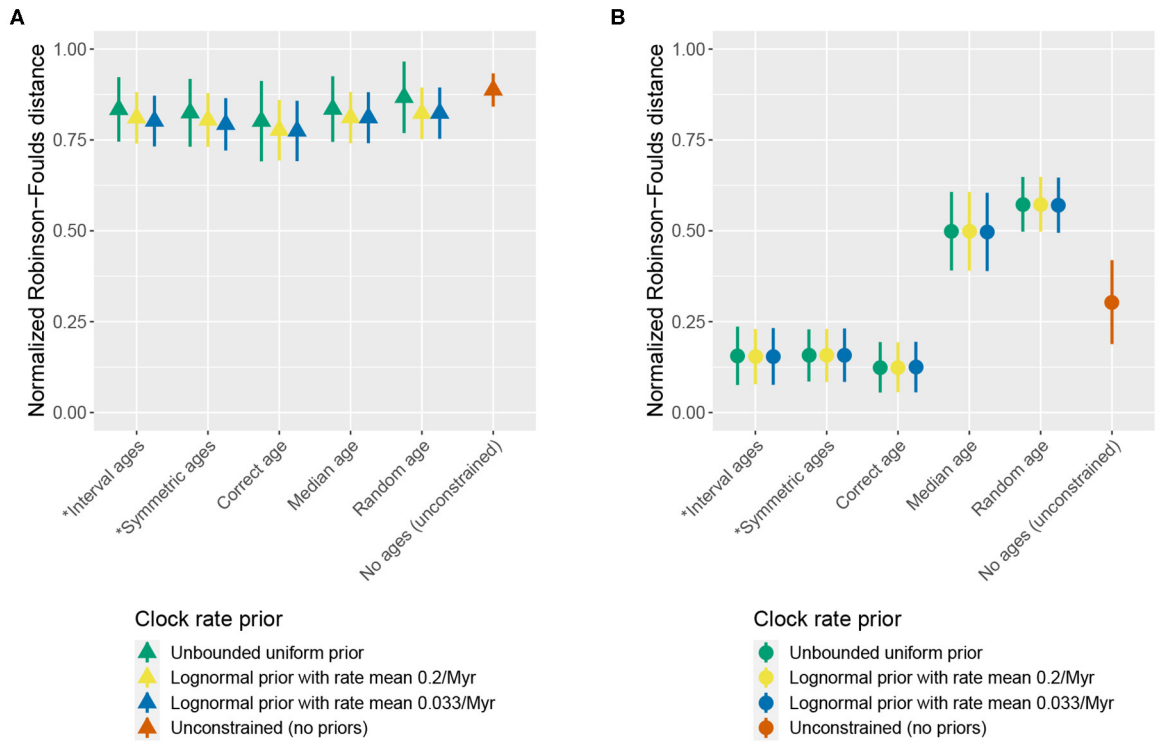


FIGURE 2 | Impact of the clock prior on topology. The mean and standard deviation of the normalized Robinson-Foulds distance between the true simulated tree and trees inferred during MCMC across all replicates are shown for different clock rate priors, different age handling methods, and different fossil sampling settings [low **(A)** vs. high **(B)**]. Ages sampled as part of the MCMC are marked by (*). **(A)** Low fossil sampling and character length 30. **(B)** High fossil sampling and character length 300.

sampling was low ($n = 30$) the inferred trees were relatively far from the true tree, as measured by RF distance, irrespective of fossil age handling approach or fossil sampling parameters. Overall, higher character and fossil sampling both led to increased accuracy (i.e., lower RF distances) across all scenarios, with the best estimates obtained when both character and fossil sampling were high (**Figure 5**). The positive effects of increased fossil or character sampling were also greater when fossil ages were fixed to the truth or co-estimated, while estimates obtained when fossil ages were fixed to median or random ages remained inaccurate even with high sampling. When fossil and character sampling were both high, using the correct fossil ages or estimating the ages performed much better than using incorrect fossil ages.

Differences in accuracy between time calibrated and unconstrained tree inferences were also linked to variation in character sampling (**Figure 2**). For low or intermediate character sampling ($n = 30$ or 300) combined with low fossil sampling, or for low character sampling ($n = 30$) combined with high fossil sampling, the FBD inference outperformed the unconstrained inference, irrespective of the fossil age handling method. In contrast, for increased character or fossil sampling ($n = 3,000$ combined with low fossil sampling and $n = 300$ or 3,000 combined with high fossil sampling), the unconstrained inference outperformed the FBD model when fossil ages were fixed to incorrect ages. The FBD model outperformed the unconstrained inference under intermediate sampling scenarios ($n = 3,000$ combined with low fossil sampling and $n = 300$ combined with high fossil sampling) when fossil ages were fixed to the correct ages or co-estimated. When fossil and character sampling were both high the results obtained using both constrained and unconstrained analyses converged on the true tree, provided fossil ages were fixed to correct ages or co-estimated.

The results are summarized in **Table 2**. Overall, these results indicate that increasing the amount of data does not compensate for the errors introduced by fixing fossil ages to incorrect values. On the contrary, these errors have a much larger impact when using larger datasets, to the point that discarding the fossil ages entirely leads to better estimates of the topology than using incorrect fixed ages.

3.2. Empirical Dataset

The Maximum Clade Credibility (MCC) trees obtained with interval ages, median ages, random ages or unconstrained analysis using our empirical brachiopod dataset are shown in **Figure 6**. The parameter estimates obtained under different age handling methods are shown in **Figure S5**. All OTUs belonging to the same species were constrained to be monophyletic. However, the posterior support for these nodes may be lower than 1.0. This is due to the fact that the clade [A1(A2)], where A1 is a sampled ancestor of A2, represents a different

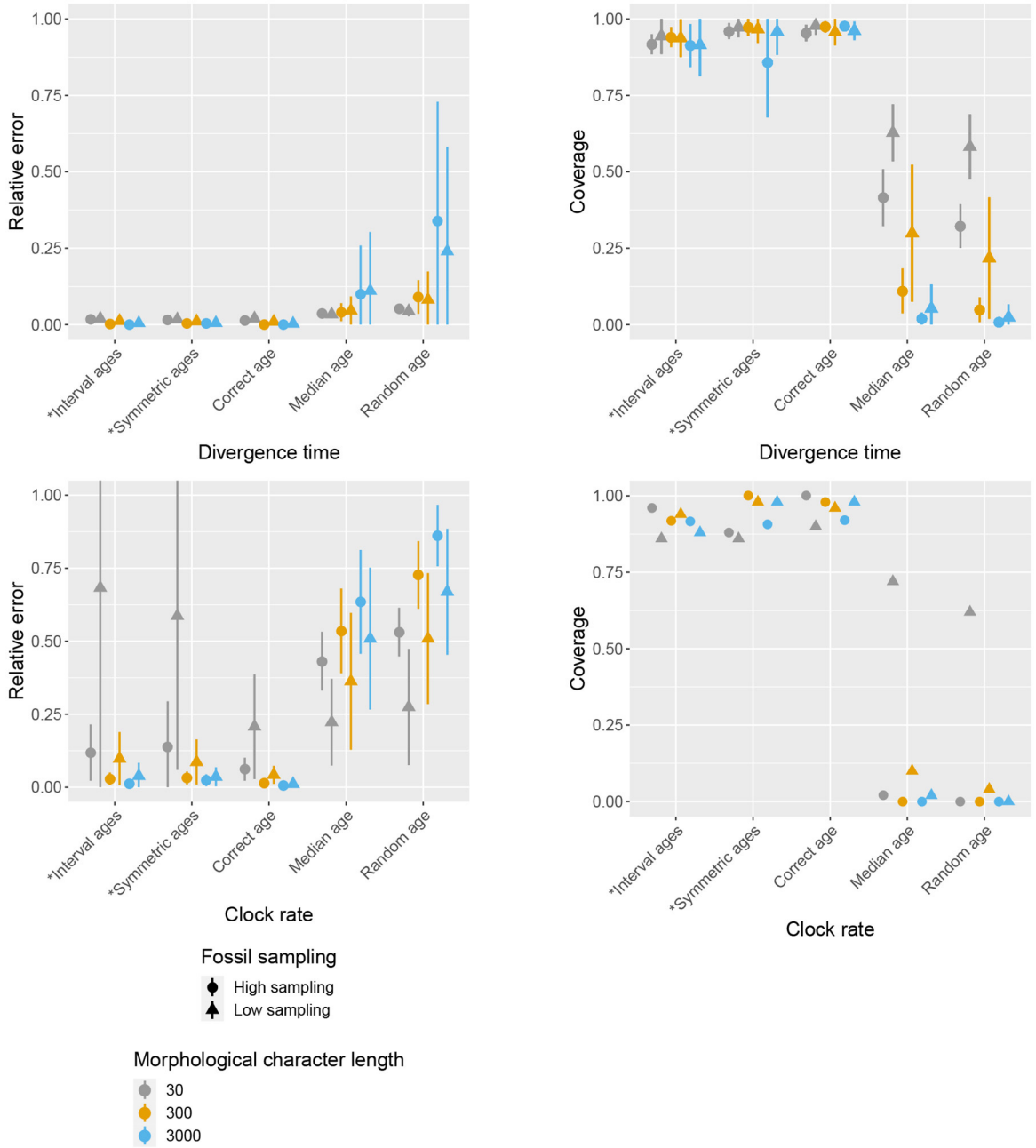


FIGURE 3 | Impact of character and fossil sampling on divergence times and clock rate. The mean and standard deviation of the relative error of median posterior estimates **(left)**, and the mean 95% HPD coverage **(right)** are shown for different age handling methods, different character sampling, and different fossil sampling settings. Ages sampled as part of the MCMC are marked by (*).

realization of the FBD process than the clade (A1,A2), and so they are counted separately when calculating the posterior support of nodes using an MCC tree summary method such as TreeAnnotator.

The MCC trees obtained with the three methods for handling fossil ages are all almost identical in terms of their topology, with the exception of the placement of *Floweria arctostriata* in the random ages tree, and the node support is consistent across all three analyses. The median root ages are slightly different, with the median root age for the interval ages analysis the youngest, but only by a few million years (**Figure 6**, **Figure S2**). The MCC for the unconstrained analysis supports some of the same sister taxa with similar support values, and larger subclades are broadly consistent with several exceptions for specific taxa. Particularly notable is the derived placement of *Floweria becraftensis* in the unconstrained analysis. This

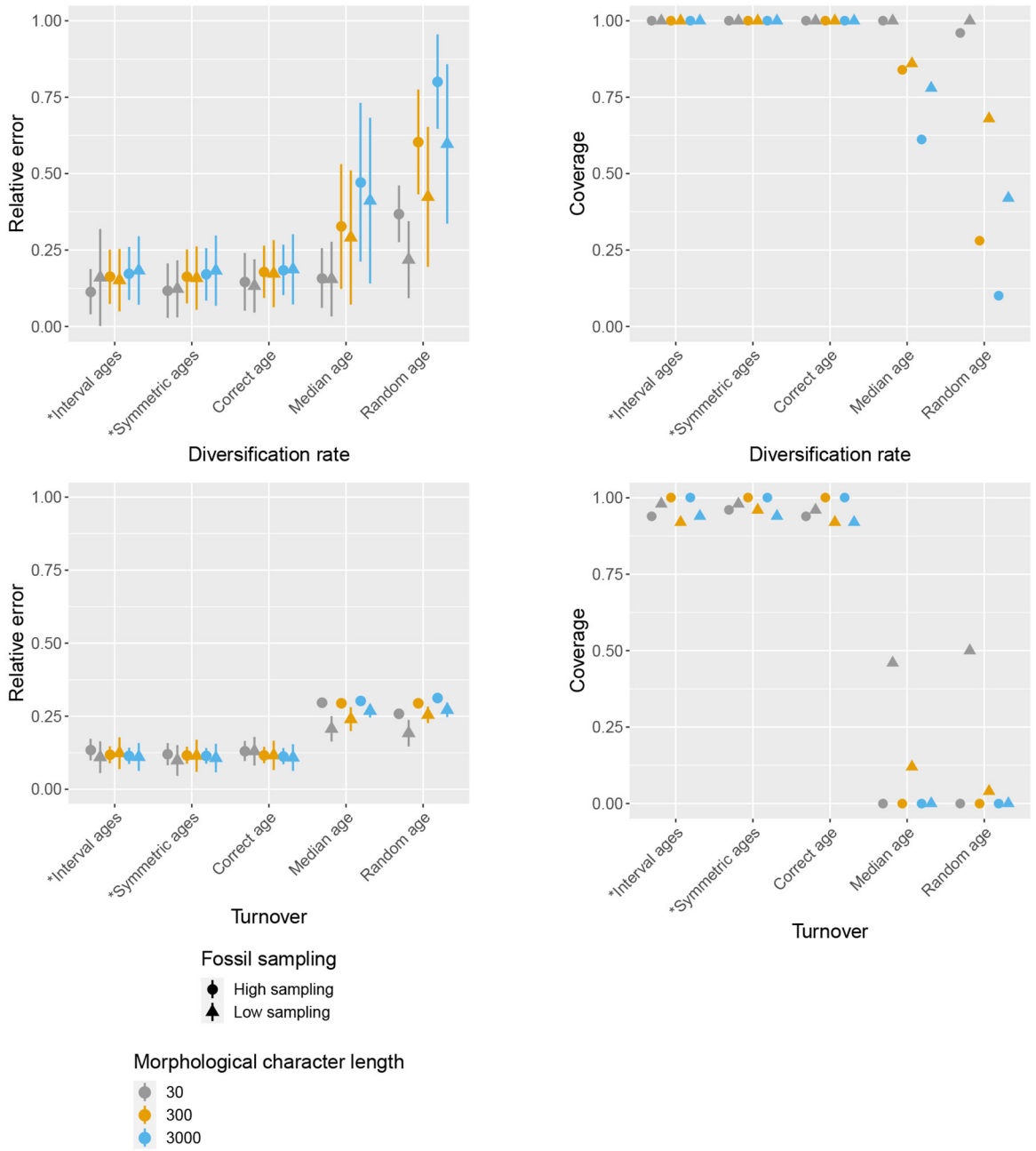


FIGURE 4 | Impact of character and fossil sampling on diversification and turnover. The mean and standard deviation of the relative error of median posterior estimates **(left)**, and the mean 95% HPD coverage **(right)** are shown for different age handling methods, different character sampling and different fossil sampling settings. Ages sampled as part of the MCMC are marked by (*).

species is among the oldest of the clade, and when fossil ages are included in the analysis, it is commonly placed as sister to the rest. Including the outgroup in the FBD analyses did not impact the estimated ingroup topology or divergence times (**Figure S6**). Overall, these results match the output expected based on our simulations, given the low taxon and character sampling, and fossil age uncertainty associated with this dataset.

4. DISCUSSION

The FBD model can be used to estimate time-calibrated trees under a range of scenarios. Our goal was to examine the impact of stratigraphic age uncertainty in FBD model analyses for datasets that are characteristic of fully extinct clades, such as Paleozoic marine invertebrate groups. Our survey of empirical data confirms that datasets associated with taxonomic groups

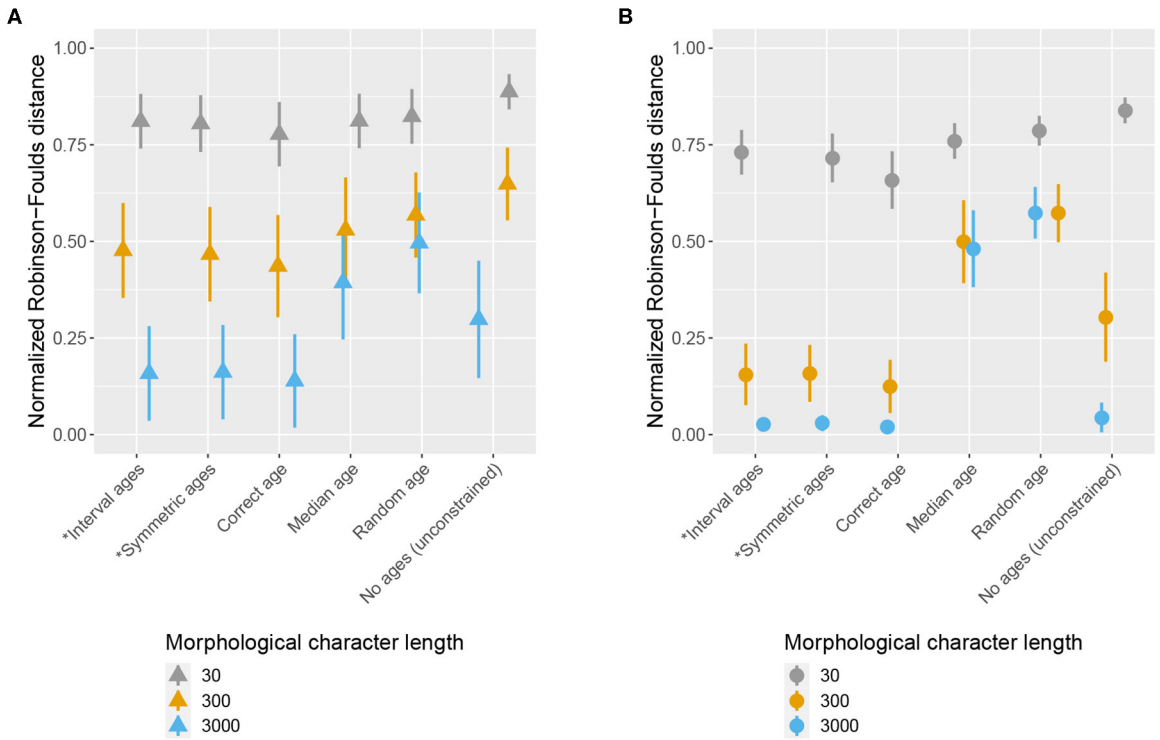


FIGURE 5 | Impact of character and fossil sampling on topology. The mean and standard deviation of the normalized Robinson-Foulds distance between the true simulated tree and trees inferred during MCMC across all replicates are shown for shown for different age handling methods, different character sampling, and different fossil sampling settings [low **(A)** vs. high **(B)**]. Ages sampled as part of the MCMC are marked by (*). **(A)** Low fossil sampling, **(B)** High fossil sampling.

TABLE 2 | Impact of fossil and character sampling on the estimates obtained using the FBD model with different age handling methods vs. unconstrained (i.e., non-clock) inference.

Character \ Fossil	Low sampling	High sampling
Low sampling	- No effect of age handling method on parameter and age estimates - FBD outperforms unconstrained inference on topological accuracy	- Higher error and/or lower coverage on parameter and age estimates with incorrect ages compared to estimated ages - FBD outperforms unconstrained inference on topological accuracy
High sampling	- Higher error and/or lower coverage on parameter and age estimates with incorrect ages compared to estimated ages - FBD with estimated ages outperforms unconstrained inference on topological accuracy - Unconstrained inference outperforms FBD with incorrect ages on topological accuracy	- Much higher error and/or lower coverage on parameter and age estimates with incorrect ages compared to estimated ages - FBD with estimated ages outperforms unconstrained inference on topological accuracy - Unconstrained inference outperforms FBD with incorrect ages on topological accuracy

from this time period typically have a small number of both taxa and phylogenetic characters. The age uncertainty associated with fossil samples from this time period is also relatively high (12 Myr on average, compared with a typical time span of 50 Myr for the full dataset in our example studies). Our results demonstrate the importance of incorporating stratigraphic age uncertainty into phylogenetic dating analyses on these datasets, rather than the popular practice of fixing fossil ages to a value from within the known interval of uncertainty, e.g., using the mean or a random value (**Figures 3**, **4**, **Figures S2, S3**).

Our results build on the findings of previous work, where it was shown that fixing fossil ages to incorrect values can lead to inaccurate estimates of divergence times under the FBD model when using topological constraints to place the fossils (Barido-Sottani et al., 2019a). This previous study focused on a scenario where the aim was to estimate divergence times among extant species using molecular data. No character data was available for fossil samples but it was assumed that strong prior information was available to constrain the topology. Here, we assumed that the phylogenetic position of fossil samples was

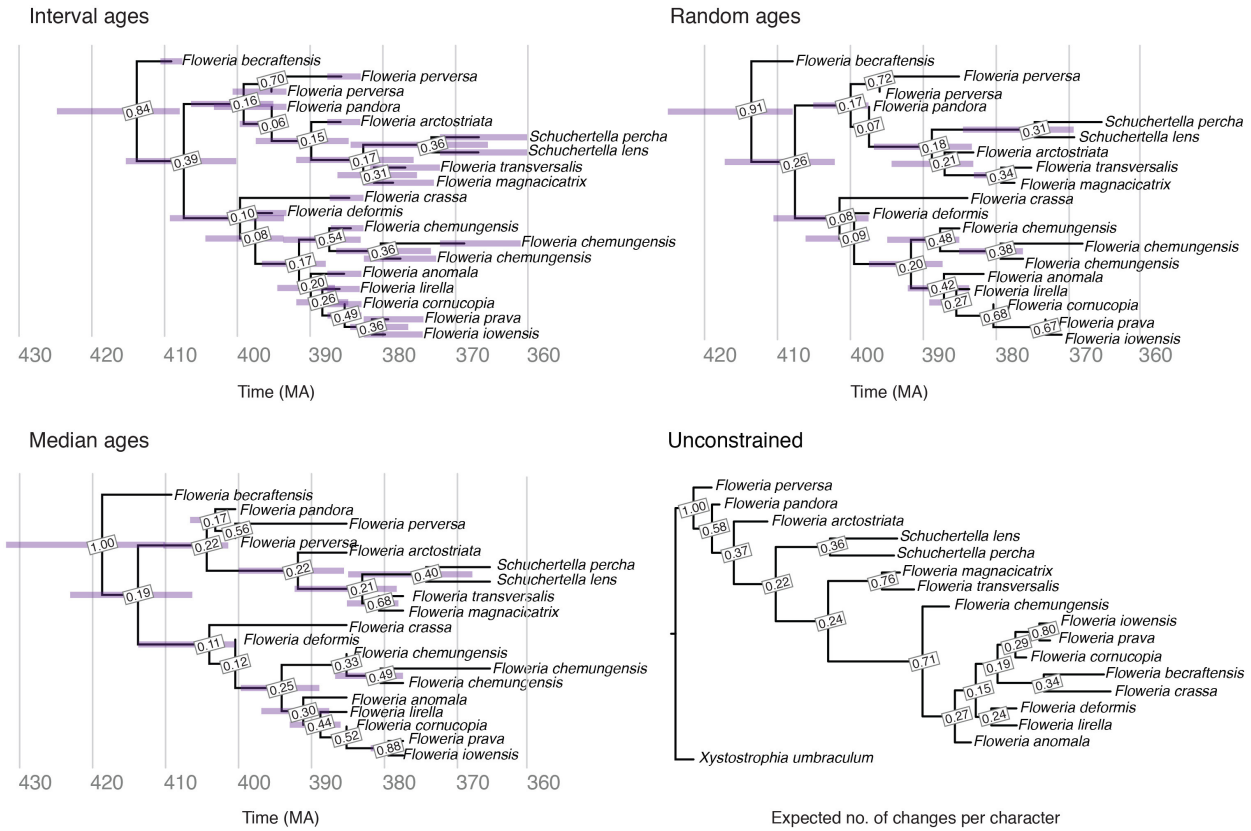


FIGURE 6 | Brachiopod MCC trees obtained using the FBD analysis with interval ages, median ages and random ages, and unconstrained analysis. Posterior support is shown for each node for all trees. Error bars on the FBD trees show the 95% HPD interval for the age of each node, as well as the age of each fossil in the tree with interval ages.

unknown and used morphological data to co-estimate topology along with divergence times. The results of our simulations show that in addition to recovering inaccurate divergences times, mishandling fossil age uncertainty can also result in the wrong tree (**Figures 2**, **5**).

We did not examine the impact of non-uniform fossil recovery, though this is known to decrease performance of the FBD model if unaccounted for (Heath et al., 2014; Luo et al., 2019; O'Reilly and Donoghue, 2019). Overall, our simulated datasets were designed to represent a best-case scenario for a fully extinct Paleozoic clade. We anticipate that additional, unaccounted-for model violations, such as non-uniform fossil recovery, would increase the errors in topology and divergence times estimates reported in this study.

Similarly, we did not examine the impact of morphological model violations such as rate heterogeneity among characters or the effects of non-uniform missing character data. A recent study suggested that even large deviations from the true model may have limited impact on divergence time estimates using total-evidence dating under the uniform tree model (Klopfstein et al., 2019). However, none of their simulation scenarios excluded molecular data and thus these findings may not be applicable to fully extinct clades. That said, the overall number of phylogenetic characters may be more of a concern for extinct clades, given the large degree of uncertainty associated with small matrices.

Small character matrices can be due to low taxon sampling, low character sampling, or both. The effect of both has been examined in previous studies. For example, simulations focused on unconstrained (i.e., non time-calibrated) Bayesian inference have shown that small morphological matrices (e.g., 100 characters or less) will result in highly uncertain trees (O'Reilly and Donoghue, 2017; Puttick et al., 2017). Similarly, several simulation studies have demonstrated the importance of having sufficient fossil sampling in order to recover reliable estimates of divergence times using the FBD model (Heath et al., 2014; O'Reilly and Donoghue, 2019). Luo et al. (2019) examined the combined effects of fossil and character sampling on total-evidence estimates of time and topology, including a scenario that used morphological data only. Similar to our findings, their results show that increasing both the number of fossil samples and morphological characters leads to better estimates of time and topology, in terms of accuracy and precision. They also compared the use of fixed vs. co-estimated fossil ages, where the age of fossils were fixed to the truth or ages were estimated from within the known interval of uncertainty. They found no strong differences in the estimated node ages when co-estimating fossil ages, which is coherent with our simulation scenarios, in which

we observe very little difference in accuracy when comparing true vs. co-estimated fossil ages.

Our simulations also show that the inclusion of fossil age information can improve the inferred topology regardless of the size of the matrix, if fossil age uncertainty is handled appropriately (**Figures 2**, **5**). On the other hand, excluding age information is preferable to using incorrect fossil ages even when using large morphological matrices. Thus, stratigraphic age uncertainty must be taken into account in order to fully benefit from the inclusion of fossil sampling times in the analysis. Note that we focus on the impact of stratigraphic age uncertainty, and not any uncertainty associated with the total duration over which a species is observed in the fossil record—that is, the stratigraphic range of a species (Hopkins et al., 2018). Different approaches to handling stratigraphic range durations have been shown recently to introduce errors into phylogenetic dating using the uniform tree model (Püschel et al., 2020). This type of data may be more appropriately modeled using the FBD *range* process (Stadler et al., 2018). However, we emphasize that the start and end of species ranges will also be associated with fossil age uncertainty, which will be essential to consider, regardless of the tree model (see also O'Reilly et al., 2015).

Assuming that fossil age uncertainty is handled appropriately, our results based on simulated and empirical data indicate that the priority for improving topology and divergence times should be to increase matrix size. However, some clades are naturally small or rare. For these clades, even with complete taxon sampling, the size of the dataset will remain small. The best course of action then may be to increase the taxonomic scope of the study and to sample more broadly. In the case of fossil clades, small numbers of characters may reflect the paucity of morphological trait data available from some groups whose record is characterized by exoskeletal or shell elements exhibiting minimal morphological variation. However, small matrix size might also reflect the historical circumstances in which these data were generated: many matrices surveyed (**Appendix S1**) were constructed for parsimony analysis where the focus was on the selection of phylogenetically informative characters and not necessarily intended to represent an exhaustive survey of the preserved variation. Moreover, some previous studies excluded a subset of characters from consideration because of *a priori* concerns about homoplasy (Guensburg and Sprinkle, 2003), and therefore only sampled characters they considered relevant or taxonomically significant. In this regard, it is conceivable that many published matrices may be expanded by a resurvey of the taxa of interest.

In addition, continuous trait data could provide an additional source of morphological information to complement matrices of discrete characters. Models of continuous trait evolution can be used to infer topology (Parins-Fukuchi, 2017) and divergence times (Alvarez-Carretero et al., 2019). Continuous data has also been shown to capture higher phylogenetic signal compared to discrete characters and can result in more accurate trees (Parins-Fukuchi, 2018). It is worth noting that >30% of the characters in our empirical example using brachiopods are continuous characters broken down into discrete states. If the process of discretization results in a loss of phylogenetic information, then tree inference and divergence time estimation could potentially be improved by modeling discrete and continuous characters separately.

We note that previous simulations examining the performance of both unconstrained vs. time calibrated Bayesian phylogenetic inference tend to use a minimum of 100 characters, which is >3 times the size of datasets available for many fossil invertebrate groups. Matrices of only 20–30 characters, which are widely used in the literature, may contain too much uncertainty for other methodological choices to matter. Thus, we must be realistic about the degree of uncertainty expected when the number of phylogenetic characters sampled is low. All approaches to constructing summary trees are problematic when there is a lot of uncertainty in the posterior and all summary trees should be interpreted with caution (O'Reilly and Donoghue, 2017). In conclusion, we show that as more phylogenetically informative data become available, fixing the fossil ages to incorrect values can lead to important errors. Sampling fossil ages as part of the inference recovers estimates similar to those obtained when fixing the ages to the correct values. Consequently, we recommend incorporating stratigraphic age uncertainty when conducting analyses using the FBD process.

DATA AVAILABILITY STATEMENT

The extended SA package is available on GitHub, https://github.com/CompEvol/sampled-ancestors and via the BEAST2 package manager. The R scripts used to simulate and analyse the data, as well as the configuration files used to run BEAST2 and RevBayes, are available on Zenodo, https://doi.org/10.5281/zenodo.3825626.

AUTHOR CONTRIBUTIONS

JB-S and NT implemented the study, ran the simulations, and analyzed the results. MH and DW assembled the empirical datasets and analyzed the empirical results. TS provided feedback on the study design and results. RW designed the study and interpreted the results. All authors contributed to writing the manuscript.

ACKNOWLEDGMENTS

This is Paleobiology Database official publication #373. JB-S was supported by funds from the National Science Foundation (US), grant DBI-1759909. DW acknowledges support from the Gerstner Scholars Fellowship and the Gerstner Family Foundation, the Lerner-Gray Fund for Marine Research, and the Richard Gilder Graduate School, American Museum of Natural History, as well as a Norman Newell Early Career Grant from the Paleontological Society. RW was funded by the ETH Zürich Postdoctoral Fellowship and Marie Curie Actions for People COFUND programme.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2020.00183/full#supplementary-material

REFERENCES

Alvarez-Carretero, S., Goswami, A., Yang, Z., and Dos Reis, M. (2019). Bayesian estimation of species divergence times using correlated quantitative characters. *Syst. Biol.* 68, 967–986. doi: 10.1093/sysbio/syz015

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

Barido-Sottani, J., Aguirre-Fernández, G., Hopkins, M. H., Stadler, T., and Warnock, R. (2019a). Ignoring stratigraphic age uncertainty leads to erroneous estimates of species divergence times under the fossilized birth-death process. *Proc. R. Soc. Lond. B Biol. Sci.* 286:20190685. doi: 10.1098/rspb.2019.0685

Barido-Sottani, J., Pett, W., O'Reilly, J. E., and Warnock, R. C. (2019b). Fossilsim: an r package for simulating fossil occurrence data under mechanistic models of preservation and recovery. *Methods Ecol. Evol.* 10, 835–840. doi: 10.1111/2041-210X.13170

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., et al. (2014). Beast 2: A software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* 10:e1003537. doi: 10.1371/journal.pcbi.1003537

Drummond, A. J., and Stadler, T. (2016). Bayesian phylogenetic estimation of fossil ages. *Philos. Trans. R. Soc. B Biol. Sci.* 371:20150129. doi: 10.1098/rstb.2015.0129

Gavryushkina, A., Heath, T. A., Ksepka, D. T., Stadler, T., Welch, D., and Drummond, A. J. (2017). Bayesian total-evidence dating reveals the recent crown radiation of penguins. *Syst. Biol.* 66, 57–73. doi: 10.1093/sysbio/syw060

Gavryushkina, A., Welch, D., Stadler, T., and Drummond, A. J. (2014). Bayesian inference of sampled ancestor trees for epidemiology and fossil calibration. *PLoS Comput. Biol.* 10:e1003919. doi: 10.1371/journal.pcbi.1003919

Grimm, G. W., Kapli, P., Bomfleur, B., McLoughlin, S., and Renner, S. S. (2015). Using more than the oldest fossils: dating osmundaceae with three bayesian clock approaches. *Syst. Biol.* 64, 396–405. doi: 10.1093/sysbio/syu108

Guensburg, T., and Sprinkle, J. (2003). The oldest known crinoids (early ordovician, utah) and a new crinoid plate homology system. *Bull. Am. Paleontol.* 364:1–43. Available online at: https://www.biodiversitylibrary.org/page/27980785#page/677/mode/1up

Heath, T. A., Huelsenbeck, J. P., and Stadler, T. (2014). The fossilized birth-death process for coherent calibration of divergence-time estimates. *Proc. Natl. Acad. Sci. U.S.A.* 111, E2957–E2966. doi: 10.1073/pnas.1319091111

Höhna, S., Landis, M. J., Heath, T. A., Boussau, B., Lartillot, N., Moore, B. R., et al. (2016). RevBayes: bayesian phylogenetic inference using graphical models and an interactive model-specification language. *Syst. Biol.* 65, 726–736. doi: 10.1093/sysbio/syw021

Hopkins, M. J., Bapst, D. W., Simpson, C., and Warnock, R. C. (2018). The inseparability of sampling and time and its influence on attempts to unify the molecular and fossil records. *Paleobiology* 44, 561–574. doi: 10.1017/pab.2018.27

Klopfstein, S., Ryser, R., Corio, M., and Spasejovic, T. (2019). Mismatch of the morphology model is mostly unproblematic in total-evidence dating: insights from an extensive simulation study. *bioRxiv* 679084. doi: 10.1101/679084

Larabee, F. J., Fisher, B. L., Schmidt, C. A., Matos-Maraví, P., Janda, M., and Suarez, A. V. (2016). Molecular phylogenetics and diversification of trap-jaw ants in the genera Anochetus and Odontomachus (Hymenoptera: Formicidae). *Mol. Phylogenet. Evol.* 103, 143–154. doi: 10.1016/j.ympev.2016.07.024

Lee, M. S., Cau, A., Naish, D., and Dyke, G. J. (2014). Morphological clocks in paleontology, and a mid-cretaceous origin of crown aves. *Syst. Biol.* 63, 442–449. doi: 10.1093/sysbio/syt110

Lewis, P. O. (2001). A likelihood approach to estimating phylogeny from discrete morphological character data. *Syst. Biol.* 50, 913–925. doi: 10.1080/106351501753462876

Luo, A., Duchêne, D. A., Zhang, C., Zhu, C.-D., and Ho, S. (2019). A simulation-based evaluation of tip-dating under the fossilized birth-death process. *Syst. Biol.* 69, 325–344. doi: 10.1093/sysbio/syz038

Matschiner, M. (2019). Selective sampling of species and fossils influences age estimates of the fossilized birth-death model. *Front. Genet.* 10:1064. doi: 10.3389/fgene.2019.01064

Matschiner, M., Musilová, Z., Barth, J. M., Starostová, Z., Salzburger, W., Steel, M., and Bouckaert, R. (2017). Bayesian phylogenetic estimation of clade ages supports trans-atlantic dispersal of cichlid fishes. *Syst. Biol.* 66, 3–22. doi: 10.1093/sysbio/syw076

Menning, M., Alekseev, A., Chuvashov, B., Davydov, V., Devuyst, F.-X., Forke, H., et al. (2006). Global time scale and regional stratigraphic reference scales of central and West Europe, East Europe, Tethys, South China, and North America as used in the devonian-carboniferous-permian correlation chart 2003 (dcp 2003). *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 240, 318–372. doi: 10.1016/j.palaeo.2006.03.058

O'Reilly, J. E., and Donoghue, P. C. (2017). The efficacy of consensus tree methods for summarizing phylogenetic relationships from a posterior sample of trees estimated from morphological data. *Syst. Biol.* 67, 354–362. doi: 10.1093/sysbio/syx086

O'Reilly, J. E., and Donoghue, P. C. (2019). The effect of fossil sampling on the estimation of divergence times with the fossilised birth death process. *Syst. Biol.* 69, 124–138. doi: 10.1093/sysbio/syz037

O'Reilly, J. E., dos Reis, M., and Donoghue, P. C. (2015). Dating tips for divergence-time estimation. *Trends Genet.* 31, 637–650. doi: 10.1016/j.tig.2015.08.001

Parins-Fukuchi, C. (2017). Use of continuous traits can improve morphological phylogenetics. *Syst. Biol.* 67, 328–339. doi: 10.1093/sysbio/syx072

Parins-Fukuchi, C. (2018). Bayesian placement of fossils on phylogenies using quantitative morphometric data. *Evolution* 72, 1801–1814. doi: 10.1111/evo.13516

Paterson, J. R., Edgecombe, G. D., and Lee, M. S. (2019). Trilobite evolutionary rates constrain the duration of the Cambrian explosion. *Proc. Natl. Acad. Sci. U.S.A.* 116, 4394–4399. doi: 10.1073/pnas.1819366116

Pennell, M. W., Eastman, J. M., Slater, G. J., Brown, J. W., Uyeda, J. C., FitzJohn, R. G., et al. (2014). geiger v2. 0: an expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. *Bioinformatics* 30, 2216–2218. doi: 10.1093/bioinformatics/btu181

Püschel, H. P., O'Reilly, J. E., Pisani, D., and Donoghue, P. C. (2020). The impact of fossil stratigraphic ranges on tip-calibration, and the accuracy and precision of divergence time estimates. *Palaeontology* 63, 67–83. doi: 10.1111/pala.12443

Puttick, M. N., O'Reilly, J. E., Tanner, A. R., Fleming, J. F., Clark, J., Holloway, L., et al. (2017). Uncertain-tree: discriminating among competing approaches to the phylogenetic analysis of phenotype data. *Proc. R. Soc. B Biol. Sci.* 284:20162290. doi: 10.1098/rspb.2016.2290

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., and Suchard, M. A. (2018). Posterior summarization in bayesian phylogenetics using tracer 1.7. *Syst. Biol.* 67, 901–904. doi: 10.1093/sysbio/syy032

Robinson, D., and Foulds, L. (1981). Comparison of phylogenetic trees. *Math. Biosci.* 53, 131–147. doi: 10.1016/0025-5564(81)90043-2

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. *Syst. Biol.* 61, 973–979. doi: 10.1093/sysbio/sys058

Schliep, K. P. (2010). phangorn: phylogenetic analysis in R. *Bioinformatics* 27, 592–593. doi: 10.1093/bioinformatics/btq706

Slater, G. J. (2015). Iterative adaptive radiations of fossil canids show no evidence for diversity-dependent trait evolution. *Proc. Natl. Acad. Sci. U.S.A.* 112, 4897–4902. doi: 10.1073/pnas.1403666111

Stadler, T. (2010). Sampling-through-time in birth-death trees. *J. Theor. Biol.* 267, 396–404. doi: 10.1016/j.jtbi.2010.09.010

Stadler, T. (2011). Simulating trees with a fixed number of extant species. *Syst. Biol.* 60, 676–684. doi: 10.1093/sysbio/syr029

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

Stigall Rode, A. L. (2005). Systematic revision of the middle and late devonian brachiopods schizophoria (schizophoria) and "schuchertella" from North America. *J. Syst. Palaeontol.* 3, 133–167. doi: 10.1017/S1477201905001537

Wright, A. M., and Hillis, D. M. (2014). Bayesian analysis using a simple likelihood model outperforms parsimony for estimation of phylogeny from discrete morphological data. *PLoS ONE* 9:e109210. doi: 10.1371/journal.pone.0109210

Wright, D. F. (2017a). Bayesian estimation of fossil phylogenies and the evolution of early to middle paleozoic crinoids (echinodermata). *J. Paleontol.* 91, 799–814. doi: 10.1017/jpa.2016.141

Wright, D. F. (2017b). Phenotypic innovation and adaptive constraints in the evolutionary radiation of palaeozoic crinoids. *Sci. Rep*. 7:13745. doi: 10.1038/s41598-017-13979-9

Wright, D. F., and Toom, U. (2017). New crinoids from the baltic region (estonia): fossil tip-dating phylogenetics constrains the origin and Ordovician–silurian diversification of the flexibilia (echinodermata). *Palaeontology* 60, 893–910. doi: 10.1111/pala.12324

Zhang, C., Stadler, T., Klopfstein, S., Heath, T. A., and Ronquist, F. (2015). Total-evidence dating under the fossilized birth-death process. *Syst. Biol*. 65, 228–249. doi: 10.1093/sysbio/syv080

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Barido-Sottani, van Tiel, Hopkins, Wright, Stadler and Warnock. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genetics

ORIGINAL RESEARCH
published: 08 July 2020
doi: 10.3389/fgene.2020.00651

Conflict Resolution for Mesozoic Mammals: Reconciling Phylogenetic Incongruence Among Anatomical Regions

Mélina A. Celik and Matthew J. Phillips**

School of Biology and Environmental Science, Queensland University of Technology, Brisbane, QLD, Australia

The evolutionary history of Mesozoic mammaliaformes is well studied. Although the backbone of their phylogeny is well resolved, the placement of ecologically specialized groups has remained uncertain. Functional and developmental covariation has long been identified as an important source of phylogenetic error, yet combining incongruent morphological characters altogether is currently a common practice when reconstructing phylogenetic relationships. Ignoring incongruence may inflate the confidence in reconstructing relationships, particularly for the placement of highly derived and ecologically specialized taxa, such as among australosphenidans (particularly, crown monotremes), haramiyidans, and multituberculates. The alternative placement of these highly derived clades can alter the taxonomic constituency and temporal origin of the mammalian crown group. Based on prior hypotheses and correlated homoplasy analyses, we identified cheek teeth and shoulder girdle character complexes as having a high potential to introduce phylogenetic error. We showed that incongruence among mandibulodental, cranial, and postcranial anatomical partitions for the placement of the australosphenidans, haramiyids, and multituberculates could largely be explained by apparently non-phylogenetic covariance from cheek teeth and shoulder girdle characters. Excluding these character complexes brought agreement between anatomical regions and improved the confidence in tree topology. These results emphasize the importance of considering and ameliorating major sources of bias in morphological data, and we anticipate that these will be valuable for confidently integrating morphological and molecular data in phylogenetic and dating analyses.

Keywords: Mesozoic mammals, correlated homoplasy, incongruence, australosphenida, haramiyida, Multituberculata

OPEN ACCESS

Edited by:
Rosane Garcia Collevatti, Universidade Federal de Goiás, Brazil

Reviewed by:
Iker Irisarri, National Museum of Natural Sciences (MNCN), Spain
Edivaldo Herculano Correa De Oliveira, Evandro Chagas Institute, Brazil

****Correspondence:***
Mélina A. Celik
melina.celik@gmail.com
Matthew J. Phillips
m9.phillips@qut.edu.au

Specialty section:
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Genetics

Received: *03 November 2019*
Accepted: *28 May 2020*
Published: *08 July 2020*

Citation:
Celik MA and Phillips MJ (2020) Conflict Resolution for Mesozoic Mammals: Reconciling Phylogenetic Incongruence Among Anatomical Regions. Front. Genet. 11:0651. doi: 10.3389/fgene.2020.00651

INTRODUCTION

The mammalian crown group includes monotremes and therians (marsupials and placentals) and all extinct descendants of their most recent common ancestor. Mammals share characteristics such as lactation from mammary glands, hair (at least ancestrally), enucleate red blood cells, a muscular diaphragm, and a jaw joint formed by the dentary (mandible) and the squamosal (cranium). Defining Mammalia and its membership among the modern fauna is trivial. However, extending

this task to fossils has proven far more elusive. From mammaliaform origins in the Triassic to the Jurassic, early mammals underwent profound morphological changes. These included the evolution of several character complexes that were upheld as high-value markers for defining mammalian taxonomic exclusion or inclusion but that have since been revealed to be homoplastic. For example, tribosphenic molars, which combine piercing/slicing and grinding functions, and detachment of the middle ear ossicles from the dentary have both now been shown to have evolved independently several times (Allin, 1975; Chow and Rich, 1982; Wang et al., 2001; Martin and Luo, 2005; Luo et al., 2007; Luo, 2011; Ramírez-Chaves et al., 2016).

Over the past three decades, the task of inferring relationships among Mesozoic mammals has shifted emphasis from intuitive form/function assessments of character complexes (*e.g.*, Kemp, 1983; Szalay, 1993) to analyses of morphological character matrices (*e.g.*, Rowe, 1988; Luo et al., 2002; Rougier et al., 2007; Luo et al., 2015). The hope with these increasingly large taxon sets of ever more numerous and finely distinguished characters is that the stochastic error is reduced, and long branches are sufficiently broken up by taxon sampling to tease apart phylogenetic signal from homoplasy. In practice, these theoretical benefits are tempered by extensive character incompleteness, temporal gaps in the fossil record, and models that do not closely reflect morphological evolution, such as assuming constant rather than episodic evolutionary rates across lineages and character complexes. Nevertheless, phylogenetic inference has been enhanced by a recent explosion in sampling Mesozoic mammals (especially from China), including early members of many clades (Meng, 2014; Luo et al., 2017; Zhou et al., 2019).

The Generalized Insectivore Backbone of Mammal Phylogeny

This study focuses on reconstructing the affinities of dentally specialized or otherwise ecologically highly derived mammals. This endeavor nevertheless requires a well-resolved and reliable backbone of more generalized mammals. Plesiomorphic mammalian ecospace was occupied by relatively small insectivore/carnivores (Jenkins and Parrington, 1976; Kermack and Kermack, 1984), which also provided the basic body plan for major new waves of mammalian diversifications throughout the Mesozoic, from the Triassic morganucodont-like predecessors to the Cretaceous ancestors of the modern clades of Marsupialia and Placentalia. The ancestors of these therian clades (and potentially Mammalia) have typically been considered to have been terrestrial-scansorial (*e.g.*, Jenkins, 1970; Kermack and Kermack, 1984; Krebs, 1991; Szalay, 1994). However, the ecological diversity of dietarily plesiomorphic Mesozoic mammals has turned out to be far more disparate for their locomotor modes, including fossorial (Luo and Wible, 2005; Luo et al., 2015), semi-aquatic (Ji et al., 2006), arboreal, and even gliding mammaliaforms (Luo, 2007).

It is a general consensus that taxa with ecological similarities tend to be close and even cluster in their morphospace distribution, as demonstrated by recent empirical studies (Chen and Wilson, 2015; Grossnickle, 2017; Benevento et al., 2019). In theory, this should assist with discriminating synapomorphy from homoplasy in phylogenetic inference (Hendy and Penny, 1989). There is less obvious potential for large-scale and correlated mandibulodental convergence if sampling only generalized insectivores. However, there is potential for cranial and postcranial convergence among some taxa, for example, with independent ecospace shifts into fossorial, semi-aquatic, or highly arboreal niches. There is also potential for parallelism (as distinct from convergence) between generalized insectivore lineages following evolutionary trends that were initiated before they diverged. Examples of such evolutionary momentum among generalized insectivores have been suggested for trends toward tribospheny (Butler, 1990; Davis, 2011) and more parasagittal posture (Ji et al., 1999). Overall, however, relative morphological conservatism and stationarity (similar evolutionary processes operating across lineages) should enhance congruence among anatomical regions and resolution overall for the placements of well-sampled, ecologically generalized insectivores in comparison to more ecologically derived taxa, such as monotremes, haramiyidans, and multituberculates.

As it turns out, most inferences of Mesozoic mammal phylogeny broadly agree on the relationships among at least the well-sampled, small insectivorous/carnivorous mammal clades (Szalay et al., 1993; Luo et al., 2002; Rougier et al., 2011). To summarize, morganucodonts and docodonts, which both retained the primitive dentary-attached middle ear bones and sprawling posture, are placed outside of crown Mammalia. The potentially paraphyletic eutriconodonts retained plesiomorphic "in-line" molar cusps but have more derived crania and postcrania and are placed on the therian stem lineage between docodonts and spalacotherioids (which have "reverse triangle" molar cusps). Closer still to the divergence of marsupial and placental therians are successive ranks of extinct clades of the cladotherian group, such as *Henkelotherium*, *Vincelestes*, and *Peramus*, among which can be traced the early development of the posterior (talonid) component of tribosphenic molars (Butler, 1990) and partly to fully coiled cochlea with internal bony structure to support the hearing organ for improved high- and low-frequency hearing (Manoussaki et al., 2008; Luo et al., 2011; Manley, 2018).

Dietarily Apomorphic Mammals

Ecological deviations from the shrew-like archetype among Mesozoic mammals evolved across several lineages. We will later consider cranial and postcranial characters, but here we pay particular attention to the impact of substantial dietary apomorphy on inferring Mesozoic mammal relationships. The reasons for this are twofold: (1) All three major clades at the crux of current debates surrounding the temporal origin and the taxonomic composition of the mammalian crown group are mired in arguments over molar cusp homology. A key argument is whether the multi-cuspate molars and omnivory–herbivory adapted craniomandibular geometries of haramiyidans and multituberculates evolved convergently (Luo et al., 2015, 2017) or are indicative of shared ancestry as "allotherians" (Bi et al., 2014; Meng et al., 2014). The near-tribosphenic teeth

of the Cretaceous monotremes, *Steropodon* and *Teinolophos*, and their stem australosphenidan relatives are also thought to have evolved independently of therian tribosphenic teeth (Luo et al., 2001). Some authors (Archer et al., 1993) have even preferred to use alternative akid (blade) terminology rather than cusp terminology for fossil monotremes to side-step the uncertainty in interpreting cusps on the trigonid and the talonid of molars; and (2) there is a strong prior expectation for extensive, correlated mandibulodental convergence associated with dietary evolution based on comparative anatomy of modern mammals and developmental genetics (Kangas et al., 2004; Springer et al., 2013). Moreover, mandibulodental characters comprise 41% of the data matrix that we employ (based on Huttenlocker et al., 2018), and so correlated homoplasy among these characters could present a profound phylogenetic bias.

The oldest australosphenidans appear to have been relatively generalized insectivores, including the earliest monotremes (*e.g.*, *Teinolophos*). However, many mandibulodental characters scored for australosphenidans (including upper dental characters) were unable to be scored for the plesiomorphic taxa but were scored only for the highly specialized platypuses and echidna. It is in this context of potential phylogenetic influence that we refer to australosphenidans as dietarily or dentally apomorphic, alongside multituberculates and haramiyidans. The phylogenetic placement of dietarily apomorphic taxa has a strong bearing on the age of Mammalia. At present, the oldest fossils that can reliably be placed within the mammalian crown and can thus be employed to calibrate the monotreme-therian divergence are Middle Jurassic (∼163 Ma) in age, such as the cladotherian *Amphitherium* (*e.g.*, dos Reis et al., 2012), or perhaps slightly older, based on the australosphenidan *Asfaltomylos* (Rauhut et al., 2002). If, however, early haramiyidans such as *Haramiyavia* are closely related to multituberculates and are crown mammals (*e.g.*, Bi et al., 2014; Krause et al., 2014), then Mammalia has a Triassic origin, pre-dating ∼201 Ma. We aim to quantify the distribution of correlated homoplasy across the skeleton that arises upon the inclusion of dietarily derived taxa in the mammal tree. This may further inform the validity of combining data from different regions or support the preferential use of certain anatomical regions.

Australosphenida and the Monotremes

Living monotremes include the semi-aquatic platypus (*Ornithorhynchus anatinus*) and the semi-fossorial echidnas (Tachyglossidae), all of which present extremely specialized morphological features. The divergence of monotremes from therians has been placed as remotely as among primitive therapsids (*e.g.*, Simpson, 1959; Macintyre, 1967) to as recently as grouping with marsupials (*e.g.*, Gregory, 1947; Kühne, 1973, 1977). More recent developments substantially narrow this range. One of these, as we will shortly outline, is the discovery of stem monotreme fossils, and the other is relaxed clock molecular dating, which strongly favors monotremes diverging just 20-40 million years prior to the divergence of marsupials and placentals (*e.g.*, Kullberg et al., 2008; Phillips et al., 2009; Meredith et al., 2011). These timing estimates are consistent with recent morphological phylogenies, in which monotremes almost always diverge from therians closely before or after eutriconodonts (*e.g.*, Rougier et al., 1996; Meng et al., 2006; Rowe et al., 2008), sometimes with other dentally divergent taxa, such as multituberculates (*e.g.*, Meng et al., 2003) or the pseudo-tribosphenic shuotheriids (*e.g.*, Luo et al., 2007).

Stem monotremes include the Cretaceous *Steropodon* (Archer et al., 1985), *Ausktribosphenos*, and *Teinolophos* (Rich et al., 1997, 1999) and their Jurassic cousins, *Ambondro* from Madagascar (Flynn et al., 1999) and *Asfaltomylos* (Rauhut et al., 2002) and *Henosferus* (Rougier et al., 2007) from Patagonia. This discovery of a Gondwanan radiation of (near)tribosphenic mammals precipitated the classification of a new infraclass, Australosphenida (Luo et al., 2001, 2002; Kielan-Jaworowska et al., 2004). Although these taxa are known predominantly from incomplete dentaries and dentitions, these early australosphenidans, which were less ecologically and morphologically specialized than living monotremes, have already overturned interpretations of key character complexes that were central to earlier debates on the placement of monotremes. In particular, the absence of near-tribosphenic molars in living monotremes led to a notion that monotremes would not be close to the ancestry of therians (*e.g.*, Rowe, 1988). Conversely, the detachment of the middle ear bones from the mandible could be interpreted as evidence for linking monotremes to therians if fossils had been excluded (*e.g.*, Miao, 1993). After the discoveries of tribosphenid-like molars in toothed monotremes from the Cretaceous (Archer et al., 1985; Rich et al., 2016) and the finding that the middle ear was still attached to the mandible (Rich et al., 2016), the absence of these features in living monotremes is now shown to be the result of evolutionary convergence in the ear and reversal of tribosphenic teeth in living monotremes.

The near-complete humerus from a stem monotreme, *Kryoryctes cadburyi* (Pridmore et al., 2005), gives the first fossil insights on the evolution of the humero-ulnar articulation in australosphenidans (Phillips, 2002; Pridmore et al., 2005). Modern monotreme humeri have a bulbous ulnar condyle superficially similar to those of early mammaliaformes and multituberculates. In *Kryoryctes*, the humero-ulnar articulation is intermediate between these convex ulnar condyles and the pulley-like ulnar trochlea morphology of modern therians. The concavity of this trochlea, although wide and shallow, is limited to the dorsal/posterior aspect, as for putative stem therians such as eutriconodonts (*e.g.*, *Jeholodens*, Ji et al., 1999). Thus, the monotreme ulnar condyle may not be homologous with those of early mammaliaformes but instead derived from an ulnar trochlea, perhaps similar to those of close relatives of therians. Thus, even the fragmentary remains from Mesozoic australosphenidans are confirming long-held beliefs (Gregory, 1947; Phillips et al., 2009) that much of modern monotreme morphology is highly derived in association with their ecology rather than indicative of phylogeny. A key question now is whether the recent consensus on australosphenidan affinities lying close to eutriconodonts is underpinned by phylogenetic signal agreement across anatomical regions or is an emergent statistical property of averaging over incongruent phylogenetic signals.

Combined Evidence and Interrogation of Homoplasy

Homoplasy among mammals is widespread across all anatomical regions (Sánchez-Villagra and Williams, 1998; Ji et al., 1999; Sansom and Wills, 2017). The so-called total evidence approach, in which all relevant characters are combined, has often been recommended on an assumption that signals from shared ancestry would prevail over homoplasies to recover the true phylogeny. The expectation is that the phylogenetic estimates may be more accurate when all characters are combined, even in the face of significant incongruence between data partitions (Kluge, 1989; O'Leary et al., 2003; Fitzhugh, 2006; Asher, 2007; Mounce et al., 2016). A series of new discoveries of early mammal fossils preserved with basicrania (*e.g.*, Rougier and Novacek, 1998) and postcrania (*e.g.*, Ji et al., 1999) have made it feasible to combine anatomical regions for total evidence phylogenetic analyses of Mesozoic mammals (*e.g.*, Luo et al., 2002), but whether different anatomical regions would yield heterogeneous signals (Naylor and Adams, 2001) has received little examination for Mesozoic mammals. For paleontological studies, a total evidence approach can also help to add the less complete taxa that are nonetheless important for tracing the evolution of particular characters and for other reasons, *e.g.*, being the earliest-known member of a clade (Luo et al., 2002; Kielan-Jaworowska et al., 2004).

Arguments against total evidence, in favor of taxonomic congruence (*e.g.*, Miyamoto and Fitch, 1995; Naylor and Adams, 2001), are concerned that combining data partitions to "democratize" evidence ignores the issue of incongruence. Such masking of interaction effects is anathema for most fields of science and statistical analysis of variance, wherein understanding incongruence has long been the preferred path to resolution (Brown, 1975). Even among advocates of total evidence, there is room for excluding some characters from phylogenetic inference. Kluge (1989), for example, wrote that "including all relevant evidence can be seen as a harmless activity, unless one is prepared to argue *a priori* that certain evidence will confound the analysis and must therefore be eliminated." The strongest advocacy for exploring incongruence and excluding data partitions has come from molecular phylogenetics, for which it has often been a relatively trivial matter to move beyond identifying the presence of incongruence to isolating the sources of misleading phylogenetic signals and identifying the underlying processes. Examples include compositional biases in mitochondrial DNA transition substitutions, which erroneously group monotremes with marsupials (Phillips and Penny, 2003), and convergent selection pressures phylogenetically grouping echolocating bats and dolphins based on the hearing gene, *Prestin* (Liu et al., 2010).

Molecular incongruence (such as noted above) at phylogenetic levels beyond the reach of deep coalescence and introgressive hybridization tends to stand out against a background of congruence among multiple, unlinked genes or is revealed by improved substitution modeling. These luxuries may be less applicable to morphological phylogeny. However, some morphological character complexes may be expected to be less reliable. Molar cusp patterns, for instance, often possess a combination of rapid evolution (which erodes phylogenetic signal) and functional/developmental correlations that can overwhelm the remaining phylogenetic signal (Luo et al., 2001; Kangas et al., 2004; Ramírez-Chaves et al., 2016). However, non-phylogenetic sources of character covariation in vertebrate morphology are numerous and commonplace, including (but not limited to) allometry, pleiotropic–developmental, and functional correlations associated with ecological convergence or evolutionary trends (parallelism), and their respective influences vary dramatically across skeletal regions (Goswami et al., 2014; Cardini et al., 2015). Hence, complex distributions of phylogenetic incongruence may emerge across anatomical region partitions (Sánchez-Villagra and Williams, 1998; Ji et al., 1999), which is not conducive to teasing apart correlated homoplasy from phylogenetic signal.

Maximum likelihood and Bayesian inference methods that attempt to model the evolutionary process can provide an improved probabilistic framework for tracing character transformation (*e.g.*, Lewis, 2001; Ronquist, 2004; Lartillot and Poujol, 2011). In particular, they can accommodate evolutionary rate variation among regional partitions and taxa, including some forms of heterotachy (Kolaczkowski and Thornton, 2004)—shifts in character-specific evolutionary rates over time. However, the current models are not robust to functionally and developmentally correlated parallelism and convergence of character complexes. Most models assume that substitutions are independent and identically distributed. There are exceptions, such as the doublet model, which allows for correlation between paired RNA stem sites (*e.g.*, Tsagkogeorga et al., 2009; Phillips et al., 2010). However, evolutionary covariation among morphological characters associated with allometry, other developmental correlations, functional constraints, and selection cannot be so simply modeled.

By partitioning analyses between anatomical regions, we hope to better accommodate evolutionary rate variation across characters and lineages and isolate major sources of correlated homoplasy. Moreover, we develop a parsimony-based method to infer the relative magnitude of correlated homoplasy among anatomical partitions that is attributable to specified taxa. Thus, by interrogating incongruence, we hope to better understand its biological underpinning and more robustly inform the placement of ecologically derived mammals.

MATERIALS AND METHODS

Data

Our morphological dataset is based on Huttenlocker et al. (2018). The matrix consists of 537 morphological characters for 103 extinct mammaliaformes. Five deeper-diverging cynodonts are used as outgroups to root the phylogeny. Within australosphenidans, coding changes were made on contentious mandibular characters related to the presence of a postdentary trough and Meckel's groove, following Ramírez-Chaves et al. (2016). We also included the near-complete humerus from *Kryoryctes*, which provides the only stem monotreme postcranial

information. Two characters (body size and mandibular depth) that are phylogenetically informative among monotremes were included (from Phillips et al., 2009). The coding changes and justifications are provided in the **Supplementary Material**.

Taxon sampling for our initial analyses is informed by our aim to explore how incongruence between mandibulodental, cranial, and postcranial partitions varies with and without the inclusion of dentally specialized multituberculates, haramiyidans, and australosphenidans (for which crown monotremes contribute many mandibulodental and almost all cranial and postcranial characters). This requires a well-resolved backbone phylogeny of generalized insectivore/carnivores that are well sampled for characters across each of the three regional partitions. A near-complete, dietarily plesiomorphic (NCDP) taxon set was identified, which includes 27 taxa that are inferred to have been predominantly insectivorous/carnivorous. Taxa were only included in the corresponding $NCDP_{537}$ dataset if they were represented by >55% completeness for the 537 characters, including >35% completeness for each of the three regions or alternatively >65% completeness overall if >35% completeness across only two regions. Inclusion in the latter case also required local stability on the overall tree (>95% maximum parsimony and maximum likelihood bootstrap support) in agreement with general consensus among recent studies. This phylogenetic stability criterion allows such taxa to be phylogenetically constrained for analyses on the region for which they are poorly complete so as to not unduly influence other taxonomic placements or homoplasy metrics for these regions. Including such taxa breaks up long branches and informs character transformation when data are missing among their close relatives.

Initial Exploratory Analyses

A maximum parsimony (MP) backbone phylogeny was initially reconstructed in PAUP 4.0b10 (Swofford, 2002) for the 27 NCDP taxa. We focus on dietary plesiomorphy (generalized insectivore/carnivores) because postcranial data are sparse for many of the more derived taxa to be included later, and therefore locomotory plesiomorphy cannot be consistently inferred. The NCDP sampling comprised *Thrinaxodon*, *Pachygenelus*, and *Sinoconodon* as outgroups and then morganucodonts, docodonts, spalacotherioid symmetrodonts, eutherians, and metatherians, along with two potentially paraphyletic groups, "eutriconodonts" and "eupantotheres."

Initial MP bootstrap analyses were also run on Australosphenida and then for the multicuspate clades, Multituberculata and Haramiyida. In each case, *Thrinaxodon*, *Pachygenelus*, *Sinoconodon*, and morganucodonts were retained as outgroups. One of the aims of the initial MP bootstrap analyses was to compare the phylogenetic resolving power, with all characters unordered or with 72 of the multistate characters ordered (see the **Supplementary Material**). Character ordering can potentially enhance the phylogenetic signal by effectively increasing the steps associated with evolving more distinct character states, under certain assumptions (Slowinski, 1993). MP bootstrap support comparisons are shown in **Supplementary Table S1** and focus on groupings that are resolved in at least one of the ordered/unordered treatments, with between 60 and 95% bootstrap probability (BP). Other groupings are not considered because there is little benefit from either treatment if both give near full support or both are unable to resolve relationships. Among these comparisons, greater mean phylogenetic resolution was recovered under the ordered character treatment for the NCDP taxon set and for the Australosphenida taxon set (with and without *Kryoryctes*). Only the Haramiyida and Multituberculata analysis resulted in effectively the same support under both treatments. As such, from here on, we focus on analyses in which those 72 characters are treated as ordered, unless otherwise stated.

The MP bootstrap analyses were run with heuristic searches using random sequence addition with TBR branch swapping for 20 replicates across 1,000 pseudo-replications (full bootstrapping). In addition to the MP analyses, the phylogenies were also inferred under maximum likelihood (ML) and Bayesian inference for the NCDP taxa, including all 537 characters (we refer to this matrix as $NCDP_{537}$) and a 53-taxon dataset that augmented these generalized insectivore/carnivores with the australosphenidans, multituberculates, and haramiyidans (we refer to this matrix as 53_{537}) to denote both the number of taxa and characters. The ML analyses were conducted in IQ-TREE v1.6.11 (Nguyen et al., 2015) using the Mk + Gamma models (Yang, 1994; Lewis, 2001). One shortcoming of IQ-TREE is that the ordered and the unordered characters are separately partitioned, which increases parameterization with all branch lengths independently estimated for partitions (-sp option). The -spp option reduces parameterization because branch lengths are instead only proportionally scaled across partitions. Unfortunately, -spp precludes the biological realism of lineage-specific variation in branch lengths across partitions, which we expect for taxa that are, for example, dentally derived but postcranially plesiomorphic. As such, our main concern is whether it is more appropriate to separately partition the ordered and the unordered characters (six partitions: two for each of the mandibulodental, cranial, and postcranial regions) or employ only the three regional partitions with all characters unordered. Based on the primary 53_{537} dataset, the corrected Akaike information criterion (AIC) and the Bayesian information criterion (BIC) scores favor the three-partition scheme (3sp: $-\ln L$ = 8,430.2836, corrected AIC = 18,126.6648, BIC = 18,683.5067) over the six-partition scheme (6sp: $-\ln L$ = 8,078.5163, corrected AIC = 654,605.0326, BIC = 19,702.3355). Thus, we focus on the "3sp" ML analyses, which treat all characters as unordered. Results with ordered and unordered characters partitioned, but with proportionally scaled partition branch lengths, are provided in the **Supplementary Material** and **Supplementary Table S2**. Ultrafast bootstrap approximation (1,000 replicates) was used for assessing ML clade support.

Bayesian inference allowed ordered and unordered characters to be modeled within the same partition and was conducted with MrBayes 3.2.6 (Huelsenbeck and Ronquist, 2001). The Mkv model (Lewis, 2001) was employed with gamma-distributed rates across sites (G) for each of the mandibulodental, cranial, and postcranial partitions. Two independent analyses were run with three Markov chain Monte Carlo chains for 5,000,000

generations. The trees were sampled every 5,000 generations, with the first 25% discarded as burn-in. Clade frequencies across the two independent runs reached convergence (average standard deviation of split frequencies < 0.01), and the estimated sample sizes for the tree prior, likelihood, and posterior estimates for tree lengths and rate alpha parameters were >200 (Tracer v1.7.1; Rambaut and Drummond, 2014).

In all our exploratory analyses, the controversial haramiyids (or multituberculates), *Vintana* and Hahnodontidae (including *Cifelliodon*) were highly unstable in the tree (as outlined in our results) and substantially affected support at other nodes. Except where otherwise stated, these taxa from the 53_{537} data were subsequently excluded, resulting in the 51_{537} dataset.

Testing Phylogenetic Congruence Between Anatomical Regions

The phylogenetic incongruence between anatomical region partitions was initially tested within a maximum likelihood framework. The 537 morphological characters were partitioned into three anatomical regions to determine the phylogenetic signal across the data: mandibulodental (220 characters), cranial (183 characters), and postcranial (134 characters). We subsequently consider correlated homoplasy among more localized sub-regions; however, these broader regional partitions provide greater statistical power. Tree topologies were estimated for the overall dataset (with ML models partitioned) and for each of the three anatomical regions. Congruence between topologies was assessed using Kishino–Hasegawa (KH), and approximately unbiased (AU) tests were implemented in IQ-TREE. Initially, this incongruence testing was applied to the $NCDP_{537}$ and 51_{537} datasets. Our focus is on phylogenetic signals for relationships among the major groups and, therefore, to reduce the influence of alternative placements within major clades and control for this across all comparisons, constraints were applied within each of Morganucodonta, Docodonta, Gobiconodonta, Spalacotheroidea, Metatheria, Eutheria, Australosphenida, Multituberculata, and Haramiyida for groupings that received $\geq$ 90% bootstrap support on the overall dataset, in agreement with the general consensus among recent studies. The bootstrap support criterion was relaxed for constraining several taxa within the dietarily apomorphic clades, for which very few sampled characters limited the statistical power. Constraint trees are provided in the **Supplementary Material**. KH and AU incongruence testing was further applied to the placement of Australosphenida specifically on all internal branches for major groups on the $NCDP_{537}$ tree.

Bayesian inference incongruence testing was undertaken in MrBayes with the same constraints applied as for the ML analyses described above. In these analyses, the mandibulodental, cranial, and postcranial characters were again modeled separately under Mkv + G as partitions that included both ordered and unordered characters. For each of the $NCDP_{537}$ and 51_{537} datasets, the three partitions were initially constrained to share the same topology. Analyses were then run with different topologies allowed for each anatomical region partition. The 95% higher posterior densities (HPDs) for marginal likelihood were compared between these topologically linked and unlinked analyses in Tracer.

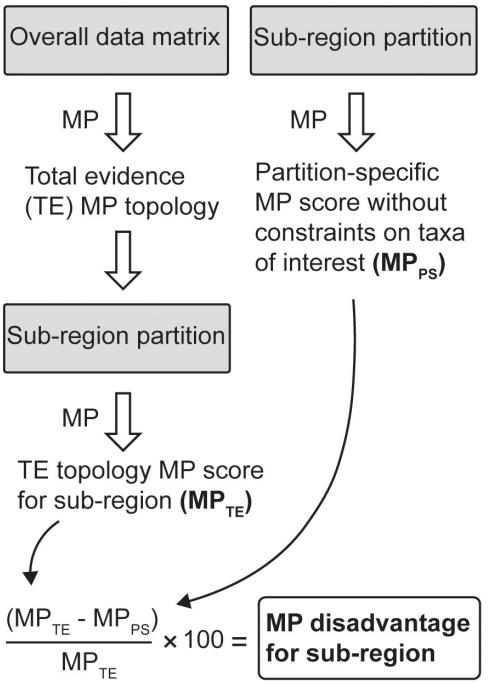


FIGURE 1 | Workflow for calculating the maximum parsimony disadvantage for a data partition with the tree constrained to the total evidence topology, relative to unconstrained placements for taxa of interest (in this case, australosphenidans, multituberculates, and haramiyidans).

Homoplasy Within Anatomical Regions

To more closely identify potential sources of correlated homoplasy induced by including the apomorphic clades, we present an MP-based metric, "MP disadvantage." The method is set out in **Figure 1**. First, we partitioned the data into 10 finer sub-regions (with the number of characters indicated in parentheses): mandibular (34), cheek teeth (163), other dental characters (23), basicranial (117), calvariaviscerocranial (68), shoulder girdle (24), axial (16), pelvic girdle (13), forelimb (17), and hindlimb (62). For each sub-region, the taxa within the 53_{537} dataset that were scored for fewer than 10 characters were deleted. For the remaining taxa, the most parsimonious trees were inferred in PAUP on the full 51_{537} dataset, and the minimum number of tree steps for these overall favored (total evidence) topologies was then inferred on the relevant sub-region data only, giving the total evidence MP score (MP_{TE}) for that sub-region. The minimum-length tree was then inferred for this taxon set on the sub-region data alone, without topological constraints on the taxa of interest and thus giving the partition-specific MP score (MP_{PS}) for that sub-region. We refer to the percentage tree-length difference between these MP_{TE} and MP_{PS} tree scores as the MP disadvantage of a sub-region being constrained to the total evidence topology.

Maximum parsimony disadvantage is an indicator of the correlated homoplasy among sub-regions that is attributable to including the australosphenidans, multituberculates, and haramiyidans since the $NCDP_{537}$ backbone constraint was employed for MP inference of the overall and sub-region trees. This constraint tree includes relationships among the 27 near-complete, dietarily plesiomorphic taxa that attained >90% MP bootstrap support (and are also generally agreed upon in recent analyses). Sub-regions with fewer than 20 characters (forelimb, axial, and pelvic girdle) were deemed to be unreliable and were not included in the main analysis. Fewer taxa could be included for these three sub-regions (*i.e.*, with at least 10 characters sampled), and furthermore, these sub-regions provided high variation around the expected values (**Supplementary Table S3**). Since evolutionary modularity studies (*e.g.*, Goswami et al., 2009) suggest that mammalian forelimbs, axial skeletons, and pelvic girdles are not functionally or developmentally closely correlated, it is also not appropriate for these sub-regions to be combined.

Correlated Homoplasy Reduction and Extension to Less Complete Taxa

The cheek teeth and shoulder girdle sub-regions were identified as contributing disproportionately high levels of correlated homoplasy upon adding haramiyidans, australosphenidans, and multituberculates into the backbone phylogeny of generalized insectivore/carnivores. To assess the impact of these sub-regions on phylogeny, we re-ran the 51-taxon analyses with the cheek teeth (163 characters) and shoulder girdle (24 characters) excluded, leaving 350 characters. A more inclusive 78-taxon dataset was also compiled upon lowering the taxon completeness requirement to at least 15% of the 350 characters. This allowed the effect of excluding major apparent sources of correlated homoplasy to be evaluated for a broader phylogenetic context, including for the placement of several important but less well-known taxa close to the mammalian and therian crown nodes, such as the proposed sister of australosphenidans, the Shuotheriidae (Kielan-Jaworowska et al., 2002), and the putatively oldest eutherian, *Juramaia* (Luo et al., 2011). Phylogenetic inference of these new 51_{350} and 78_{350} datasets employed MP, ML, and Bayesian inference, as described above.

RESULTS

Exploratory Phylogenetic Analyses

Initial phylogenetic analyses on the 27 highly complete "dietarily plesiomorphic" ($NCDP_{537}$) dataset reconstructed a well-supported mammalian backbone phylogeny of generalized insectivore/carnivores (**Figure 2A**). The modern consensus is retrieved with Eutheria, Metatheria, Cladotheria, Trechnotheria, and Theriimorpha all strongly supported, and moving stemwise, more inclusive clades that successively include docodonts, morganucodonts, and then *Sinoconodon* all received maximum bootstrap support (BP_{MP} and BP_{ML}) and Bayesian posterior probability (BPP). There is uncertainty regarding the relationships and potential paraphyly of the eutriconodonts; however, they are globally stable in that, among these generalized insectivore/carnivore taxa, eutriconodonts are all placed on the trechnothere stem lineage with maximum BP/BPP.

Separate MP bootstrap analyses were conducted for each of the main "dietarily apomorphic" taxa, Australosphenida, Multituberculata, and Haramiyida, rooted with stem-mammal outgroups. Strong support was recovered for key groupings within each of these clades (see the **Supplementary Material**). Within Australosphenida, the monotreme clade was recovered with 91% BP and, in turn, formed a clade (78% BP) with the other Australian australosphenidans, *Ausktribosphenos*, and *Bishops*. *Kryoryctes* was recovered among monotremes in all analyses with moderate support (>73% BP) despite only being included for 10 characters and not yet including several apomorphies with living monotremes that were noted by Pridmore et al. (2005) (see the **Supplementary Material**). Hence, we are confident of the placement of *Kryoryctes* with monotremes.

To enhance phylogenetic comparability and incongruence testing between mandibulodental, cranial, and postcranial regions, we focused on taxa that are well sampled across each of these regions. However, we also included several taxa that contribute much-needed information on otherwise less well-sampled sub-regions and that are locally stable on the overall dataset. These taxa can be reliably constrained in regional analyses so as to not unduly influence incongruence testing. The global stability of *Kryoryctes* and several mandibulodental taxa within Australosphenida (*Asfaltomylos*, *Ambondro*, *Ausktribosphenos*, *Bishops*, *Teinolophos*, and *Steropodon*) permits their inclusion for further analyses. Similarly, *Kuehneodon*, which is one of the oldest and most mandibulodentally complete multituberculates, did not meet the initial inclusion criteria but was stably placed as sister to *Rugosodon* (>90% BP) and therefore included in phylogenetic and incongruence testing analyses.

The controversial haramiyids or multituberculates, *Vintana* (Krause et al., 2014) and Hahnodontidae [including *Cifelliodon* (Huttenlocker et al., 2018)], were highly unstable in the tree. Although these two enigmatic taxa could contribute importantly to cranial character sampling, they are otherwise poorly known. MP and ML analyses respectively placed this Hahnodontidae–*Vintana* grouping outside Mammalia [nested within Haramiyida as sister to Eleutherodontidae, *sensu* Huttenlocker et al. (2018)] and as sister to multituberculates, nested well within Mammalia. In contrast, our Bayesian inference analysis recovered weakly supported placements of *Vintana* with multituberculates and, separately, Hahnodontidae within Haramiyida. These two widely separated local optima also prevented MrBayes runs from converging on a global optimum. Improved sampling of non-cranial material from these two taxa or further cranial material from haramiyidans may resolve these affinities. However, our primary aim is to demonstrate and identify incongruence between data partitions. To isolate the impact of incongruence on phylogenetic estimates and to reduce the uncertainty and possible errors associated with *Vintana* and Hahnodontidae, these taxa were excluded.

With *Vintana* and Hahnodontidae excluded, phylogenetic analyses of the resulting 51_{537} dataset recovers Multituberculata and Haramiyida as reciprocally monophyletic [as in Huttenlocker et al. (2018)], with 93–100% support in ML and Bayesian

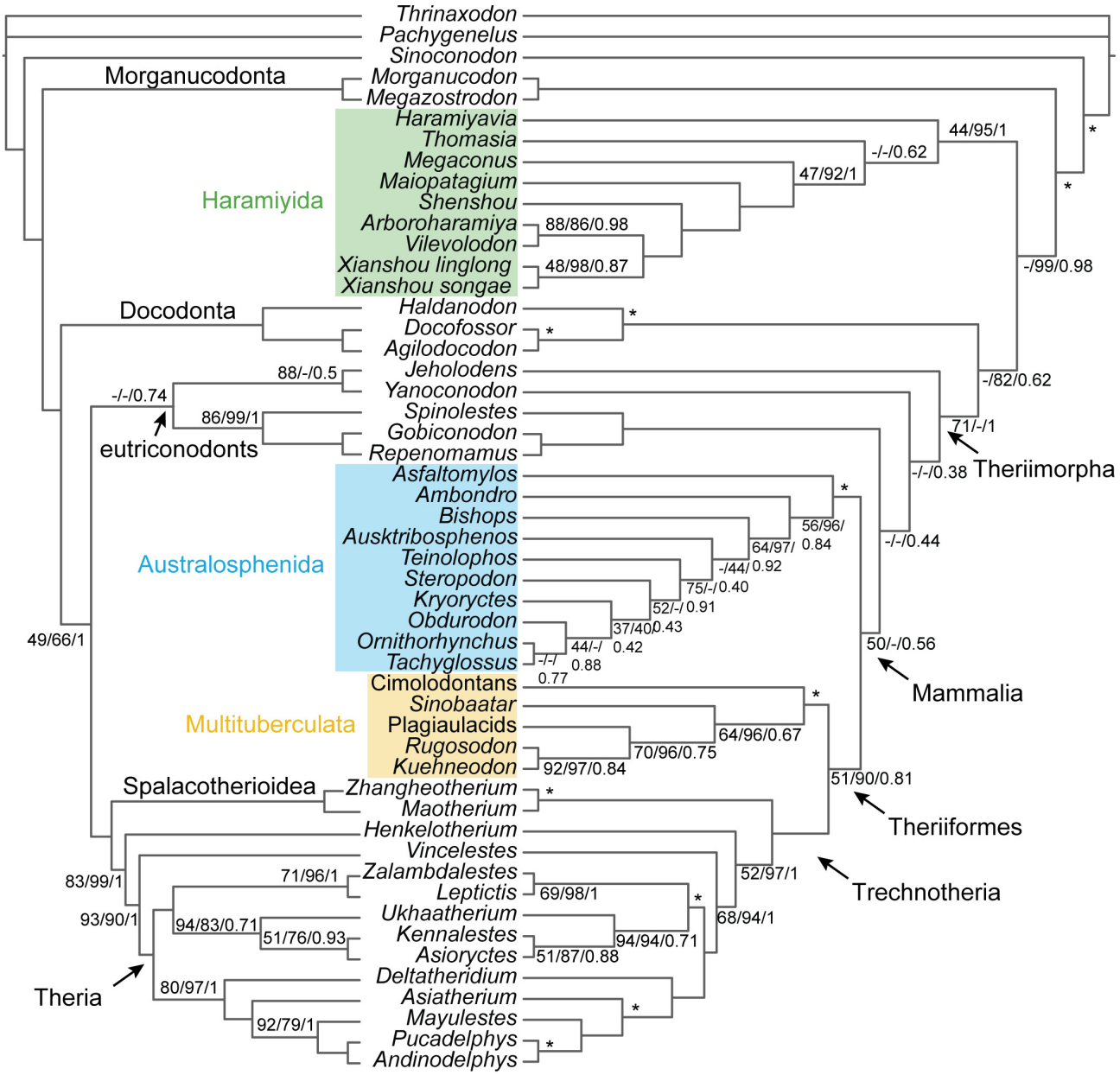


FIGURE 2 | (A) $NCDP_{537}$ and **(B)** 51_{537} phylogenies. The Bayesian inference topologies are shown, but with support values provided respectively for maximum parsimony-ordered (bootstrap), maximum likelihood (bootstrap), and Bayesian interference (Bayesian posterior probability) only for nodes at which at least one of these measures is < 95%. The asterisks indicate the constrained nodes (see **Supplementary Material** for the constraint used and **Supplementary Figure S1** for the results from the unconstrained analyses). The dashes represent the branches which are not supported.

inference, but with only 44–50% MP bootstrap support (**Figure 2B**). An important difference here may be that the ML and the Bayesian analyses partition the data by anatomical regions and allow evolutionary rates to vary across characters, thus effectively conferring greater weight to the influence of slower evolving characters. These more conserved characters typically retain more phylogenetic signal relative to non-phylogenetic signals, at least at deeper divergences (Philippe et al., 2000).

Including the dietarily apomorphic taxa with the highly complete plesiomorphic taxa substantially eroded the previously strong support for branching orders along the backbone of the tree, from crown Theria stemwards (**Figures 2A,B**). The resulting 51_{537} tree contains what is essentially a six-taxon polytomy

TABLE 1 | Kishino–Hasegawa tests in IQ-TREE on the individual partition datasets (mandibulodental, cranial, and postcranial), assessing the congruence between the topology favored for each partition and the total evidence topology for $NCDP_{537}$, 51_{537}, or 51_{350}.

	$NCDP_{537}$		51_{537}		51_{350}	
Data used	*P*-value	Δln*L*	*P*-value	Δln*L*	*P*-value	Δln*L*
Mandibulodental	0.312	6.544	**0.0005**	82.768	0.0913	10.638
Cranial	0.138	6.1598	**0.0301**	21.766	0.0696	15.311
Postcranial	**0.0383**	7.3127	0.498	0.186	0.202	4.2128

Bold values indicate the rejection ($P < 0.05$) of the $NCDP_{537}$, 51_{537}, or 51_{350} topologies on the mandibulodental, cranial, or postcranial characters.

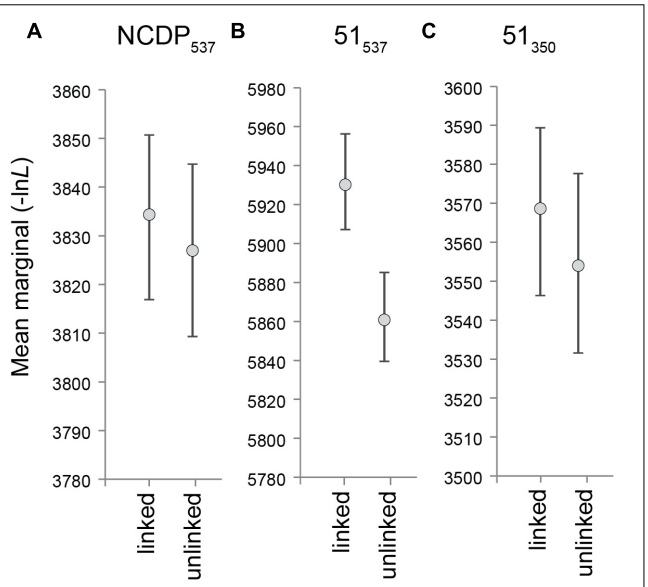


FIGURE 3 | Bayesian inference marginal likelihood means and 95% higher posterior densities (averaged over two runs) for **(A)** $NCDP_{537}$, **(B)** 51_{537}, and **(C)** 51_{350}, with the tree topology linked and unlinked across the three anatomical regions. ln*L* advantage = $(M_{Llinked} - M_{Lunlinked})/M_{Lunlinked}$.

that includes Multituberculata, Trechnotheria, Australosphenida, and three eutriconodont lineages, *Yanoconodon*, *Jeholodens*, and gobiconodontids. However, Theriiformes (Multituberculata and Trechnotheria) was recovered at >80% support by ML and Bayesian analyses of the 51_{537} data (**Figure 2B** and **Supplementary Figure S1**). Paraphyly of eutriconodonts was favored, with Gobiconodontidae, *Yanoconodon*, and *Jeholodens* falling successively deeper. Australosphenida was recovered as sister to Theriiformes (Multituberculata and Trechnotheria), albeit with weak support (34–50% BP_{MP}, 32% BP_{ML}, and 0.56–0.77 BPP). One further instability in the tree concerns the sister relationship to mammals (including eutriconodonts), which alternated between Docodonta or Haramiyida.

Testing Phylogenetic Congruence Between Anatomical Regions

We investigated whether phylogenetic uncertainty in the affinities of the apomorphic australosphenidans, haramiyidans, and multituberculates relative to the 27 near-complete dietarily plesiomorphic taxa is substantially contributed to by phylogenetic incongruence between the mandibulodental, cranial, and postcranial anatomical regions. We found only a minor evidence of incongruence among the major $NCDP_{537}$ clades. KH testing indicated that the topology of the overall (combined data) tree was not rejected in ML analyses of either the mandibulodental or cranial datasets but was rejected with the postcranial data ($P = 0.0383$, **Table 1**). Even in this case, the topological difference between the overall and the postcranial $NCDP_{537}$ ML trees was minor, with the placement of *Yanoconodon* among eutriconodonts differing by a single branch step. Moreover, the analysis of $NCDP_{537}$ in MrBayes provides widely overlapping likelihood 95% HPDs for treatments regardless of whether the topologies are linked or allowed to differ between the three anatomical region partitions (**Figure 3A**). This indicates close congruence between trees inferred from the different regional partitions for the $NCDP_{537}$ data.

The inclusion of australosphenidans, haramiyidans, and multituberculates with the near-complete dietarily plesiomorphic taxa (51_{537}) dramatically increased the incongruence among partitions. This is particularly salient for the Bayesian inference analyses in which the regional partitions are linked *versus* unlinked, with the 95% HPDs for the likelihoods becoming widely separated (**Figure 3B**). In KH testing, the overall (combined data) 51_{537} ML topology was rejected on both the mandibulodental ($P = 0.0005$) and the cranial data ($P = 0.0301$). The overall 51_{537} ML tree is not rejected on the postcranial data ($P = 0.4980$). However, the postcranial 51_{537} ML tree is rejected with both the mandibulodental and the cranial datasets ($P < 0.0001$). The strong phylogenetic incongruence between the anatomical regions is predominantly attributable to their alternative placements of the apomorphic australosphenidans, haramiyidans, and multituberculates. This is indicated by KH testing, providing almost identical results even without the backbone constraint being applied (see **Supplementary Table S4**).

The topological manifestations of incongruence between the overall 51_{537} ML tree and regional trees for the placements of the dietarily apomorphic taxa on the NCDP backbone are shown in **Figure 4**. On the overall 51_{537} ML tree (**Figure 4A**), Haramiyida is strongly excluded from crown Mammalia (>99% BP). Similar haramiyidan placements (adjacent to Docodonta) are favored for the cranial region and for the postcranial region (albeit with australosphenidans deeper in the postcranial tree; **Figure 4D**). However, for the mandibulodental region, haramiyidans and multituberculates group together (99% BP, **Figure 4B**). Indeed when reciprocal monophyly was not enforced for these two multicuspate orders, haramiyidans were paraphyletic on the mandibulodental tree, with multituberculates sister to eleutherodontid haramiyidans.

Two australosphenidan placements are similarly likely on the overall 51_{537} dataset, as sister to Theriimorpha, or one step closer to therians, as sister to Theriiformes. However, the ML trees in **Figure 4** show that the individual regions favor widely separated placements for Australosphenida. On the mandibulodental data

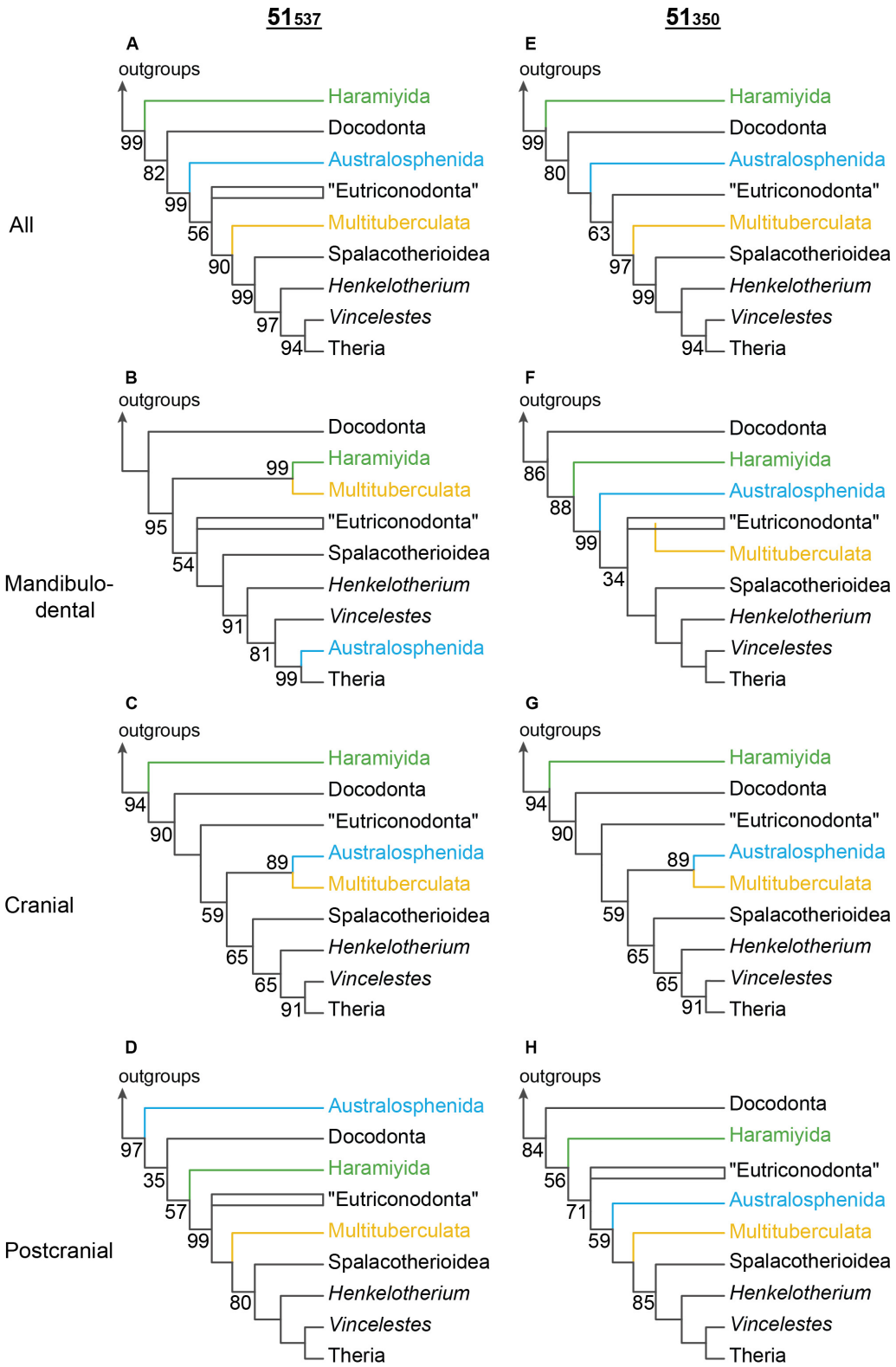


FIGURE 4 | Alternative phylogenetic positions for Haramiyida (green), Australosphenida (blue), and Multituberculata (orange) for the 51_{537} and the 51_{350} datasets on **(A,E)** all characters combined, and separately **(B,F)** mandibulodental, **(C,G)** cranial, and **(D,H)** postcranial. These groups were constrained to be monophyletic, given that some taxa that are informative for the placement of the group have too few characters for one or more anatomical regions. The maximum likelihood bootstrap (BP_{ML}) is represented only at nodes < 100%. The cranial trees are identical, since the same data is employed for this partition for both 51_{537} and 51_{350}.

(**Figure 4B**), australosphenidans are sister to crown Theria (99% BP), and on the cranial data (**Figure 4C**) they group with multituberculates (89% BP). On the postcranial data, australosphenidans diverge from very deep in the tree, outside a weakly supported (35% BP) clade that includes Theriimorpha, Haramiyida, and Docodonta (**Figure 4D**).

The regional incongruence for the placement of australosphenidans does not depend on the inclusion of the enigmatic haramiyidans and multituberculates. We compared alternative hypotheses for the placement of australosphenidans on the NCDP backbone phylogeny (without multituberculates and haramiyidans). Maximum likelihood analyses on the 537-character matrix (**Figure 5B**) favored Australosphenida being placed (1) as sister to crown therians on the mandibulodental data, (2) as sister to Trechnotheria on the cranial data and, (3) far deeper on the postcranial data, as sister to Theriimorpha plus docodonts. Notably, the favored placement of Australosphenida as sister to Trechnotheria on the full 537-character matrix is rejected at $P < 0.05$ on both the mandibulodental and the postcranial data.

Investigating Finer-Scale Regional Homoplasy

We investigated correlated homoplasy induced by the inclusion of the apomorphic clades at a finer anatomical scale among the 537-character matrix. MP disadvantage was calculated for each of the 10 sub-regions as the percentage difference between the MP tree score on that sub-region data compared to the MP score on the same sub-region data, but with the tree constrained to the total evidence topology favored for the overall 537-character dataset.

The highest MP disadvantages were attributable to the shoulder girdle (20.0%) and mandibular (12.0%) sub-regions, and the lowest MP disadvantages were attributable to the basicranial (4.4%) and the hindlimb (4.0%) sub-regions (**Figure 6A**). A power curve was fitted to control MP disadvantage for the number of characters in each sub-region (**Figure 6A**). The resulting (inferred/expected) "corrected" MP disadvantage ratios are highest for the shoulder girdle (1.74) and the cheek teeth (1.66) and are lowest for the hindlimb (0.55) and "other dental" (0.75) sub-regions (**Figure 6B**). All other sub-regions with sufficient character sampling had corrected MP disadvantage ratios between 0.79 and 1.24, close to the expected value of 1. Given this finding, we revisited our primary phylogenetic and incongruence analyses with the cheek teeth and the shoulder girdle characters excluded, leaving a 350-character dataset (51_{350}).

Correlated Homoplasy Reduction

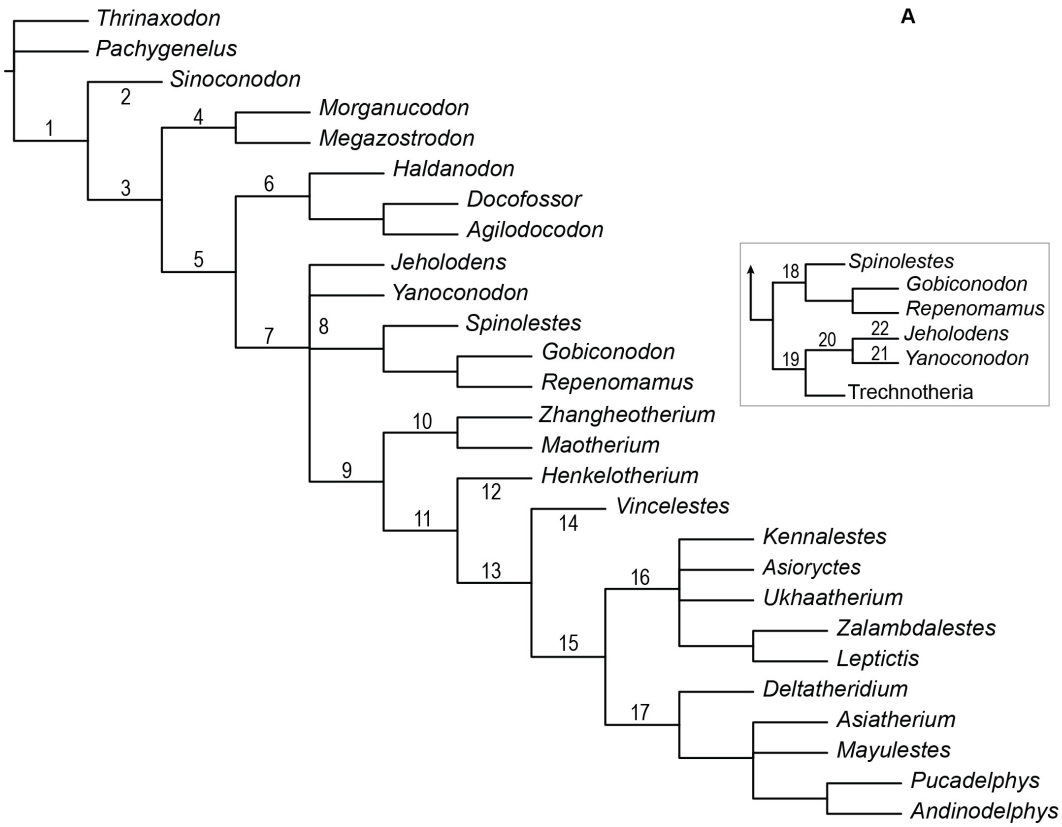
Excluding the cheek teeth and the shoulder girdle characters in the MP analysis of the 51_{350} dataset barely reduced the overall homoplasy index (HI) from 0.52 to 0.49. However, the objective was more specifically to exclude major sources of correlated homoplasy predicted to affect the placements of Australosphenida, Haramiyida, and Multituberculata. Bayesian inference analyses showed a dramatic improvement in congruence among the mandibulodental, cranial, and postcranial region partitions. This is clearly shown in **Figure 3**, where 95% HPDs for tree likelihood are compared between analyses with topology linked or unlinked (free to vary across partitions). The HPDs are widely separated for the 51_{537} dataset (**Figure 3B**), indicating a significant incongruence between regional phylogenetic signals. With the cheek teeth and the shoulder girdle characters excluded, the 95% HPDs for tree likelihood are brought into a wide overlap for the 51_{350} dataset (**Figure 3C**).

Maximum likelihood analyses of the 51_{350} dataset (which excludes the correlated-homoplastic cheek teeth and shoulder-girdle characters) yield notably different results from those of the 51_{537} dataset (**Figure 4**). For the mandibulodental partition, the multituberculate–haramiyidan and australosphenidan–crown therian groupings recovered with the 51_{537} data (**Figure 4B**) are both rejected with the 51_{350} data (**Figure 4F**). The placements recovered with the 51_{350} dataset (**Figure 4F**) are instead more consistent with analyses of the full skeletal dataset: multituberculates within Theriimorpha and Australosphenida were excluded from Theriimorpha. Exclusion of the shoulder girdle characters also brought the 51_{537} postcranial placement of australosphenidans (previously outside Theriimorpha and Docodonta, **Figure 4D**) into closer agreement with the overall tree, as sister to Theriiformes (**Figure 4H**).

With the regional partitions combined (51_{350}), but modeled separately, the exclusion of the cheek teeth and the shoulder girdle characters resulted in the same (51_{537}) placement of haramiyidans—falling outside of Mammalia (**Figures 4A,E**). However, 51_{350} provided increased confidence for grouping multituberculates with trechnotheres (BP_{ML} from 90 to 97%; BPP from 0.81 to 0.99). Moreover, the placements of Australosphenida as sister to Theriiformes or Theriimorpha (Theriiformes plus eutriconodonts) are not rejected for either 51_{537} or 51_{350}, although the exclusion of the shoulder and the cheek teeth characters shifts support further in favor of excluding australosphenidans from Theriimorpha.

Comparisons of alternative ML placements of australosphenidans alone on the NCDP backbone are particularly instructive for the mandibulodental and the postcranial regions. With the full character set (51_{537}), the accepted placements (not rejected at $P = 0.05$) on these two regional datasets are widely separated. In contrast, upon the exclusion of the cheek teeth and the shoulder girdle characters, the (51_{350}) mandibulodental and postcranial partitions both favor the same placement for australosphenidans (as sister to Theriimorpha, **Figure 5C**). This placement of australosphenidans as sister to Theriimorpha was also favored on the combined 51_{350} data and was congruent with the cranial data.

The loss of statistical power for resolving australosphenidan affinities with the mandibulodental data after the cheek teeth characters were excluded (**Figure 5C**) does not translate as diminished resolution for 51_{350} when the regions are combined. All australosphenidan placements within Trechnotheria or deeper than the theriimorph stem are rejected in IQ-TREE KH-testing at $P < 0.05$ for both the 537- and 350-character datasets.



	B 537				C 350			
Tree	**All**	**MD**	**Cranial**	**Postcranial**	**All**	**MD**	**Cranial**	**Postcranial**
1	0.00135	5.33E-06	1.85E-47	0.00719	0.0173	0.0909	9.48E-64	8.39E-05
2	0.00147	0.00167	8.51E-33	0.0063	0.0118	0.0568	5.72E-54	0.000765
3	0.00146	4.26E-06	0.000626	0.0279	0.00218	0.191	0.000118	0.173
4	0.000352	0.000992	0.000544	0.002091	0.00216	0.193	9.42E-05	0.000156
5	0.000885	0.00247	0.0022	**0.97**	0.00245	0.466	0.00058	1.72E-12
6	0.0143	0.00165	0.008	0.0552	0.00718	0.494	0.0062	0.367
7	0.55	1.65E-06	0.564	0.0963	**0.742**	**0.734**	0.532	**0.849**
8								
9	**0.584**	0.0346	**0.843**	6.68E-65	0.401	0.36	**0.852**	0.391
10	5.98E-40	0.0276	0.118	5.14E-05	0.000532	0.147	0.142	2.64E-71
11	0.00689	0.127	0.375	7.95E-06	4.03E-05	0.512	0.39	1.04E-83
12	2.80E-06	0.0257	0.0974	8.42E-06	0.000484	0.24	0.133	1.02E-55
13	0.0158	0.14	0.146	4.17E-05	0.000871	0.238	0.172	4.50E-07
14	5.83E-08	0.127	0.00855	1.57E-20	0.000167	0.0722	0.0165	1.81E-44
15	0.019	**0.904**	0.0254	5.91E-06	5.06E-07	0.118	0.0166	1.19E-08
16	0.000112	0.354	0.00251	7.78E-05	2.48E-65	0.489	7.29E-06	1.74E-08
17	0.000325	0.105	0.0025	6.58E-06	4.46E-65	0.119	2.43E-07	8.68E-07
18	6.90E-05	8.65E-05	0.312	0.000264	0.125	0.164	0.119	0.335
19	4.19E-05	0.000773	0.503	6.48E-05	0.0543	0.162	0.417	6.19E-32
20	0.00276	0.000547	0.161	6.44E-06	0.00262	0.164	0.295	7.52E-06
21	1.64E-05	0.00224	0.133	0.0013	7.21E-05	0.031	0.116	0.0503
22	5.53E-07	0.00167	0.136	0.00106	3.78E-05	0.0348	0.116	0.0176

FIGURE 5 | Monotreme placement. Kishino–Hasegawa tests (in IQ-TREE) for the alternative placements of Australosphenida on **(A)** the NCDP mammal backbone constraint phylogeny for the mandibulodental, cranial, and postcranial data for **(B)** the full character set and **(C)** excluding cheek teeth and shoulder girdle characters. The placement of Australosphenida is rejected in red ($P < 0.05$), not rejected in green ($P > 0.1$), and weakly rejected when not highlighted ($P = 0.05$–0.1). Placement 8 corresponds to the placement of Australosphenida within or adjacent to eutriconodonts and is inclusive of placements 7 and 8, and any placements with *Jeholodens* or *Yanoconodon*.

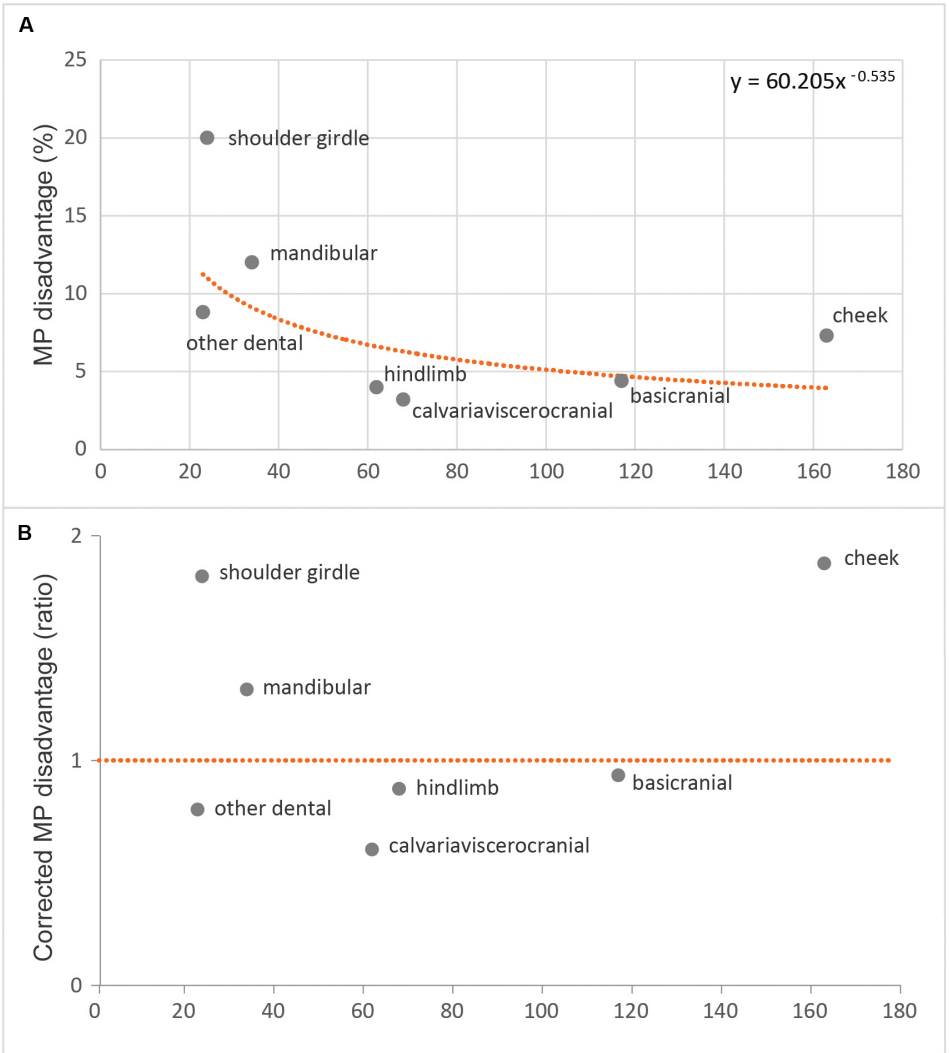


FIGURE 6 | Maximum parsimony disadvantage for each sub-region **(A)** expressed as a percentage and regressed across sub-regions as a power curve and **(B)** corrected maximum parsimony disadvantage for those same values, but compared as a ratio relative to their expected values from the power curve regression.

With the addition of multituberculates, an Australosphenida–Multituberculata grouping was rejected initially at $P = 0.218$, but with this result strengthening slightly to $P = 0.158$ following the exclusion of the cheek teeth and the shoulder girdle characters. Rejection of the Allotheria (Haramiyida–Multituberculata) hypothesis is a far stronger result based on the 51_{350} data ($P = 0.001$) compared with 51_{537} data ($P = 0.044$).

Extension to Less-Complete Taxa

Character incompleteness within the more taxonomically inclusive 78_{350} dataset invalidates the further examination of incongruence among regions. However, the increased taxon sampling confirms and generally enhances statistical support for the placements of australosphenidans, multituberculates, and haramiyidans when compared with the 51_{350} dataset (**Figure 7**, cf. **Figure 4E**) and recovers a broadly similar topology to that of Huttenlocker et al. (2018). The additional taxon sampling and exclusion of the sub-regions contributing disproportionately high levels of correlated homoplasy also strengthened the support for grouping Hahnodontidae and the Gondwanathere, *Vintana*, as well as for their position within Haramiyida as sister to Eleutherodontidae (98% BP_{ML} and 0.95 BPP).

The 78_{350} data (**Figure 7**) allows us to examine the phylogenetic implications of excluding the cheek teeth and the shoulder girdle characters for several less complete taxa that are nevertheless important for reconstructing mammalian evolution. Of particular note, the oldest proposed eutherian, *Juramaia* (Luo et al., 2011), was excluded from crown Theria (73% BP_{ML} and 0.77 BPP). The proposed sister taxon of Australosphenida, the Shuotheriidae [see Luo et al. (2007)], was instead placed as sister to Mammalia, although the group composed of Australosphenida, *Fruitafossor*, and Theriimorpha, which define Mammalia in this case, is not confidently resolved (57% BP_{ML} and 0.94 BPP). Analyses on this extended 78_{350} dataset also retrieved a stronger support for placing Australosphenida (and

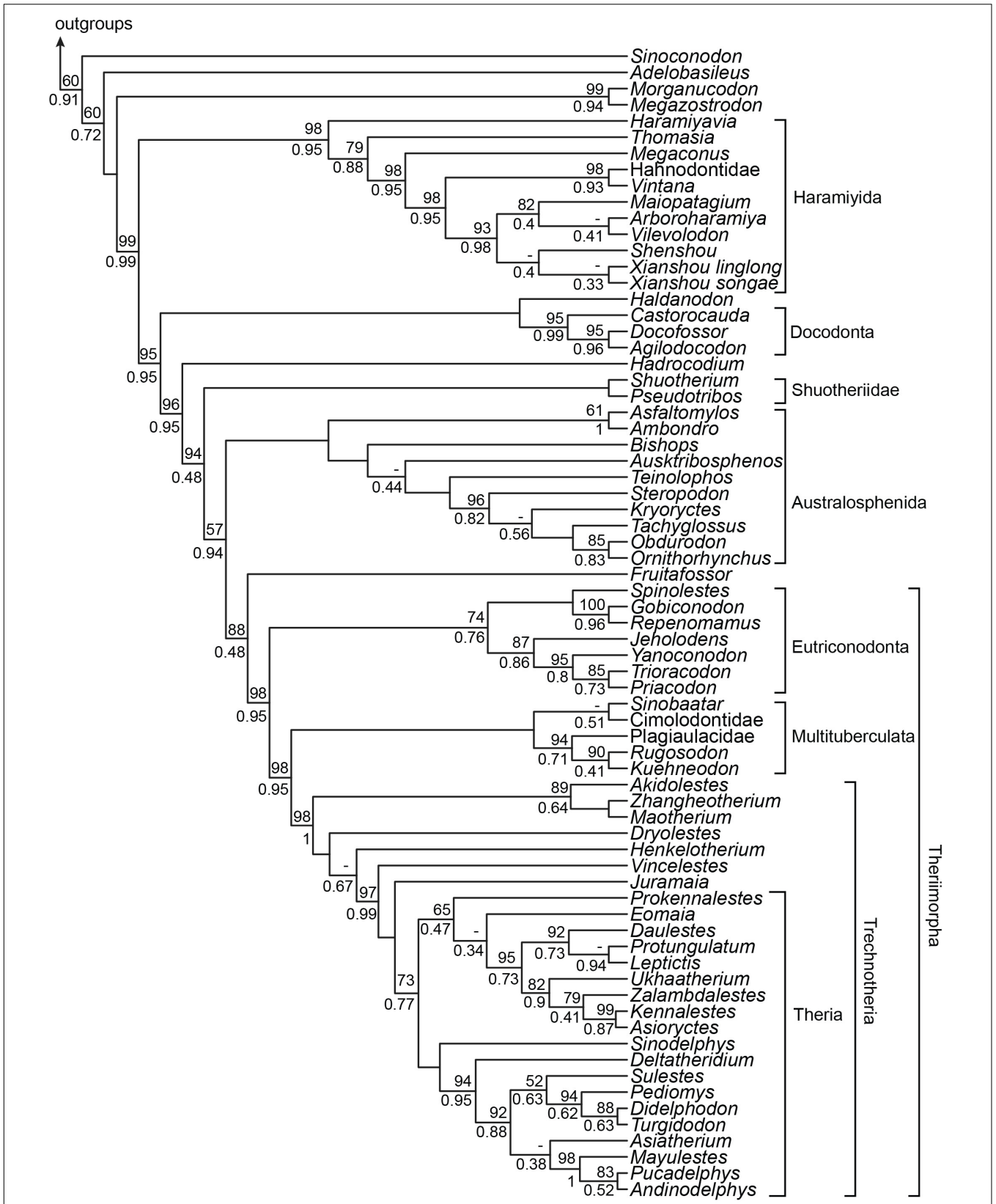


FIGURE 7 | Bayesian tree based on the 78_{350} dataset. The Bayesian posterior probability (BPP) and maximum likelihood bootstrap (BP_{ML}) are represented only at nodes where BPP < 1 and $BP_{ML} < 100\%$. The dashes represent branches which are not supported in the maximum likelihood analysis.

Fruitafossor) outside of Theriimorpha (98% BP_{ML} and 0.95 BPP) and for eutriconodont monophyly (74% BP_{ML} and 0.76 BPP).

DISCUSSION

Phylogenetic Incongruence Between Anatomical Regions for Ecologically Apomorphic Groups

Complex organisms are able to efficiently evolve through fitness landscapes because developmental modularity results in correlated evolution among characters (Lande and Arnold, 1983; Wagner, 1996; Kemp, 2007). Characters can be functionally and/or developmentally linked by genetic pleiotropy, heterochrony (*e.g.*, paedomorphosis), allometric lines of least resistance, and adaptive covariance. The linked characters are therefore unlikely to evolve fully independently. Such evolutionarily correlated character complexes tend to inflate both true phylogenetic signals (*i.e.*, due to shared ancestry) and non-phylogenetic signals (homoplasy). However, taxa that are more developmentally or functionally divergent are of course more likely to express these correlations as non-phylogenetic signals. Analogous phylogenetic biases are well studied for molecular data, for example, with DNA base composition (Sueoka, 1995; Phillips and Penny, 2003; Gibson et al., 2004), but this issue remains underexplored with morphology. However, from this theoretical expectation, we hypothesized that monotremes and the dietarily apomorphic (omnivorous–herbivorous) multituberculates and haramiyidans would contribute more substantially to phylogenetic incongruence among anatomical regions than the generalized insectivore/carnivore mammalian backbone phylogeny. As anticipated, there is broad phylogenetic agreement between mandibulodental, cranial, and postcranial partitions among the generalized insectivore/carnivores, but the inclusion of ecologically apomorphic clades of australosphenidans (including monotremes), multituberculates, and haramiyidans leads to a significant incongruence in ML hypothesis testing (**Table 1**) and a significant likelihood advantage for unlinking topologies for anatomical region partitions in Bayesian inference (**Figure 3B**).

Three regional relationships in **Figure 4** stand out as being several tree steps distant from their full-skeletal affinities. These are the mandibulodental groupings of (1) haramiyidans with multituberculates (Allotheria) and (2) australosphenidans with crown therians, as well as (3) the deep placement of australosphenidans on the postcranial data. While previous research has not quantified the incongruence induced by ecologically apomorphic Mesozoic mammals, the patterns are consistent with previous observations. The potential for dental and jaw geometry convergence between the multicuspate, omnivorous–herbivorous "allotherians" has been robustly discussed (*e.g.*, Jenkins et al., 1997; Butler, 2000; Zhou et al., 2013; Zheng et al., 2013; Bi et al., 2014; Luo et al., 2015). Also, early suggestions of close homology between the molar morphology of australosphenidans (or stem monotremes) and therians (*e.g.*, Archer et al., 1985; Rich et al., 1997) are less probable in view of more recent fossil finds, and the independent evolution of near-tribosphenic molars in therians and australosphenidans is well founded in form/function cusp pattern models (*e.g.*, Luo et al., 2001; Rougier et al., 2007; Davis, 2011).

Postcranial character complexes can have homoplasies (*e.g.*, Ji et al., 1999, Meng et al., 2017), and widespread postcranial homosplasy has been suggested for monotremes (Gregory, 1947). Shoulder girdle morphology and some upper appendicular features of monotremes appear to be superficially similar to those of ancient cynodonts (*e.g.*, Macintyre, 1967; Gambaryan et al., 2015). Such similarities are homoplastic when mapped on phylogeny as monotremes (and other crown mammals) are widely separated from non-mammaliaform cynodonts, according to modern phylogenetic interpretations that consider a wide range of new Mesozoic mammaliaforms that were unknown 50 years ago (**Figure 4**). The historical argument for monotremes to be related to some cynodonts (Macintyre, 1967) is also inconsistent with relaxed-clock molecular dating for the divergence of monotremes and therians (*e.g.*, Kullberg et al., 2008; Meredith et al., 2011; Phillips, 2015). Developmental studies on the monotreme and the marsupial shoulder girdles (*e.g.*, Klima, 1973) leave open the possibility that the monotreme condition is partly paedomorphic, while the extreme humeral long-axis rotation emphasis in monotremes associated with fossorial/swimming activity offers functional arguments for an evolutionary reversal upon earlier cynodont conditions (Phillips et al., 2009). In addition, the *Kryoryctes* humerus (Pridmore et al., 2005) suggests reversal from trochlea-like to condylar ulnar articulation. A somewhat similar transformation (at least with a shallow trochlea condition) may also have occurred in marsupial moles [noting the Miocene and modern species figured in Archer et al., 2011)]. Developmental genomics may offer a pathway to testing the hypotheses of homology *versus* homoplasy involving monotremes and other extant mammals (or reptiles), as has been done for female reproductive organs (Wagner and Lynch, 2005).

It is encouraging that our MP disadvantage analyses identified the same two anatomical character complexes (cheek teeth and shoulder girdle) that had prior expectations (*e.g.*, Kangas et al., 2004; Phillips et al., 2009; Ramírez-Chaves et al., 2016) for substantial correlated homoplasy upon the inclusion of australosphenidans, multituberculates, and haramiyidans with the generalized insectivore/carnivores. It is not obvious from standard homoplasy metrics, such as the HI, that either the cheek teeth or the shoulder girdle is especially problematic; of the 10 sub-regions, they are ranked only fifth and seventh for HI (**Supplementary Table S3**). HI may be a poorer indicator of phylogenetic inaccuracy. If homoplastic transformations are randomly distributed, then increasing homoplasy may reduce phylogenetic resolution but may not manifest as statistical inconsistency. MP disadvantage instead considers the difference between MP scores on a given partition, for when the tree is unconstrained, compared to when the tree is constrained to the topology inferred from all partitions. The more phylogenetically correlated the homoplasy, the larger the MP disadvantage. Beyond the present study, it will be important to further explore the relationship between MP disadvantage and the relative size of the partition to the overall dataset (which we correct for in

Figure 6B). Resampling procedures may also permit confidence intervals for MP disadvantage.

Mesozoic Mammal Phylogeny and the Constituency of the Crown Group

The relationships between the major groups of generalized insectivore/carnivores that were included in the near-completely sampled, dietarily plesiomorphic dataset ($NCDP_{537}$) are well established (*e.g.*, Kielan-Jaworowska et al., 2004; Rougier et al., 2011; Martin et al., 2015). Here we confirm strong support for their branching order, from non-mammaliaform cynodonts right through to eutherians and metatherians (**Figure 2A**)—with or without the cheek teeth and the shoulder girdle characters and with extended taxon sampling (**Figure 7**). However, by investigating phylogenetic signal among broad-scale anatomical regions (**Figures 4A–D**), we show that the affinities of three dietarily apomorphic clades, multituberculates, haramiyidans, and australosphenidans (including monotremes), are less stable than total evidence approaches might imply.

There is a common argument that combining all relevant evidences provides the best phylogenetic estimator (Asher, 2007; Mounce et al., 2016). Similar to advocacy for parsimony over likelihood, this simpler solution may well be appropriate when underlying evolutionary processes are largely unknown (Kolaczkowski and Thornton, 2004). However, our interrogation of phylogenetic signal variation across partitions reveals extreme incongruence (**Figures 3B**, **4B–D**, **5**). To ignore phylogenetic incongruence in favor of the total evidence approach to combine the incongruent data partitions is akin to ignoring interaction effects in a statistical analysis of variance. Instead we have identified elevated levels of correlated homoplasy among the cheek teeth and the shoulder girdle (**Figure 6**). Excluding these data was remarkably effective for bringing a phylogenetic congruence between the mandibulodental, cranial, and postcranial partitions (**Figures 3C**, **4F–H**, **5**). Moreover, reducing the overall dataset from 537 to 350 characters did not cost precision. In fact, in some cases, phylogenetic resolution was greatly enhanced, such as for rejecting Allotheria (Multituberculata–Haramiyida) at $P = 0.001$ (51_{350}) compared with $P = 0.044$ (51_{537}). This strong result in favor of multituberculates grouping with trechnotheres and haramiyidans falling outside crown Mammalia was foreshadowed by the analysis of Ramírez-Chaves et al. (2016), with molar characters excluded. A note of caution is nevertheless warranted for the relative placements of these multicuspate taxa. Even beyond arguments for and against dental convergence, there is substantial variation and debate regarding non-dental characters among haramiyidans and multituberculates (Bi et al., 2014; Luo et al., 2015; Meng et al., 2018). Resolving character scoring and including further cranial material from undoubted haramiyidans will be important for confirming our placement of multituberculates within and haramiyidans outside Mammalia.

Our efforts to identify the affinities of Australosphenida are broadly reflective of most matrix-based analyses over the past decade, which tend to place this Gondwanan clade either as the sister of Theriiformes (trechnotheres and multituberculates) or a further step stemward, also outside of eutriconodonts. Our incongruence findings do reveal instability in this near-consensus, and an Australosphenida–Multituberculata clade is not yet convincingly rejected; however, there is progress toward resolution. In particular, exclusion of the sources of substantial correlated homoplasy (**Figures 4**, **5**) and expanded taxon sampling (**Figure 7**) both strengthen the support for australosphenidans falling outside of Theriimorpha. Thus, trechnotheres, multituberculates, and eutriconodonts would all be crown mammals. The case is weaker for the somewhat fossorial and dentally simplified *Fruitafossor* (Luo and Wible, 2005) and for the pseudotribosphenic shuotheriids (**Figure 7**). Discovery of substantial skeletal material from *Pseudotribos* (Luo et al., 2007) added weight to the earlier hypothesis of Kielan-Jaworowska et al. (2002) that shuotheriids were the sister group of australosphenidans. The primary arguments were based on tentative interpretations of molar morphology and seemingly more robust (although ambiguous) shoulder girdle synapomorphies. These are notably the same character complexes that our MP disadvantage analysis identified as the most likely to be unreliable for inferring the affinities of highly apomorphic Mesozoic mammals. Hence, we suggest that placement of shuotheriids with australosphenidans (or even within Mammalia) requires further testing, ideally on more complete cranial and postcranial material.

On balance our incongruence testing, identification of sources of elevated correlated homoplasy, and partitioned ML and Bayesian phylogenetic inference offer increased confidence in the relationships of the main Mesozoic clades of generalized insectivorous/carnivorous mammals and stronger support (albeit with caveats) for the placements of australosphenidans, multituberculates, and haramiyidans. We also identify where additional caution is required, for example, we join several other studies (Krause et al., 2014; Averianov, 2015; Sweetman et al., 2017) in questioning the placement of *Juramaia* as a eutherian. The implication for molecular dating is that more completely known crown therians from the ∼125 Ma Jehol biota may provide a safer Early Cretaceous minimum bound for calibrating the marsupial–placental divergence than does the purportedly Jurassic *Juramaia*. Current progress toward improved inference of phylogeny and dating with appropriate calibrations (Phillips and Fruciano, 2018) present an opportunity to more accurately trace the ecological ancestry of early mammals. This, in turn, can shed light on the biotic and abiotic drivers of convergent dietary and locomotory evolution among Mesozoic mammals.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding authors.

AUTHOR CONTRIBUTIONS

MP conceived the study. MC and MP compiled and analyzed the dataset and wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Australian Research Council Discovery Projects (DP150104659 and DP170103227 awarded to MP).

ACKNOWLEDGMENTS

We thank the reviewers for providing valuable comments that substantially improved the manuscript. We also thank Michel Laurin, Gilles Didier, and Rachel Warnock for their efforts in organizing the "Timetrees" symposium at IPC5.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.00651/full#supplementary-material

REFERENCES

Allin, E. F. (1975). Evolution of the mammalian middle ear. *J. Morphol.* 147, 403–438.

Archer, M., Beck, R., Gott, M., Hand, S., Godthelp, H., and Black, K. (2011). Australia's first fossil marsupial mole (Notoryctemorphia) resolves controversies about their evolution and palaeoenvironmental origins. *Proc. Biol. Sci.* 278, 1498–1506. doi: 10.1098/rspb.2010.1943

Archer, M., Flannery, T., Ritchie, A., and Molnar, R. (1985). First mesozoic mammal from Australia—an early Cretaceous monotreme. *Science* 318, 363–366. doi: 10.1038/318363a0

Archer, M., Murray, P., Hand, S., and Godthelp, H. (1993). "Reconsideration of monotreme relationships based on the skull and dentition of the Miocene Obdurodon dicksoni," in *Mammal Phylogeny*, eds F. S. Szalay, M. J. Novacek, and M. C. McKenna (New York, NY: Springer), 75–94. doi: 10.1007/978-1-4613-9249-1_7

Asher, R. J. (2007). A web-database of mammalian morphology and a reanalysis of placental phylogeny. *BMC Evo. Biol.* 7:108. doi: 10.1186/1471-2148-7-108

Averianov, A. O. (2015). Taxonomic revision of tribosphenic mammals from the lower cretaceous antlers formation of Texas and Oklahoma. USA. *Proc. Zool. Inst. RAS* 319, 141–181.

Benevento, G. L., Benson, R. B. J., and Friedman, M. (2019). Patterns of mammalian jaw ecomorphological disparity during the Mesozoic/Cenozoic transition. *Proc. R. Soc. B* 286:20190347. doi: 10.1098/rspb.2019.0347

Bi, S., Wang, Y., Guan, J., Sheng, X., and Meng, J. (2014). Three new Jurassic euharamiyidan species reinforce early divergence of mammals. *Nature* 514, 579–584. doi: 10.1038/nature13718

Brown, M. B. (1975). Exploring interaction effects in the ANOVA. *J. R. Stat. Soc. Ser. C* 24, 288–298.

Butler, P. M. (1990). Early trends in the evolution of tribosphenic molars. *Biol. Rev.* 65, 529–552. doi: 10.1111/j.1469-185x.1990.tb01237.x

Butler, P. M. (2000). Review of the early allotherian mammals. *Acta Palaeontol. Pol.* 45, 317–342.

Cardini, A., Polly, D., Dawson, R., and Milne, N. (2015). Why the long face? Kangaroos and wallabies follow the same 'rule'of cranial evolutionary allometry (CREA) as placentals. *Evol. Biol.* 42, 169–176. doi: 10.1007/s11692-015-9308-9

Chen, M., and Wilson, G. P. (2015). A multivariate approach to infer locomotor modes in Mesozoic mammals. *Paleobiology* 41, 280–312. doi: 10.1017/pab.2014.14

Chow, M., and Rich, T. H. (1982). *Shuotherium dongi*, n. gen. and sp., a therian with pseudo-tribosphenic molars from the Jurassic of Sichuan. China. *Aust. Mammal.* 5, 127–142.

Davis, B. M. (2011). Evolution of the tribosphenic molar pattern in early mammals, with comments on the "dual-origin" hypothesis. *J. Mammal. Evol.* 18:227. doi: 10.1007/s10914-011-9168-8

dos Reis, M., Inoue, J., Hasegawa, M., Asher, R. J., Donoghue, P. C. J., and Yang, Z. (2012). Phylogenomic datasets provide both precision and accuracy in estimating the timescale of placental mammal phylogeny. *Proc. R. Soc. B Biol. Sci.* 279, 3491–3500. doi: 10.1098/rspb.2012.0683

Fitzhugh, K. (2006). The 'requirement of total evidence' and its role in phylogenetic systematics. *Biol. Philos.* 21, 309–351. doi: 10.1007/s10539-005-7325-2

Flynn, J. J., Parrish, J. M., Rakotosamimanana, B., Simpson, W. F., and Wyss, A. R. (1999). A middle Jurassic mammal from Madagascar. *Nature* 401, 57–60. doi: 10.1038/43420

Gambaryan, P. P., Kuznetsov, A. N., Panyutina, A. A., and Gerasimov, S. V. (2015). Shoulder girdle and forelimb myology of extant monotremata. *Russ. J. Theriol.* 14, 1–56. doi: 10.15298/rusjtheriol.14.1.01

Gibson, A., Gowri-Shankar, V., Higgs, P. G., and Rattray, M. (2004). A comprehensive analysis of mammalian mitochondrial genome base composition and improved phylogenetic methods. *Mol. Biol. Evol.* 22, 251–264. doi: 10.1093/molbev/msi012

Goswami, A., Smaers, J. B., Soligo, C., and Polly, P. D. (2014). The macroevolutionary consequences of phenotypic integration: from development to deep time. *Philos. Trans. R. Soc. B Biol. Sci.* 369, 20130254. doi: 10.1098/rstb.2013.0254

Goswami, A., Weisbecker, V., and Sánchez-Villagra, M. R. (2009). Developmental modularity and the marsupial-placental dichotomy. *J. Exp. Zool. Part BMol. Dev. Evol.* 312, 186–195. doi: 10.1002/jez.b.21283

Gregory, W. K. (1947). The monotremes and the palimpsest theory. *Bull. Am. Museum Nat. Hist.* 88, 1–52.

Grossnickle, D. M. (2017). The evolutionary origin of jaw yaw in mammals. *Sci. Rep.* 7:45094.

Hendy, M., and Penny, D. (1989). A framework for the quantitative study of phylogenetic data. *Syst. Zool.* 38, 297–309.

Huelsenbeck, J. P., and Ronquist, F. (2001). MRBAYES: bayesian inference of phylogenetic trees. *Bioinformatics* 17, 754–755. doi: 10.1093/bioinformatics/17.8.754

Huttenlocker, A. K., Grossnickle, D. M., Kirkland, J. I., Schultz, J. A., and Luo, Z. X. (2018). Late-surviving stem mammal links the lowermost cretaceous of North America and Gondwana. *Nature* 558, 108–123.

Jenkins, F. A. Jr., and Parrington, F. R. (1976). The postcranial skeletons of the Triassic mammals Eozostrodon, Megazostrodon and Erythrotherium. *Philos. Trans. R. Soc. Lon. B* 273, 387–431. doi: 10.1098/rstb.1976.0022

Jenkins, F. A. (1970). Limb movements in a monotreme (*Tachyglossus aculeatus*): a cineradiographic analysis. *Science* 168, 1473–1475. doi: 10.1126/science.168.3938.1473

Jenkins, F. A., Gatesy, S. M., Shubin, N. H., and Amaral, W. W. (1997). Haramiyids and Triassic mammalian evolution. *Nature* 385, 715–718. doi: 10.1038/385715a0

Ji, Q., Luo, Z., and Ji, S. A. (1999). A Chinese triconodont mammal and mosaic evolution of the mammalian skeleton. *Nature* 398, 326–330. doi: 10.1038/18665

Ji, Q., Luo, Z. X., Yuan, C. X., and Tabrum, A. R. (2006). A swimming mammaliaform from the middle jurassic and ecomorphological diversification of early mammals. *Science* 311, 1123–1127. doi: 10.1126/science.1123026

Kangas, A. T., Evans, A. R., Thesleff, I., and Jernvall, J. (2004). Nonindependence of mammalian dental characters. *Nature* 432, 11–14.

Kemp, T. S. (1983). The relationships of mammals. *Zool. J. Linn. Soc.* 77, 353–384.

Kemp, T. S. (2007). The origin of higher taxa: macroevolutionary processes, and the case of the mammals. *Acta Zool.* 88, 3–22. doi: 10.1111/j.1463-6395.2007.00248.x

Kermack, D. M., and Kermack, K. A. (1984). *The Evolution of Mammalian Characters*. Sydney: Croom Helm.

Kielan-Jaworowska, Z., Cifelli, R. L., and Luo, Z. (2002). Dentition and relationships of the Jurassic mammal Shuotherium. *Acta Palaeontol. Pol.* 47, 479–486.

Kielan-Jaworowska, Z., Cifelli, R. L., and Luo, Z.-X. (2004). *Mammals From the Age of Dinosaurs: Origins, Evolution, and Structure*. New York, NY: Columbia University Press.

Klima, M. (1973). Die Frühentwicklung des Schültergürtels und des Brustbeins bei den Monotremen (Mammalia: Prototheria). *Adv. Anat. Embryol. Cell Biol.* 47, 1–80. doi: 10.1007/978-3-662-06649-2

Kluge, A. G. (1989). A concern for evidence and a phylogenetic hypothesis of relationships among Epicrates (*Boidae*. *Serpentes*). *Syst. Biol.* 38, 7–25. doi: 10.1093/sysbio/38.1.7

Kolaczkowski, B., and Thornton, J. W. (2004). Performance of maximum parsimony and likelihood phylogenetics when evolution is heterogeneous. *Nature* 431, 980–984. doi: 10.1038/nature02917

Krause, D. W., Hoffmann, S., Wible, J. R., Kirk, E. C., Schultz, J. A., von Koenigswald, W., et al. (2014). First cranial remains of a gondwanatherian mammal reveal remarkable mosaicism. *Nature* 515, 512–517. doi: 10.1038/nature13922

Krebs, B. (1991). *Das skelett von Henkelotherium Guimarotae gen. et sp. nov. (Eupantotheria, Mammalia) aus dem Oberen Jura von Portugal*. Berlin: Selbstverlag Fachbereich Geowissenschaften. Series: Berliner geowissenschaftliche Abhandlungen Bd 133, 1–110.

Kühne, W. G. (1973). The systematic position of monotremes reconsidered (*Mammalia*). *Z. Morphol. Tiere* 75, 59–64. doi: 10.1007/bf00723669

Kühne, W. G. (1977). On the Marsupionta, a reply to Dr. Parrington. *J. Nat. Hist.* 11, 225–228. doi: 10.1080/00222937700770141

Kullberg, M., Hallström, B. M., Arnason, U., and Janke, A. (2008). Phylogenetic analysis of 1.5 *Mbp* and platypus EST data refute the Marsupionta hypothesis and unequivocally support Monotremata as sister group to Marsupialia/Placentalia. *Zool. Scr.* 37, 115–127. doi: 10.1111/j.1463-6409.2007.00319.x

Lande, R., and Arnold, S. J. (1983). The measurement of selection on correlated characters. *Evolution* 37, 1210–1226. doi: 10.1111/j.1558-5646.1983.tb00236.x

Lartillot, N., and Poujol, R. (2011). A phylogenetic model for investigating correlated evolution of substitution rates and continuous phenotypic characters. *Mol. Biol. Evol.* 28, 729–744. doi: 10.1093/molbev/msq244

Lewis, P. O. (2001). A likelihood approach to estimating phylogeny from discrete morphological character data. *Syst. Biol.* 50, 913–925. doi: 10.1080/106351501753462876

Liu, Y., Cotton, J. A., Shen, B., Han, X., Rossiter, S. J., and Zhang, S. (2010). Convergent sequence evolution between echolocating bats and dolphins. *Curr. Biol.* 20, R53–R54.

Luo, Z. X. (2007). Transformation and diversification in early mammal evolution. *Nature* 450, 1011–1019. doi: 10.1038/nature06277

Luo, Z.-X. (2011). Developmental patterns in Mesozoic evolution of mammal ears. *Annu. Rev. Ecol. Evol. Syst.* 42, 355–380. doi: 10.1146/annurev-ecolsys-032511-142302

Luo, Z.-X., Cifelli, R. C., and Kielan-Jaworowska, Z. (2001). Dual origin of tribosphenic mammals. *Nature* 409, 53–57. doi: 10.1038/35051023

Luo, Z.-X., Gatesy, S. M., Jenkins, F. A., Amaral, W. W., and Shubin, N. H. (2015). Mandibular and dental characteristics of Late Triassic mammaliaform *Haramiyavia* and their ramifications for basal mammal evolution. *Proc. Natl. Acad. Sci. U.S.A.* 112, E7101–E7109.

Luo, Z. X., Ji, Q., and Yuan, C. X. (2007). Convergent dental adaptations in pseudotribosphenic and tribosphenic mammals. *Nature* 450, 93–97. doi: 10.1038/nature06221

Luo, Z.-X., Kielan-Jaworowska, Z., and Cifelli, R. L. (2002). In quest for a phylogeny of Mesozoic mammals. *Acta Palaeontol. Pol.* 47, 1–78.

Luo, Z. X., Meng, Q. J., Grossnickle, D. M., Liu, D., Neander, A. I., Zhang, Y. G., et al. (2017). New evidence for mammaliaform ear evolution and feeding adaptation in a jurassic ecosystem. *Nature* 548, 326–329. doi: 10.1038/nature23483

Luo, Z. X., and Wible, J. R. (2005). A late jurassic digging mammal and early mammalian diversification. *Science* 308, 103–107. doi: 10.1126/science.1108875

Luo, Z.-X., Yuan, C.-X., Meng, Q.-J., and Ji, Q. (2011). A Jurassic eutherian mammal and divergence of marsupials and placentals. *Nature* 476, 442–445. doi: 10.1038/nature10291

Macintyre, G. T. (1967). Foramen pseudovale and quasi-mammals. *Evolution* 21, 834–841. doi: 10.1111/j.1558-5646.1967.tb03437.x

Manley, G. A. (2018). The foundations of high-frequency hearing in Early Mammals. *J. Mamm. Evol.* 25, 155–163. doi: 10.1007/s10914-016-9379-0

Manoussaki, D., Chadwick, R. S., Ketten, D. R., Arruda, J., Dimitriadis, E. K., and O'Malley, J. T. (2008). The influence of cochlear shape on low-frequency hearing. *Proc. Natl. Acad. Sci. U.S.A.* 105, 6162–6166. doi: 10.1073/pnas.0710037105

Martin, T., and Luo, Z. X. (2005). Homoplasy in the mammalian ear. *Science* 307, 861–862. doi: 10.1126/science.1107202

Martin, T., Marugán-Lobón, J., Vullo, R., Martín-Abad, H., Luo, Z. X., and Buscalioni, A. D. (2015). A Cretaceous eutriconodont and integument evolution in early mammals. *Nature* 526, 380-384. doi: 10.1038/nature14905

Meng, J. (2014). Mesozoic mammals of China: implications for phylogeny and early evolution of mammals. *Nat. Sci. Rev.* 1, 521–542. doi: 10.1093/nsr/nwu070

Meng, J., Bi, S., Wang, Y., Zheng, X., and Wang, X. (2014). Dental and mandibular morphologies of *Arboroharamiya* (*Haramiyida*, *Mammalia*): a comparison with other haramiyidans and Megaconus and implications for mammalian evolution. *PLoS One* 9:e113847. doi: 10.1371/journal.pone.0113847

Meng, J., Bi, S., Zheng, X., and Wang, X. (2018). Ear ossicle morphology of the Jurassic euharamiyidan *Arboroharamiya* and evolution of mammalian middle ear. *J. Morphol.* 279, 441–457. doi: 10.1002/jmor.20565

Meng, J., Hu, Y., Wang, Y., Wang, X., and Li, C. (2006). A mesozoic gliding mammal from northeastern China. *Nature* 444, 889–893. doi: 10.1038/nature05234

Meng, J. I. N., Hu, Y., Wang, Y., and Li, C. (2003). The ossified Meckel's cartilage and internal groove in Mesozoic mammaliaforms: implications to origin of the definitive mammalian middle ear. *Zool. J. Linn. Soc.* 138, 431–448. doi: 10.1046/j.1096-3642.2003.00064.x

Meng, Q.-J., Grossnickle, D. M., Liu, D., Zhang, Y. G., Neander, A. I., Ji, Q., et al. (2017). New gliding mammaliaforms from the Jurassic. *Nature* 548, 291–296. doi: 10.1038/nature23476

Meredith, R. W., Janeèka, J. E., Gatesy, J., Ryder, O. A., Fisher, C. A., and Teeling, E. C. (2011). Impacts of the Cretaceous terrestrial revolution and KPg extinction on mammal diversification. *Science* 334, 521–524. doi: 10.1126/science.1211028

Miao, D. (1993). "Cranial morphology and multituberculate relationships," in *Mammal Phylogeny. Mesozoic Differentiation, Multituberculates, Monotremes, Early Therians, and Marsupials*, eds F. S. Szalay, M. J. Novacek, and M. C. McKenna (New York, NY: Springer), 63–74. doi: 10.1007/978-1-4613-9249-1_6

Miyamoto, M. M., and Fitch, W. M. (1995). Testing species phylogenies and phylogenetic methods with congruence. *Syst. Biol.* 44, 64–76.

Mounce, R. C. P., Sansom, R., and Wills, M. A. (2016). Sampling diverse characters improves phylogenies: craniodental and postcranial characters of vertebrates often imply different trees. *Evolution* 70, 666–686. doi: 10.1111/evo.12884

Naylor, G. J., and Adams, D. C. (2001). Are the fossil data really at odds with the molecular data/morphological evidence for Cetartiodactyla phylogeny reexamined. *Syst. Biol.* 50, 444–453. doi: 10.1080/106351501300318030

Nguyen, L. T., Schmidt, H. A., Von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/molbev/msu300

O'Leary, M. A., Gatesy, J., and Novacek, M. J. (2003). Are the dental data really at odds with the molecular data? Morphological evidence for whale phylogeny (re) reexamined. *Syst. Biol.* 52, 853–864. doi: 10.1093/sysbio/52.6.853

Philippe, H., Lopez, P., Brinkmann, H., Budin, K., Germot, A., Laurent, J., et al. (2000). Early–branching or fast–evolving eukaryotes? An answer based on slowly evolving positions. *Proc. R. Soc. BBiol. Sci.* 267, 1213–1221. doi: 10.1098/rspb.2000.1130

Phillips, M. J. (2002). *Monotremata, Marsupialia and Placentalia: Inferring Phylogenetic Relationships From Molecular and Morphological Data*. Palmerston North: Massey University.

Phillips, M. J. (2015). Four mammal fossil calibrations: balancing competing palaeontological and molecular considerations. *Palaeontol. Electron.* 18, 1–16. doi: 10.26879/490

Phillips, M. J., Bennett, T. H., and Lee, M. S. Y. (2009). Molecules, morphology, and ecology indicate a recent, amphibious ancestry for echidnas. *Proc. Natl. Acad. Sci. U.S.A.* 106, 17089–17094. doi: 10.1073/pnas.0904649106

Phillips, M. J., and Fruciano, C. (2018). The soft explosive model of placental mammal evolution. *BMC Evol. Biol.* 18:104. doi: 10.1186/s12862-018-1218-x

Phillips, M. J., Gibb, G. C., Crimp, E. A., and Penny, D. (2010). Tinamous and moa flock together: mitochondrial genome sequence analysis reveals independent losses of flight among ratites. *Syst. biol.* 59, 90–107. doi: 10.1093/sysbio/syp079

Phillips, M. J., and Penny, D. (2003). The root of the mammalian tree inferred from whole mitochondrial genomes. *Mol. Phylogenet. Evol.* 28, 171–185. doi: 10.1016/s1055-7903(03)00057-5

Pridmore, P. A., Rich, T. H., Vickers-Rich, P., and Gambaryan, P. P. (2005). A tachyglossid-like humerus from the early cretaceous of South-Eastern Australia. *J. Mamm. Evol.* 12, 359–378. doi: 10.1007/s10914-005-6959-9

Rambaut, A., and Drummond, A. J. (2014). *Tracer. 2007 Version1.4*.

Ramírez-Chaves, H. E., Weisbecker, V., Wroe, S., and Phillips, M. J. (2016). Resolving the evolution of the mammalian middle ear using Bayesian inference. *Front. Zoo.* 13:39. doi: 10.1186/s12983-016-0171-z

Rauhut, O. W. M., Martin, T., Ortiz-Jaureguizar, E., and Puerta, P. (2002). A jurassic mammal from South America. *Nature* 416, 165–168. doi: 10.1038/416165a

Rich, T. H., Hopson, J. A., Gill, P. G., Trusler, P., Rogers-Davidson, S., and Morton, S. (2016). The mandible and dentition of the Early Cretaceous monotreme *Teinolophos trusleri*. *AlcheringaAust. J. Palaeontol.* 40, 475–501. doi: 10.1080/03115518.2016.1180034

Rich, T. H., Vickers-Rich, P., Constantine, A., Flannery, T. F., Kool, L., and van Klaveren, N. (1997). A tribosphenic mammal from the Mesozoic of Australia. *Science* 278, 1438–1442. doi: 10.1126/science.278.5342.1438

Rich, T. H., Vickers-Rich, P., Constantine, A. E., Flannery, T. F., Lesley, K., and Van Klaveren, N. (1999). Early Cretaceous mammals from Flat Rocks, Victoria, Australia. *Records Queen Victoria Museum* 106, 1–35.

Ronquist, F. (2004). Bayesian inference of character evolution. *Trends Ecol. Evol.* 19, 475–481. doi: 10.1016/j.tree.2004.07.002

Rougier, G. W., Apesteguía, S., and Gaetano, L. C. (2011). Highly specialized mammalian skulls from the Late Cretaceous of South America. *Nature* 479, 98. doi: 10.1038/nature10591

Rougier, G. W., Garrido, A., Gaetano, L., Puerta, P. F., Corbitt, C., and Novacek, M. J. (2007). First Jurassic triconodont from South America. *Am. Mus. Novit.* 3580, 1–17. doi: 10.1206/0003-0082(2007)3580[1:fjtfsa]2.0.co;2

Rougier, G. W., and Novacek, M. J. (1998). Early mammals: Teeth, jaws and finally a skeleton! *Curr. Biol.* 8, R284–R287. doi: 10.1016/s0960-9822(98)70174-5

Rougier, G. W., Wible, J. R., and Hopson, J. A. (1996). Basicranial anatomy of Priacodon fruitaensis (*Triconodontidae*, *Mammalia*) from the Late Jurassic of Colorado, and a reappraisal of mammaliaform interrelationships. *Am. Mus. Novit.* 3183, 1–38.

Rowe, T. (1988). Definition, diagnosis, and origin of Mammalia. *J. Vertebr. Paleonto.* 8, 241–264. doi: 10.1080/02724634.1988.10011708

Rowe, T., Rich, T. H., Vickers-Rich, P., Springer, M., and Woodburne, M. O. (2008). The oldest platypus and its bearing on divergence timing of the platypus and echidna clades. *Proc. Natl Acad. Sci. U.S.A.* 105, 1238–1242. doi: 10.1073/pnas.0706385105

Sánchez-Villagra, M. R., and Williams, B. A. (1998). Levels of homoplasy in the evolution of the mammalian skeleton. *J. Mamma. Evol.* 5, 113–126.

Sansom, R. S., and Wills, M. A. (2017). Differences between hard and soft phylogenetic data. *Proc. R. Soc. BBiol. Sci.* 284:20172150. doi: 10.1098/rspb.2017.2150

Simpson, G. G. (1959). Mesozoic mammals and the polyphyletic origin of mammals. *Evolution* 13, 405–414. doi: 10.1111/j.1558-5646.1959.tb03026.x

Slowinski, J. B. (1993). "Unordered" versus "ordered" characters. *Syst. Bio.* 42, 155–165. doi: 10.1093/sysbio/42.2.155

Springer, M. S., Meredith, R. W., Teeling, E. C., and Murphy, W. J. (2013). Technical comment on "The placental mammal ancestor and the post–K-Pg radiation of placentals". *Science* 341:613. doi: 10.1126/science.1238025

Sueoka, N. (1995). Intrastrand parity rules of DNA base composition and usage biases of synonymous codons. *J. Mole. Evol.* 40, 318–325. doi: 10.1007/bf00163236

Sweetman, S., Smith, G., and Martill, D. (2017). Highly derived eutherian mammals from the earliest Cretaceous of southern Britain. *Acta Palaeontol. Pol.* 62, 1–9.

Swofford, D. L. (2002). *PAUP* 4.0 b10. Phylogenetic Analysis Using Parsimony (and other methods), Version, 4, b10*.

Szalay, F. S. (1993). "Pedal evolution of mammals in the Mesozoic," in *Mammalian Phylogeny. Mesozoic Differentiation, Multituberculates, Monotremes, Early Therians and Marsupials*, eds F. S. Szalay, M. J. Novacek, and M. C. McKenna (New York, NY: Springer-Verlag), 108–128. doi: 10.1007/978-1-4613-9249-1_9

Szalay, F. S. (1994). *Evolutionary History of the Marsupials and an Analysis of Osteological Characters*. Cambridge, MA: Cambridge University Press, 481.

Szalay, F. S., Novacek, M. J., and McKenna, M. C. (eds) (1993). *Mammal Phylogeny: Mesozoic Differentiation, Multituberculates, MONOTREMES, early therians, and Marsupials*. New York, NY: Springe-Verlag.

Tsagkogeorga, G., Turon, X., Hopcroft, R. R., Tilak, M. K., Feldstein, T., Shenkar, N., et al. (2009). An updated 18S rRNA phylogeny of tunicates based on mixture and secondary structure models. *BMC Evol. Biol.* 9:187. doi: 10.1186/1471-2148-9-187

Wagner, G. P. (1996). Homologues, natural kinds and the evolution of modularity. *Am. Zool.* 36, 36–43. doi: 10.1093/icb/36.1.36

Wagner, G. P., and Lynch, V. J. (2005). Molecular evolution of evolutionary novelties: the vagina and uterus of therian mammals. *J. Exp. Zool. Part BMol. Dev. Evol.* 304, 580–592. doi: 10.1002/jez.b.21074

Wang, Y., Hu, Y., Meng, J., and Li, C. (2001). An ossified Meckel's cartilage in two Cretaceous mammals and origin of the mammalian middle ear. *Science* 294, 357–361. doi: 10.1126/science.1063830

Yang, Z. (1994). Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *J. Mol. Evol.* 39, 306–314. doi: 10.1007/bf00160154

Zheng, X., Bi, S., Wang, X., and Meng, J. (2013). A new arboreal haramiyid shows the diversity of crown mammals in the Jurassic period. *Nature* 500, 199–202. doi: 10.1038/nature12353

Zhou, C. F., Bhullar, B. A. S., Neander, A. I., Martin, T., and Luo, Z. X. (2019). New Jurassic mammaliaform sheds light on early evolution of mammal-like hyoid bones. *Science* 365, 276–279. doi: 10.1126/science.aau9345

Zhou, C.-F., Wu, S., Martin, T., and Luo, Z.-X. (2013). A Jurassic mammaliaform and the earliest mammalian evolutionary adaptations. *Nature* 500, 163–167. doi: 10.1038/nature12429

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Celik and Phillips. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genetics

REVIEW
published: 16 October 2020
doi: 10.3389/fgene.2020.506749

Can We Reliably Calibrate Deep Nodes in the Tetrapod Tree? Case Studies in Deep Tetrapod Divergences

*Jason D. Pardo[1,2], Kendra Lennie[2,3] and Jason S. Anderson[1,2]**

[1] Department of Comparative and Experimental Biology, Faculty of Veterinary Medicine, University of Calgary, Calgary, AB, Canada, [2] McCaig Institute for Bone and Joint Health, University of Calgary, Calgary, AB, Canada, [3] Department of Biological Sciences, University of Calgary, Calgary, AB, Canada

OPEN ACCESS

Edited by:
Michel Laurin,
UMR 7207 Centre de Recherche sur la Paléobiodiversité et les Paléoenvironnements (CR2P), France

Reviewed by:
David Buckley,
Autonomous University of Madrid, Spain
Roger Benson,
University of Oxford, United Kingdom

***Correspondence:**
Jason S. Anderson
janders@ucalgary.ca

Specialty section:
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Genetics

Received: 22 October 2019
Accepted: 03 September 2020
Published: 16 October 2020

Citation:
Pardo JD, Lennie K and Anderson JS (2020) Can We Reliably Calibrate Deep Nodes in the Tetrapod Tree? Case Studies in Deep Tetrapod Divergences. Front. Genet. 11:506749. doi: 10.3389/fgene.2020.506749

Recent efforts have led to the development of extremely sophisticated methods for incorporating tree-wide data and accommodating uncertainty when estimating the temporal patterns of phylogenetic trees, but assignment of prior constraints on node age remains the most important factor. This depends largely on understanding substantive disagreements between specialists (paleontologists, geologists, and comparative anatomists), which are often opaque to phylogeneticists and molecular biologists who rely on these data as downstream users. This often leads to misunderstandings of how the uncertainty associated with node age minima arises, leading to inappropriate treatments of that uncertainty by phylogeneticists. In order to promote dialogue on this subject, we here review factors (phylogeny, preservational megabiases, spatial and temporal patterns in the tetrapod fossil record) that complicate assignment of prior node age constraints for deep divergences in the tetrapod tree, focusing on the origin of crown-group Amniota, crown-group Amphibia, and crown-group Tetrapoda. We find that node priors for amphibians and tetrapods show high phylogenetic lability and different phylogenetic treatments identifying disparate taxa as the earliest representatives of these crown groups. This corresponds partially to the well-known problem of lissamphibian origins but increasingly reflects deeper instabilities in early tetrapod phylogeny. Conversely, differences in phylogenetic treatment do not affect our ability to recognize the earliest crown-group amniotes but do affect how diverse we understand the earliest amniote faunas to be. Preservational megabiases and spatiotemporal heterogeneity of the early tetrapod fossil record present unrecognized challenges in reliably estimating the ages of tetrapod nodes; the tetrapod record throughout the relevant interval is spatially restricted and disrupted by several major intervals of minimal sampling coincident with the emergence of all three crown groups. Going forward, researchers attempting to calibrate the ages for these nodes, and other similar deep nodes in the metazoan fossil record, should consciously consider major phylogenetic uncertainty, preservational megabias, and spatiotemporal heterogeneity,

preferably examining the impact of working hypotheses from multiple research groups. We emphasize a need for major tetrapod collection effort outside of classic European and North American sections, particularly from the southern hemisphere, and suggest that such sampling may dramatically change our timelines of tetrapod evolution.

Keywords: tetrapod, prior constraint, node age prior, fossil record bias, phylogeny

INTRODUCTION

Modern biodiversity is generally organized into large, relatively ancient, clades (i.e., Amniota, Mammalia, and Reptilia) with characteristic body plans and broad ecomorphological similarity. Building a comprehensive understanding of the origin and diversification of these major taxa is a uniquely challenging research program. Often, we are studying groups that originated long ago, defined by long branches to living representatives of the clade and at its base. For example, crown Tetrapoda (the most recent ancestor of living reptiles, mammals, and amphibians) diverged from its most recent living clade, the lungfish (Takezaki and Nishihara, 2017) over 400 million years ago (e.g., Zhu and Yu, 2002), leaving a long stem occupied by a diversity of fossil species that document important evolutionary events such as the acquisition of limbs and digits and emergence on land. As a result, these major taxa are often quite distinct from their closest living relatives, making it difficult to isolate specific intrinsic and extrinsic drivers that may explain their success. Intrinsic factors typically refer to heritable factors that govern a population's ability to generate new forms through evolutionary novelties, changes in evolvability, and developmental canalization (Hendrikse et al., 2007) or ability of established forms to maximize fitness in a range of possible environments through ecology, physiology, plasticity, and functional morphology (Schluter, 2000). Extrinsic factors, on the other hand, typically refer to large-scale changes in the overall state of the Earth's biosphere, including changes in biogeographic connectivity due to plate tectonics (San Mauro et al., 2005; Pyron, 2014), shifts in nutrient or oxygen availability (Ward et al., 2006), shifts in global climate (Chaboureau et al., 2014), and global mass extinction events that either serve as discrete events, which culled global diversity (Raup and Sepkoski, 1982; Sallan and Coates, 2010) or vacated niches to permit subsequent diversification of survivors (e.g., Field et al., 2018). Given that these hypothesized extrinsic factors explicitly invoke geological or macroecological conditions that existed at a specific time in Earth's history, testing a relationship between these factors and the evolution of major taxa requires precise, accurate constraints on the timing of the origin and diversification of those taxa.

As the fossil record is incomplete, it is often impossible to directly use fossils to establish tight constraints on the origin of major taxa. To address this problem, a series of methods have been devised to use relative difference in molecular sequence between two taxa to estimate the age of the divergence between those taxa. These methods, collectively termed "the molecular clock," integrate paleontological data (as node calibration dates) and molecular data (as sequence divergence or estimated branch length) to produce estimates of the ages of all nodes on a phylogenetic tree. Although early implementation of these methods was highly procedural and prone to multiplication of error (Graur and Martin, 2004), newer approaches have re-envisioned node calibration dates as a range of prior probabilities for the age of a node ("node priors"), allowing coestimation of tree topology and age (Stadler and Yang, 2013), potentially improving precision of node estimates. Furthermore, a series of *a posteriori* methods have been created to assess quality of individual node calibrations within a set of calibrations. In these approaches, the quality of individual calibrations is tested by comparing how well each calibration can predict the ages of all other calibrations (Near and Sanderson, 2004; Stadler and Yang, 2013; Heath et al., 2014). Integration of these methods into phylogenetic analysis has even been used as a means of discerning between phylogenetic trees (Lee and Yates, 2018; King and Beck, 2019) and for dating the age of specific fossils for critical evaluation (King and Beck, 2019). These methods are appealing because practitioners are free to engage with mathematically tractable patterns in the data rather than engage in taxonomic arguments of otherwise narrow interest, based on broadly inaccessible and subjective debates on the importance of specific anatomical features for inference of phylogeny. Conversely, these analytical approaches effectively give hypotheses of rate of change veto power over the estimated fossil age and taxonomic ID that serve as the primary data used to test those hypotheses. This has led to an emerging analytical pipeline that selects trees or calibration ages *a posteriori*, and in doing so excludes or reinterprets primary data that inconveniently conflicts with the overall pattern of results (King and Beck, 2019). Although the long-term utility of these methods remains to be seen, *a priori* assessment of the quality of *a priori* node calibrations must retain logical primacy in assessing the quality of a molecular clock (Hedges et al., 2018; Morris et al., 2018).

Node-age calibrations themselves require a detailed assessment of the fossil record to identify the earliest member of a given clade. Identifying the earliest members of a clade requires substantial specialist knowledge of the anatomy of the group, how variation in that anatomy corresponds with the crown group, and the temporal distribution of fossils that exhibit that anatomy. This specialist knowledge from paleontology is often far outside the expertise of molecular phylogeneticists. To facilitate easy access to this knowledge, compendia of node calibration dates have been assembled first by Benton and Donoghue (2006) and more recently by Benton et al. (2015). These compendia present a list of node minima and maxima for many clades in the tree of life and typically claim a lack of ambiguity over these proposed node calibration ages. These compendia are widely treated as expert-vetted calibration points in molecular clock studies (Feng et al., 2017; Hime et al., 2020), with little to no direct consultation with experts. This assumes several things: that paleontological experts address phylogeny in

a manner consistent with usage by molecular clock approaches, that compendia such as Benton et al. (2015) accurately report consensus between paleontological workers and stability of the underlying tree, and that stability of age estimates reflects biological processes recorded in molecular data.

To date, discussions refining best practices in node calibration have focused on ensuring that fossils chosen as node-age calibrations fall with certainty within the crown group, that their precise stratigraphic resolution is provided, and that this precise stratigraphic resolution is placed into an explicit numerical framework (Parham and Irmis, 2008; Parham et al., 2012). However, considerably less attention has been given to factors influencing calibrations of deeper nodes indicating the divergence of major clades. These nodes are important because they often serve as external bounds on node age interpolation (Duchêne et al., 2014) and because their position deep within the tree of life means they are likely to appear frequently in studies using a molecular clock (Müller and Reisz, 2005; Chen et al., 2015; Feng et al., 2017; Hime et al., 2020). Given the importance of reliable node age calibrations in these deeper nodes, it is critical to ask whether current recommendations of best practices, and the calibrations outlined in compendia, are sufficient.

Three such nodes of interest are the deep divergences within the Tetrapoda. Tetrapoda is a monophyletic grouping that includes all descendants of the common ancestor of modern amphibians, reptiles, birds, and mammals. These make up the entirety of extant vertebrates with digited limbs. The term "Tetrapoda" is generally applied to digited members of the total group, a usage that is equivalent to Stegocephalia (Laurin et al., 2000), whereas members of the crown-group are sometimes referred to Neotetrapoda (Sues, 2019). Tetrapoda consists of two clades: the Lissamphibia and Amniota (Chen et al., 2015; Irisarri et al., 2017; Hime et al., 2020). Lissamphibians include the caecilians (Gymnophiona), frogs (Anura), and salamanders (Caudata) and are characterized by thin, permeable mucous skin. Amniota includes mammals (Mammalia), birds (Aves), and 'reptiles', and is characterized by keratinized skin and a unique extraembryonic membrane, the amnion (Reisz, 1997). Each of these clades is notable in that they are all very old (>265 Ma) and that the monophyly of each clade is not in serious contention (Chen et al., 2015; Irisarri et al., 2017; Hime et al., 2020). Additionally, in each case modern body plans are extremely different from fossil forms, to the extent that it is difficult if not impossible to identify diagnostic characters of the crown group without reference to fossil diversity.

We here review these three calibration points to understand how current best practices for node calibration may fail to guide calibration of Palaeozoic nodes. We discuss how phylogenetic problems in the Palaeozoic, including node calibration, are almost entirely dependent on interpretation of morphology among fossil groups rather than reference to an independently inferred molecular phylogeny. We then explore specific features of phylogenetic uncertainty among Palaeozoic tetrapods, and how subtly different interpretations of Palaeozoic tetrapod interrelationships suggest very different timelines for the origin of these three tetrapod clades. We finally discuss general spatial and temporal patterns in the early tetrapod fossil record, and how these may bias against discovery of early members of each clade. Finally, we provide recommendations that we believe will mitigate some of the problems currently affecting these node calibrations and that may provide a framework for efforts to calibrate similar nodes in other taxa.

ORIGINS VERSUS AFFINITIES

Assigning a node age calibration requires identification of the oldest known fossil that can be assigned to an extant clade, but there is some variation in how this is done. To establish universal standards, Parham et al. (2012) outlined a set of best practices. This set of best practices focuses largely on connecting an occurrence with stratigraphic information. Less attention has been given to outlining standards for ensuring that the fossil used in a calibration in fact belongs to the clade in question; Parham et al. (2012) suggest that apomorphies be identified in the specimen used to date the clade, but do not suggest universal standards for how these apomorphies are to be chosen in the first place.

How are these apomorphies chosen in practice? To examine this, we use as an example only the calibration list of Hime et al. (2020), a recent phylogenomic analysis of lissamphibian diversification employing a molecular clock. This calibration list is chosen here as it represents one of the larger and more comprehensive calibration sets employed across early vertebrate diversity, and is largely consistent with other recent calibration sets, such as Cannatella (2015) and Feng et al. (2017). Hime et al. (2020) employ a total of nineteen calibration points. Of these, the oldest calibration point (Tetrapoda) has a minimum age of 337 Ma, whereas the youngest minimum calibration point (*Ptychadena* + *Phrynobatrachus*) is set at 25 Ma. The methodologies for choosing these calibration points are varied; several taxa (*Chunerpeton tianyiensis* and *Iridotriton hechti*) were initially assigned to nodes without an explicit phylogenetic analysis (Gao and Shubin, 2003; Evans et al., 2005), and two taxa (*Calyptocephalella pichileufensis* and an unnamed fossil ptychadenid) have never been assessed within a phylogenetic framework (Gómez et al., 2011; Blackburn et al., 2015). In these cases, assignment to a given clade is accomplished solely through comparative anatomy and reference to differential diagnoses. However, most nodes have been assessed through some manner of phylogenetic analysis. In the case of all node calibrations in the Mesozoic and Cenozoic, these phylogenetic analyses invariably include at least a subset of extant taxa. In fact, these analyses typically include a large majority of extant taxa, but nodes representing divergences in the Paleozoic differ in the constitution of the overall phylogenetic sampling. Phylogenetic analyses cited for node calibrations of the divergence of Amniota, Batrachia, Lissamphibia, and Tetrapoda do not sample a single extant taxon in any cited case (see Anderson, 2008 for further discussion).

This distinction between Paleozoic and post-Paleozoic divergences is noteworthy. Relationships of fossils used as node calibrations in the Mesozoic and Cenozoic are investigated via comparison with the specific taxa sampled for molecular

sequence data, and interrelationships between fossil taxa are generally not important for resolving phylogenetic disputes. In direct contrast, node calibrations in the Palaeozoic depend on fine interrelationships between sometimes-obscure fossil taxa with little to no direct comparison with extant organisms. This places molecular phylogeneticists in a predicament: calibration of Paleozoic nodes may require engagement with paleontological literature and contending with disputes among those workers.

NODE MINIMA: WHAT ARE THE EARLIEST REPRESENTATIVES OF THE MAJOR TETRAPOD CROWN GROUPS?

When we talk about phylogenetic uncertainty of fossils involved in node calibration, we typically have in mind a situation where there is a relatively dense phylogeny of modern taxa and the difficulty is in finding fossils that preserve sufficient diagnostic anatomy to be placed confidently into this phylogenetic framework (Patterson, 1981; Parham et al., 2012). In a situation such as this, diagnostic characteristics can be determined *a priori* through comparative anatomy of extant organisms with known phylogenetic relationships. In such an ideal case, the primary challenges are local uncertainty in phylogeny and in specific node age calibrations, and many tools, such as Bayesian tip-dating (Stadler and Yang, 2013) are designed to handle these problems by assessing these sorts of local patterns of uncertainty (uncertainty of local tree resolution, uncertainty of specific fossil age interpretations) as a range of posterior probabilities. Under these circumstances, molecular phylogeneticists do not need to engage with the paleontological record beyond identifying the taxa that need to be incorporated into an analysis.

One way out of this problem has been for paleontologists to assemble compendia of recommended nodes for use in molecular clock calibrations and fossils to use for calibration of these nodes (Benton and Donoghue, 2006; Benton et al., 2015). These compendia provide lists of nodes and fossil taxa with reference to the paleontological literature, but generally do not provide substantial discussion of the specific bases for these attributions or differences in expert opinion. This approach is generally acceptable for Mesozoic and Cenozoic divergences, where the anatomical basis for relationships between modern groups is well-understood. However, this is not the case for deep divergences in the tetrapod tree. Early tetrapod phylogeny is highly unstable and lacking in consensus. Calibration of these nodes depends on broad anatomical comparisons across the entire early tetrapod diversification, beginning in the late Devonian and extending through the early Permian. These anatomical comparisons also extend to the earliest representatives of modern amphibian lineages in the Mesozoic (Maddin et al., 2012; Ascarrunz et al., 2016; Schoch et al., 2020), as these fossils preserve generalized tetrapod anatomy not seen in modern representatives of these groups and therefore provide insight into the relationships between amphibians, amniotes, and extinct tetrapod groups. This instability manifests as two major points of controversy: (1) what are the general interrelationships of major archaic early tetrapod taxa and (2) what is the relationship between major archaic early tetrapod taxa and lissamphibians? An addendum to the second point is that some workers have questioned the inclusiveness of the lissamphibian crown group itself, depending on how convergences between modern lissamphibian orders are interpreted (Anderson et al., 2008; Pardo et al., 2017b). Furthermore, the differences between these phylogenetic hypotheses are not trivial. Different phylogenetic hypotheses of early tetrapod relationships and of lissamphibian origins represent substantially different interpretations of the nature of crown tetrapod and crown lissamphibian characters, and a resulting different timeline of tetrapod origins (**Figure 1**). Given that phylogenetic analyses treating this problem must consider the anatomy of fossils spanning approximately the first 170 Ma of tetrapod evolution (**Figure 2**) and compare hypotheses suggesting very different patterns of body plan evolution, this is not a simple problem.

Phylogenetic Context of the Amniote Crown

The phylogenetic relationships of vertebrate taxa associated with the origins of the amniote crown (i.e., the mammal-reptile split) are relatively stable (Laurin and Reisz, 1995). Amniotes are recognized as comprising two clades, the Reptilia and the Mammalia. Although some disagreement remains concerning the relationship of turtles among other reptiles (Chiari et al., 2012; Field et al., 2014; Bever et al., 2015; Schoch and Sues, 2015), there is essentially no disagreement concerning the monophyly of these two amniote clades. The phylogenetic relationships of modern amniote clades to Palaeozoic relatives are relatively stable, although some disagreements do exist.

The fossil record of total-group mammals (Synapsida) provides an exceptional record of the origin of the crown group from Palaeozoic ancestors. Broad trends in the assembly of the mammalian body plan has been reconstructed with wide consensus based on the dense record of total-group mammals (therapsids) from the late Permian and early Triassic, and confidently extending back through the Late Carboniferous "pelycosaurs" (Sidor and Hopson, 1998). These "pelycosaurs" can be assigned to several major clades, the Eupelycosauria, the Varanopidae, and the Caseasauria (Laurin and Reisz, 1995; Sidor and Hopson, 1998; Benson, 2012). Therapsids are thought to fall within the Eupelycosauria, whereas varanopids and caseasaurs are thought to represent successive outgroups to this clade (Sidor and Hopson, 1998; Benson, 2012).

The early record of synapsids has historically been relatively depauperate. The earliest definitive synapsid fossils are known from the Moscovian stage of the Carboniferous (315.2–307 Ma) of North America and the Czechia. These fossils are primarily attributable to eupelycosaurs, including *Archaeothyris florensis* (Reisz, 1972) and *Echinerpeton intermedium* (Reisz, 1972; Mann et al., 2019), although newly described fossils demonstrate the presence of a varanopid, *Dendromaia unamakiensis* from the same age (Maddin et al., 2020). Fragmentary fossils ambiguously attributable to synapids are known from the Bashkirian stage of the Carboniferous (320–315.2 Ma) of Joggins, Nova Scotia.

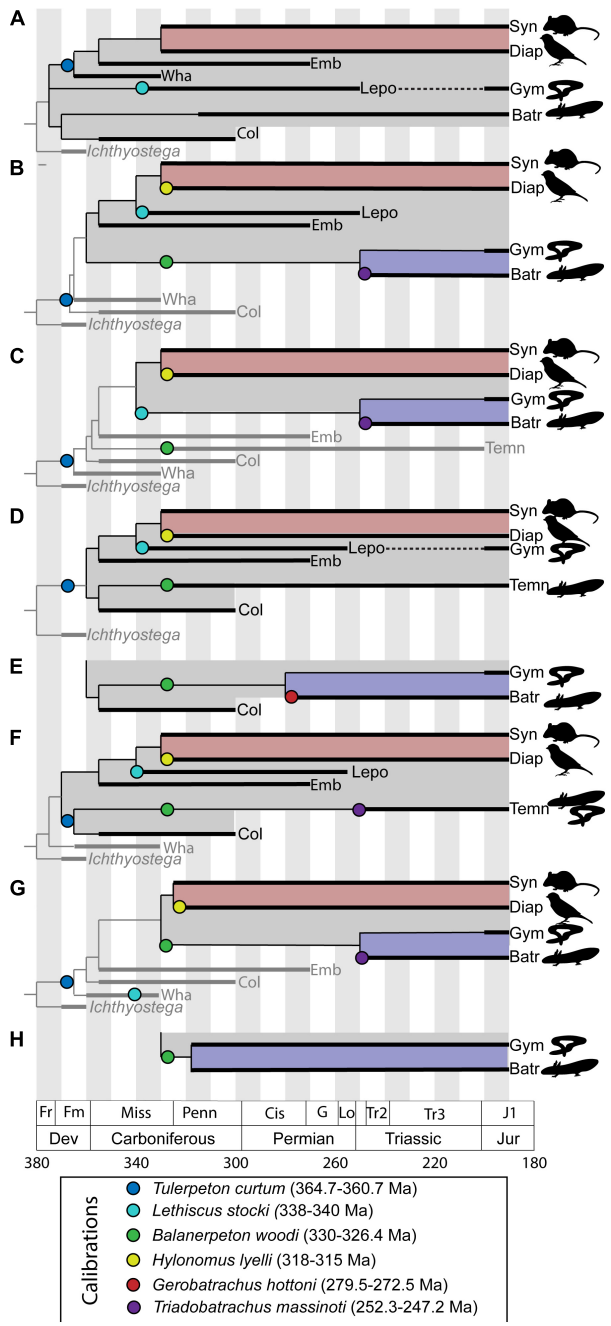


FIGURE 1 | Comparative phylogenetic hypotheses of early tetrapod interrelationships and lissamphibian affinities, showing effects on node calibrations for the age of tetrapod (gray) and lissamphibian crown (blue). **(A)** Modified from Carroll (2001); **(B)** modified from Ruta and Coates (2007); **(C)** modified from Vallin and Laurin (2004); **(D)** modified from Anderson et al. (2008); **(E)** variant of **(D)** modified from Maddin et al. (2012); **(F)** modified from Clack et al. (2017); **(G)** modified from Pardo et al. (2017a); **(H)** variant of **(G)** modified from Pardo et al. (2017b). Taxonomic abbreviations: Batr, Batrachia; Col, Colosteida; Diap, Diapsida; Emb, Embolomeri; Gym, Gymnophiona; Lepo, Lepospondyli; Syn, Synapsida; Temn, Temnospondyli; Wha, Whatcheeriida. Gray boxes delineate the tetrapod crown, red boxes delineate the amniote crown, and blue boxes delineate the lissamphibian crown (where present). Phylogenetic position of taxa often used to calibrate node ages denoted with colored circles.

Bashkirian records of synapsids were previously limited to the partial skeleton *Protoclepsydrops haplous* (Carroll, 1964, although disputed by Reisz, 1972), but it has recently been proposed that *Asaphestera platyrhis*, previously considered a tuditanomorph "microsaur," might in fact be a caseasaur synapsid from the same set of localities (Mann et al., 2020).

The precise composition of the reptile stem group is somewhat more contentious. Earliest definitive members of the reptile crown group are relatively derived stem-archosaurs, such as the proterosuchids and prolacertids of the Permo-Triassic Boundary (e.g., Benton, 1984; Evans, 1984; Dilkes, 1998; Nasbitt, 2011; Ezcurra, 2016; Simões et al., 2018). A diverse assemblage of possible stem-reptiles (claudiosaurids, weigeltosaurids, and younginids) are known from the Upper Permian, but how these relate to Carboniferous and Early Permian taxa is uncertain. Traditional Carboniferous-Permian eureptiles have been assigned to four major groups: the Araeoscelidae, Protorothyrididae, Captorhinidae, and Parareptilia (Laurin and Reisz, 1995). Of these, the Araeoscelidae is thought to be most closely related to the stem-reptiles of the Upper Permian (Laurin and Reisz, 1995; Reisz et al., 2011; Ford and Benson, 2018) and protorothyrididae are thought to represent a paraphyletic assemblage that includes both derived diapsid relatives, as well as early diverging captorhinids (Müller and Reisz, 2006). Phylogenetic treatments have variously found the parareptiles to be the sister clade of all other reptiles (Gauthier et al., 1988; Modesto et al., 2015), or slightly closer to the crown (less mesosaurs, Laurin and Reisz, 1995), or a polyphyletic assemblage of stem-reptiles, stem-turtles, or both (Bever et al., 2015; Laurin and Reisz, 1995; Modesto et al., 2015; Ford and Benson, 2020). Within this framework, the earliest hypothesized member of the reptile stem group is *Hylonomus lyelli* (**Figure 2H**) from the Bashkirian Joggins Formation of Nova Scotia (Carroll, 1964), with additional abundant reptile material preserved throughout the early Pennsylvanian (**Figures 1**, **3**).

However, despite the appearance of phylogenetic consensus, there is in fact a large amount of uncertainty and disagreement on overall phylogenetic relationships of many of these archaic taxa with respect to the two major amniote clades. The overall framework of early amniote phylogeny, and therefore the actual phylogenetic affinities of fossils typically identified as earliest crown amniotes, depends largely on distinctions made by Carroll (1964) prior to systematic phylogenetic analytical techniques, validated in part by early phylogenetic analyses (e.g., Laurin and Reisz, 1995). Internal organization of early divergences within Amniota has varied across a number of analyses, including substantial reorganization of early synapsid relationships (Benson, 2012), a possible position of protorothyridids at the base of the mammal stem (Brocklehurst et al., 2016; Matzke and Irmis, 2018), massive reorganization of early stem-reptiles (Laurin and Piñeiro, 2017; Ford and Benson, 2020), a paraphyletic Synapsida (MacDougall et al., 2018), and a possible displacement of a major clade of synapsids onto the reptile stem (Ford and Benson, 2018, 2020). This uncertainty suggests that attribution of fragmentary Bashkirian and Moscovian taxa, important for node date estimation, to either the reptile or mammal total groups may be volatile.

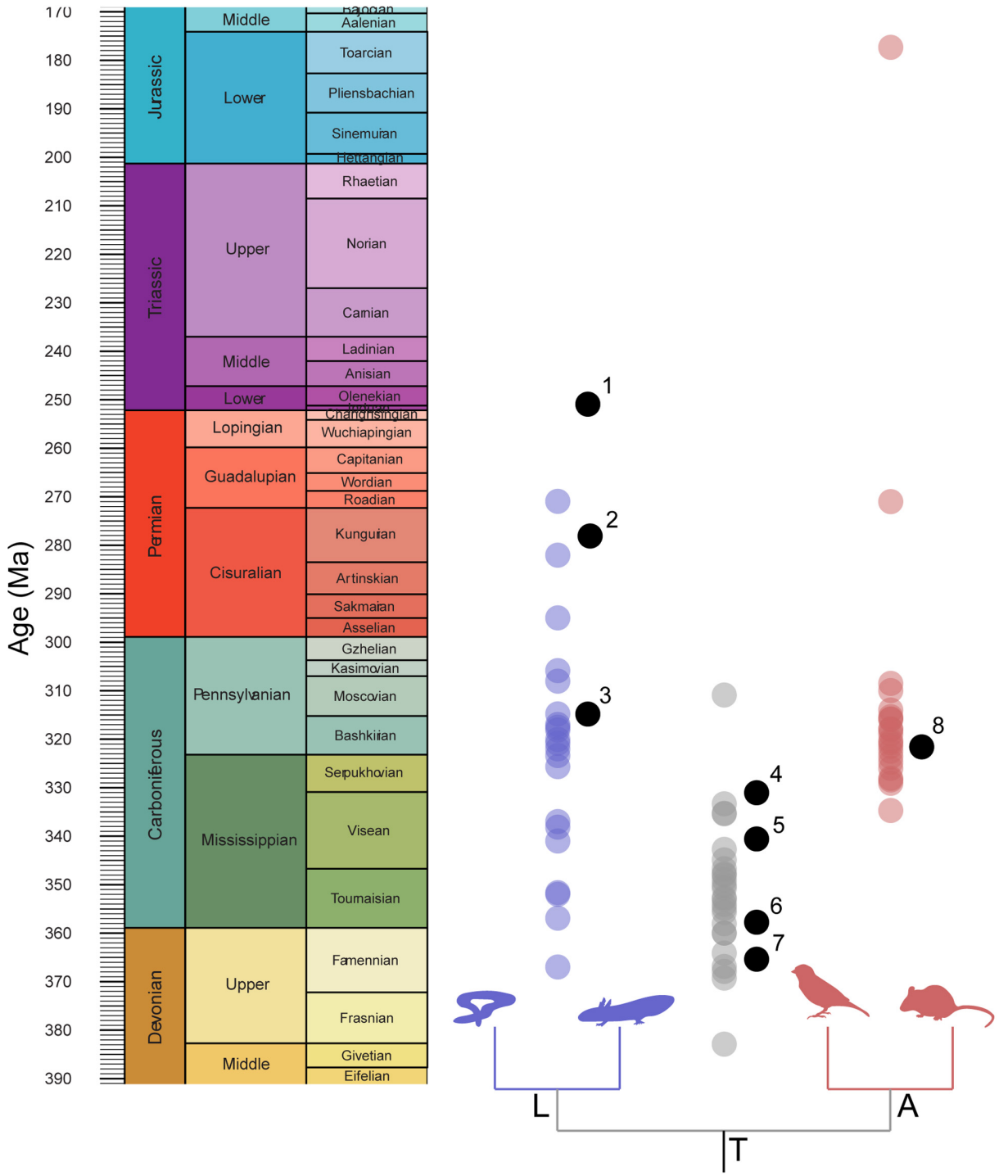


FIGURE 2 | Distribution of implied node calibration ages compared with distribution of mean node-age estimates. Node estimates drawn from timetree.org. 1, *Triadobatrachus massinoti*; 2, *Gerobatrachus hottoni*; 3, *Amphibamus grandiceps*; 4, *Balanerpeton woodi*; 5, *Lethiscus stocki*; 6, Horton Bluff tetrapod fauna; 7, *Tulerpeton curtum*; 8, *Hylonomus lyelli*.

Furthermore, some taxa not traditionally considered amniotes have appeared within the Amniota in some recent analyses. Most notably, the Recumbirostra, a group of small fossorial tetrapods traditionally classified within a larger order Microsauria and sometimes considered to be related to extant amphibians (Vallin and Laurin, 2004; Marjanović and Laurin, 2013, 2019), has recently been placed on the reptile stem based on neurocranial similarities (Pardo et al., 2015, 2017b; Szostakiwskyj et al., 2015).

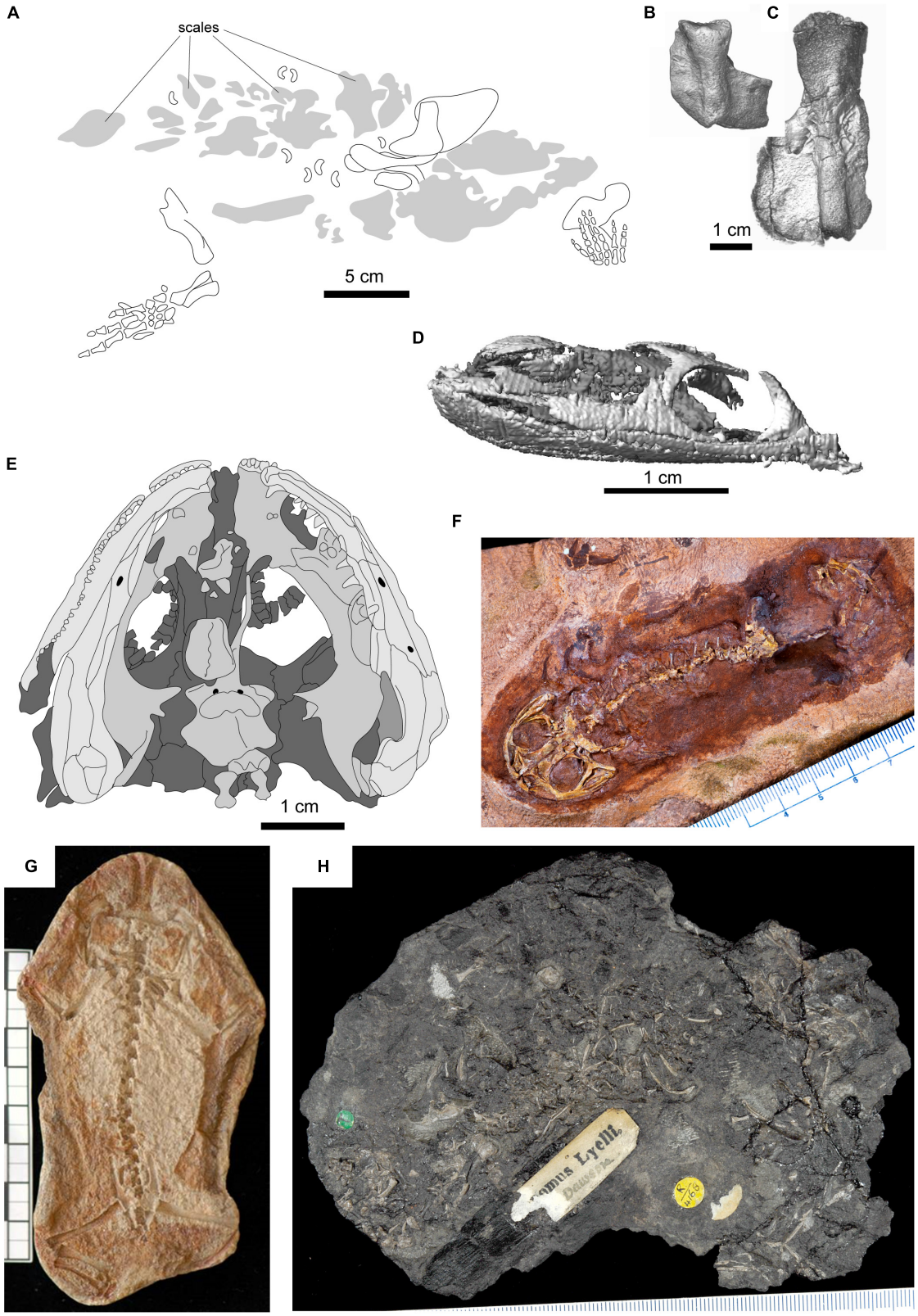


FIGURE 3 | Selected fossils representing node-age calibrations. **(A)** *Tulerpeton curtum*, after Lebedev and Coates (1995); **(B)** Horton Bluff colosteid-like taxon; **(C)** Horton Bluff embolomere-like taxon; **(D)** *Lethiscus stocki*, segmented skull based on micro-CT; **(E)** type specimen of *Balanerpeton woodi*, after Milner and Sequeira (1993); **(F)** *Gerobatrachus hottoni*; **(G)** *Triadobatrachus massinoti*; **(H)** type specimen of *Hylonomus lyelli*.

This clade is relatively diverse in the Joggins Formation (Carroll, 1966; Mann et al., 2020), including forms such as *Steenerpeton sylvae*, *Trachystegos megalodon*, and *Leiocephalikon problematicum*, and *Hylerpeton dawsoni*, all of which can be confidently assigned to recumbirostran subclades for which complete articulated fossils exist. This contrasts with the relatively fragmentary fossils attributed to *Hylonomus lyelli* (**Figure 3H**; Carroll, 1964), as well as the fragmentary and ambiguous fossils attributed to synapsids from the same locality (Carroll, 1964; Mann et al., 2020), allowing phylogenetic relationships to be assessed with greater confidence. Although recent work finding the recumbirostran diversification within the Amniota remains controversial (compare Pardo et al., 2017b with Marjanović and Laurin, 2019), recent studies continue to identify reptile-like anatomy of the recumbirostran temporal fenestra (Gee et al., 2019) and dentition (Gee et al., 2020). Additionally, the Diadectamorpha, traditionally conceptualized as the sister clade of amniotes, has been found within the mammalian total group on the basis of occipital morphology (Berman, 2000), a result that has recently received new support based on anatomy of the inner ear (Klembara et al., 2020). This latter result currently has no major implications for node calibrations, however, as the oldest diadectamorph fossils are substantially younger than most putative early crown amniotes.

Phylogenetic Context of the Amphibian Crown

Extant amphibians can be assigned to three monophyletic orders, the frogs (Anura), salamanders (Caudata), and caecilians (Gymnophiona). Although there is some uncertainty about the interrelationships of these three groups (Marjanović and Laurin, 2013), most analyses support the existence of a Batrachian clade that comprised Anura and Caudata to the exclusion of Gymnophiona (Hime et al., 2020). Molecular analyses consistently recover an amphibian clade (Lissamphibia) to the exclusion of Amniota, but a minority of phylogenetic analyses of amphibian morphology have found Gymnophiona as the sister clade of Amniota (**Figure 1D**), thus rendering Lissamphibia polyphyletic (Carroll, 2007; Anderson et al., 2008), although this has been rejected in more recent iterations of those analyses (Maddin et al., 2012; Pardo et al., 2017a). The timing of the origin of the amphibian crown group is difficult to determine, in large part because the phylogenetic context of amphibian origins remains hotly debated. Earliest representatives of all three modern lissamphibian orders are already highly derived, making it difficult to define a lissamphibian bauplan, and this lack of a clear lissamphibian bauplan has subsequently led to difficulties in placing lissamphibians into Palaeozoic diversity more generally. Both classic comparative and modern analytical approaches to the phylogenetic relationships of lissamphibian origins have found relationships between modern lissamphibians and two groups of early tetrapods, the Temnospondyli and Lepospondyli. The former comprised mostly medium to large-bodied tetrapods with a few small-bodied lineages, but share features of the braincase, palate, and limbs with modern amphibians (Anderson et al., 2008; Sigurdsen and Bolt, 2009; Fröbisch and Shubin, 2011; Maddin and Anderson, 2012; Maddin et al., 2012; Witzmann and Werneburg, 2017), whereas the latter is mostly small-bodied and generally shares patterns of cranial bone reduction and vertebral consolidation with modern amphibians (Marjanović and Laurin, 2013, 2019). General trends in sequence heterochrony have been invoked in support of both phylogenetic hypotheses (Fröbisch et al., 2007; Olori, 2013; Laurin, 2019), but the implications of these data remain unclear. Among temnospondyls, most phylogenetic analyses place lissamphibians within amphibamid dissorophoids (**Figures 1B,E**). Phylogenetic analyses finding a lepospondyl origin of lissamphibians have typically placed lissamphibians within a clade that comprised "lysorophians" and brachystelechid "microsaurs" (**Figure 1C**), which are currently recognized by most workers as recumbirostrans as discussed above (Anderson et al., 2008; Maddin et al., 2012; Pardo et al., 2017b).

These alternative hypotheses have different implications for the age of the lissamphibian crown group, even though they primarily concern the nature of the lissamphibian stem group. The earliest unambiguous lissamphibian fossil is the stem-anuran *Triadobatrachus massinoti* (**Figure 3H**) from the earliest Triassic Sakamena Formation of Madagascar (Rage and Roček, 1989; Ascarrunz et al., 2016). Early caudates appear by the Middle Triassic of Kyrgyzstan (Schoch et al., 2020), whereas the earliest unambiguous stem-gymnophionans are Jurassic in age (Jenkins et al., 2007). In phylogenetic analyses that place lissamphibians within lepospondyls, no Palaeozoic tetrapods are found within the lissamphibian crown group (Marjanović and Laurin, 2013, 2019). Phylogenetic analyses that find lissamphibians within temnospondyls intermittently do find evidence of Palaeozoic representatives of the lissamphibian crown group, however. One possible Palaeozoic crown-group amphibian is the early Permian amphibamid *Gerobatrachus hottoni* (**Figure 3F**) from the Clear Fork Group (Kungurian) of Texas, which preserves a mosaic of anatomical features typical of anurans, caudates, and more generalized temnospondyls (Anderson et al., 2008). Different phylogenetic treatments have disagreed on the placement of *Gerobatrachus*, either placing it as the sister taxon to batrachians (Anderson et al., 2008; Maddin et al., 2012) or just outside the crown group (Sigurdsen and Green, 2011) in trees that align with the Temnospondyl hypothesis. However, the inclusiveness of the lissamphibian crown group depends more generally on the position of caecilians (Anderson, 2008). Most workers have not found evidence of Palaeozoic stem-group representation of gymnophionans. Pardo et al. (2017b) identified large-scale similarities between the caecilian skull and the skulls of a different temnospondyl group, the mostly Triassic-aged rhytidostean stereospondyls. Although similar levels of cranial consolidation between gymnophionans and specialized rhytidosteans may reflect convergence in headfirst burrowers, Pardo et al. (2017b) also identified a number of major anatomical similarities that cannot be so easily dismissed, suggesting that more inclusive phylogenetic analyses are necessary to properly test hypotheses of gymnophionan origins. If this phylogenetic hypothesis is correct, it would suggest a much more inclusive lissamphibian crown group and earlier origin of the amphibian crown group (**Figure 1G**). The earliest definitive crown lissamphibians

in this phylogeny would be the dissorophoids *Amphibamus grandiceps* and an unnamed branchiosaurid from the early Moscovian Francis Creek Shale of Illinois, United States (∼315 Ma, Milner, 1982). The early temnospondyl *Eugyrinus wildi* from the Bashkirian (∼318–315 Ma) of the Lower Coal Measures Formation of Lancashire, United Kingdom, would be ambiguously assignable to the lissamphibian crown group as well (Milner, 1980).

Phylogenetic Context of the Tetrapod Crown

Whereas the composition of the amniote crown group is relatively stable, and the composition of the lissamphibian crown group is only questioned in a minority of analyses, the composition of the tetrapod crown group among early tetrapods is hugely controversial with very little consensus (Ruta et al., 2003; Ruta and Coates, 2007; Anderson et al., 2008; Marjanović and Laurin, 2013, 2019; Clack et al., 2017; Pardo et al., 2017a,b). Because of substantial changes in understanding of early tetrapod phylogeny over the past 40 years, essentially every major group of Carboniferous tetrapods has been alternately placed both within the tetrapod crown group and outside of it in different phylogenetic hypotheses (**Figure 1**). Importantly, the temporal range of some of these groups appears to extend back in time to the latest Devonian, so these differences in phylogenetic hypotheses can have major implications on the minimum node calibration age for the tetrapod crown group. Discussions of phylogenetic uncertainty in the origin of the tetrapod crown group have attributed this uncertainty to one of two major problems: (1) that different hypotheses of lissamphibian origins imply a less inclusive (Lepospondyl hypothesis) or more inclusive (Temnospondyl hypothesis) tetrapod crown group within a more stable tetrapod phylogeny (Anderson, 2008; Marjanović and Laurin, 2013), or (2) that uncertainty of deep interrelationships between major Carboniferous tetrapod lineages stems from an explosive radiation dating back to the End Devonian mass extinction (Coates et al., 2008). Both factors contribute to overall uncertainty concerning the composition of the tetrapod crown group, although this appears to be a much broader problem.

As we noted above, paleontologists have had considerable difficulty determining the immediate Paleozoic outgroups of modern lissamphibians, but two major groups of early tetrapods have been identified as credible candidates, the Lepospondyli and the Temnospondyli. Lepospondyls are a morphologically diverse group of early tetrapods with little unifying morphology aside from small body size. Temnospondyls typically all share a common bauplan but exhibit a substantial disparity of body sizes, although putative lissamphibian outgroups within Temnospondyli are also small-bodied (Fröbisch and Schoch, 2009; Pérez-Ben et al., 2018). Phylogenetic support for the two hypotheses has traditionally been roughly within a statistical margin of error (Ruta and Coates, 2007; Marjanović and Laurin, 2019) with differing implications for both pattern of lissamphibian body plan assembly and timing of the origin of the tetrapod crown group.

Traditionally, both temnospondyls and lepospondyls have been considered early diverging tetrapod clades that originated as part of an early Carboniferous tetrapod diversification. Because of a poor vertebrate record in the earliest Carboniferous, the earliest representative of this diversification has traditionally been the lepospondyl *Lethiscus stocki* (**Figure 3D**; Anderson et al., 2003; Benton et al., 2015), which would be considered a crown tetrapod under either major lissamphibian origins hypothesis (a stem-amniote under the temnospondyl hypothesis or a stem-amphibian under the lepospondyl hypothesis, **Figures 1B,C**), and *Lethiscus stocki* has been therefore conveniently recommended by paleontologists as the appropriate node calibration for the tetrapod crown (Benton and Donoghue, 2006; Benton et al., 2015). However, recent description of tetrapod faunas from earliest Carboniferous fossil deposits (Anderson et al., 2015; Clack et al., 2017, 2018) has identified many taxa within this early diversification that were thought to be characteristic of later Carboniferous or Permian faunas, demanding a more careful consideration of which of these Carboniferous forms belong to the crown group. Recent reanalysis of *Lethiscus* has shown that such reconsideration is not only justified but also necessary, as it shares a number of anatomical features with definitive Devonian stem-tetrapods not seen in the Carboniferous radiation (Pardo et al., 2017b). The earliest temnospondyl, *Balanerpeton* (Milner and Sequeira, 1993) from the Viséan (∼335 MA) of East Kirkton, Scotland, by contrast, is widely accepted in its identification and establishes the temnospondyl (sensu strictu, independent of the placement of colosteids) portion of this dichotomy.

There is some uncertainty in the overall relationships of tetrapod taxa that make up this Carboniferous radiation, but there are some broad patterns. Traditionally, the least inclusive clade including temnospondyls and modern amniotes is thought to include most if not all Carboniferous tetrapod taxa (Ruta et al., 2003; Ruta and Coates, 2007), regardless of whether amphibians originated within Temnospondyli or Lepospondyli (**Figures 1A,B**). Specifically, this clade is thought to include the Embolomeri (**Figure 3C**), a group of large to very large predatory tetrapods, which are typically considered to be more closely related to amniotes than temnospondyls but less closely related to amniotes than lepospondyls (Ruta et al., 2003). It sometimes also includes the Colosteida (**Figure 3B**), a group of aquatic elongate-bodied forms that appears as the sister group of Temnospondyli in some analyses. Thus, in most common formulations, the Temnospondyl hypothesis extends the age of the tetrapod crown group to the age of the oldest embolomere or colosteid, whereas the Lepospondyl hypothesis set the age of the tetrapod crown at the appearance of the earliest lepospondyl, the aïstopod *Lethiscus stocki*, from the middle Visean of Scotland (Marjanović and Laurin, 2013). Because fragmentary embolomere-like and colosteid-like limb elements have been recently reported from the early Tournaisian of Nova Scotia (Anderson et al., 2015), the Temnospondyl hypothesis may implicitly support an age of the tetrapod crown group at the Devonian-Carboniferous boundary.

Furthermore, within a Temnospondyl hypothesis framework, some variation in estimated age of the crown also depends on the phylogenetic position of two problematic taxa: the fragmentary Devonian tetrapod *Tulerpeton curtum* and the Whatcheeriidae,

a group of animals widespread in the lower Carboniferous (Lombard and Bolt, 1995; Clack, 1998; Warren, 2007) but present in the uppermost Devonian (Daeschler et al., 2009; Olive et al., 2016). Both *Tulerpeton* and the whatcheeriids have generally been found on the tetrapod stem in most analyses (Ruta et al., 2003; Clack et al., 2017; Pardo et al., 2017b; Marjanović and Laurin, 2019) but appear on the amniote stem in a subset of studies (e.g., one of the three trees reported by Clack et al., 2017). *Tulerpeton* consists primarily of a single articulated but headless holotype (**Figure 3A**). Because the majority of anatomical data used in phylogenetic analyses are cranial (Ahlberg and Clack, 1998; Pardo et al., 2017b), the phylogenetic placement of *Tulerpeton* depends largely on less-studied anatomy and more general anatomy of the limb elements (Lebedev and Coates, 1995), in particular the unusually shaped humeral and cylindrical femoral shafts. Conversely, *Tulerpeton* exhibits prominent adductor blades on the femur (shared with *Ichthyostega* and *Acanthostega*) and a polydactylous manus (shared with *Acanthostega* and *Ichthyostega*). Whatcheeriids, best typified by *Whatcheeria deltae* from the Visean of Iowa, United States, but also including *Pederpes finneyae* from the Tournaisian of Scotland and *Ossinodus puerhi* from the Visean of Australia, are somewhat better-known than *Tulerpeton*. *Whatcheeria* was first compared with embolomeres, considered by some to be stem-amniotes, on the basis of the deep skull and short postorbital skull table (Lombard and Bolt, 1995), although the authors acknowledged that most of these embolomere-like features only weakly support this placement. However, *Whatcheeria* also preserves many features that are either plesiomorphic or are found only in the stem-tetrapod *Ichthyostega*, including large triangular flanges on the ribs and a buccohypophyseal foramen (Bolt and Lombard, 2018). In resolving the phylogenetic relationships of both of these taxa, there are deep conflicts between character complexes and treatments. These conflicts have major implications for the timing of tetrapod origins: although inclusion of one or more of these taxa in a more derived position that temnospondyls suggests no change in the timing of crown tetrapod origins under the Lepospondyl hypothesis, this would suggest a very deep origin of tetrapods under the Temnospondyl hypothesis, emphasizing a central need to resolve the lissamphibian origins debate in order to inform deeper node calibrations within the tetrapod tree.

This debate itself depends on two major features of tetrapod phylogeny: a monophyletic Lepospondyli that is closely related to amniotes and an early divergence of Temnospondyli within the Late Devonian or early Carboniferous radiation. It increasingly appears that the former is not a settled feature of early tetrapod phylogeny. Recent redescription of a number of recumbirostrans, a clade of lepospondyls part of the previously recognized Order Microsauria (Carroll and Gaskill, 1978), has shown surprisingly reptile-like morphology of the braincase, suspensorium, and lower jaw (Pardo et al., 2015; Szostakiwskyj et al., 2015; Pardo and Anderson, 2016). In contrast, micro–computed tomography study of the aïstopod *Lethiscus stocki*, the earliest lepospondyl, has revealed extremely fishlike organization of the head (Pardo et al., 2017a), suggesting that the diverse morphology of lepospondyls may be a function of polyphyletic origins across the early tetrapod tree rather than a single adaptive radiation. Although this does not exclude the possibility that one lepospondyl group might represent the lissamphibian stem group, the likely polyphyly of lepospondyls means that supporters of the Lepospondyl hypothesis must specify which lepospondyl group they consider most closely related to lissamphibians and must identify node calibration dates accordingly. Regardless, it is unlikely that *Lethiscus* can remain the node calibration. With the exception of the ambiguous *Westlothiana lizzeae* (Smithson et al., 1993) and *Kirktonecta* (Clack, 2011b) and some fragmentary fossils attributed to microsaurs from the Serpukhovian of Goreville, Kentucky (Lombard and Bolt, 1999), there are few Mississippian lepospondyls aside from aïstopods and adelogyrinids, both of which are unlikely to be lissamphibian stem groups. The first unambiguous members of the remaining lepospondyl groups (microsaurs, nectrideans, lysorophians) are earliest Pennsylvanian in age and from the same localities as the earliest amniotes.

The broader patterns of early tetrapod phylogeny may be under dispute as well. In particular, several new lines of evidence suggest that the early Carboniferous tetrapod diversification may be limited to stem-group tetrapod lineages and that the divergence of lissamphibians and amniotes may be substantially more recent, even under the Temnospondyl hypothesis. These lines of evidence come from restudy of colosteids and embolomeres themselves and suggest an emergence of both taxa within the Devonian radiation of early tetrapods, prior to tetrapod terrestrialization. In colosteids, this has come from new comprehensive studies that have found that the colosteid skull and jaw retain many bones lost in more advanced taxa and that similarities with temnospondyls are likely superficial (Bolt and Lombard, 2001, 2010). In contrast, studies addressing embolomere anatomy have remained relatively restricted in anatomical scope. Embolomeres have often been considered early representatives of the lineage leading to amniotes, based on the deep narrow skull and large size of the Meckelian foramen in the lower jaw, among other features (Carroll, 1970). However, recent work has identified substantial conflicts between anatomical suites, suggesting that reconsideration of this scenario is necessary. Most notably, Clack (2011a) identified the presence of dermal fin rays (lepidotrichia) and bony supports (supraneural radials) in the caudal fin of a partial embolomere tail and likely presence of supraneural articulations in other more complete embolomeres, but did not address whether this would suggest an earlier divergence of embolomeres within tetrapods or a reversal in this one species. Pardo et al. (2018) drew several comparisons between the skull and braincase of aïstopods and embolomeres and identified evidence for articulation between the dorsal branchial skeleton and the otoccipital regions of both taxa. As a dorsal branchial skeleton is thought to be retained only through the fin-to-limb transition in tetrapods (Coates and Clack, 1991), this would provide further evidence for placing embolomeres on the tetrapod stem, regardless of one's hypothesis of lissamphibian origins. Indeed, recent phylogenetic treatment of endocranial data from these and other early tetrapods has found increased evidence for a closer relationship between temnospondyls and amniotes to the exclusion of both colosteids and embolomeres (Pardo et al., 2017a), possibly

indicating a much more exclusive crown group under the temnospondyl hypothesis.

It is important to note that the proceeding discussion relates to only one view of overall tetrapod phylogeny, the tree of Ruta et al. (2003) and Ruta and Coates (2007), but other hypotheses similarly struggle with these issues. Another hypothesis, that of Smithson (1985), posits a deep divergence between reptiles and lissamphibians, with reptiles descending from a long lineage from embolomeres to anthracosaurs called Reptilomorpha (note, this concept differs from that defined phylogenetically by Laurin, 2001; Vallin and Laurin, 2004). Given this hypothesis (which has not been supported by the largest computer assisted analyses conducted to date but has had some support from more limited treatments, such as Ruta and Clack, 2006), the split between reptilomorphs and batrachomorphs (the lissamphibian stem group) would be placed at least into the Viséan and possibly extend possibly into the Devonian, should *Tulerpeton* (not included in the analysis of Ruta and Clack, 2006, which they state was "not intended as an exhaustive investigation of early tetrapod relationships" [p. 49]) prove to be an embolomere.

NODE MAXIMA: IS THE EARLY TETRAPOD RECORD COMPLETE ENOUGH TO RELY ON NODE CALIBRATIONS?

Assignment of maxima (hard or soft) depends on confidence in the quality of the fossil record. Reviews suggesting hard and soft maxima for major tetrapod clades (e.g., Benton et al., 2015) generally point to faunas entirely lacking any members of these groups. Assignment of hard maxima must contend with the understanding that absence of evidence is not evidence of absence, but that continued absence after sufficient sampling may provide a degree of confidence in absence. Sampling of a crown group fossil within a fossil collection requires that four criteria are met:

(1) the crown-group animal has the same (or better) probability of being preserved in the fossil localities sampled in comparison with outgroups;
(2) known fossil localities sample the kind of local habitats where the crown-group animals lived and died;
(3) known fossil localities sample the biogeographic provinces where crown-group animals were distributed; and
(4) sampling effort is sufficient within an interval and region to assume that the crown-group animal would have been found if it were present, which itself is a function of species prevalence (e.g., Hedman, 2010).

If all four of these conditions are met and no members of the crown group are identified, it becomes more reasonable to infer that the crown group may not have originated by a specific interval. This presents a substantial challenge: when is a representative of a crown group absent from a collection or fauna because it did not exist at the time, and when is it absent from a collection or fauna because one or more of these four conditions has not been met?

Benton et al. (2015) provide three distinct justifications for soft maxima for the three major tetrapod clades reviewed here. The soft maximum for crown amniotes is set at 332.9 Ma based on the absence of crown amniotes at the East Kirkton locality within the Visean of Scotland. The soft maximum for the amphibian crown group is set at the base of the Middle Permian (272.8 Ma) and based on the absence of definitive stem amphibians in the Middle and Upper Permian rocks of Russia, China, and South Africa. The soft maximum for crown tetrapods is set at the middle Tournaisian (351 Ma) based on the presence of the whatcheeriids *Pederpes finneyae* and *Whatcheeria deltae* in Scotland and North America, respectively.

Already it should be apparent that some of these soft maxima are substantially younger than the age of the oldest member of the crown group according to different phylogenetic hypotheses. For example, if we accept that the temnospondyl *Gerobatrachus hottoni* is in fact a stem-group batrachian following Anderson et al. (2008) and Maddin et al. (2012), then the hard minimum age of the amphibian crown group must be the age of *Gerobatrachus*, which is no younger than 272.8 Ma and likely closer to 276.2 Ma, the maximum age of the Tubb Sandstone inferred by U-Pb dating of detrital zircons (Liu and Stockli, 2020). The type locality of *Gerobatrachus hottoni* is in the informal "Cedar Top Sandstone" unit of the Middle Clear Fork Group (R. Hook, pers. comm.), which sits below the Tubb Sandstone (Nelson et al., 2013a,b). The recent suggestion by Pardo et al. (2017b) that the amphibian crown group may be even more inclusive would set the hard minimum at a substantially older age (∼315 Ma) over 40 million years older than the soft maximum of Benton et al. (2015). Disagreements exist in both directions for the age of the tetrapod crown group; some studies of early tetrapod phylogeny suggest that the hard minima for the tetrapod crown group is older than the soft maximum offered by Benton et al. (2015), although recent phylogenetic analyses relying on more sophisticated treatment of endocranial anatomy suggest a much younger age of the tetrapod crown more generally (Pardo et al., 2017a, 2019; Pardo and Mann, 2018).

Additionally, the soft maxima suggested for major tetrapod clades may conflict with the criteria outlined above. These conflicts are themselves a combination of overlapping deficiencies in the vertebrate fossil record. These deficiencies are a product of systematic preservation and sampling biases in space and time and correspond to areas of great uncertainty in the fossil record of major tetrapod clades.

Preservational Heterogeneity and Small Body Size

The assembly of both the amphibian crown group and the amniote crown group are thought to have largely occurred at small body sizes (Carroll, 1982; Laurin, 2004; Kemp, 2007; Pérez-Ben et al., 2018), although the overall pattern of body size evolution in these clades is under some debate (Didier et al., 2019). Skeletal material from small vertebrates degrades more quickly than bones of larger vertebrates and is preferentially lost from the record (Behrensmeyer et al., 1979). This creates a set of related patterns that have the potential to preferentially

deplete the fossil record of early members of major tetrapod crown groups. This is apparent in the particularly large gaps in the caecilian fossil record (Evans and Sigogneau-Russell, 2001; Jenkins et al., 2007). First, this would suggest that early members of the amphibian and amniote crown groups may be expected to be absent across much of the early tetrapod fossil record even if they were present at the time. Furthermore, this would suggest that what remnants of early crown-group amphibians, amniotes, and perhaps even crown tetrapods may also be preferentially more degraded, reducing the ability of specialists to identify isolated elements of small-bodied early tetrapods to higher taxon. Finally, this would suggest that early representatives of both amphibian and amniote stem groups (and potentially crown groups) can be expected to reflect distribution of localities with exceptional preservation.

Exceptional preservation in the fossil record is itself a function of several properties of specific local or regional depositional environments. Rapid burial in an anoxic reducing environment is generally a prerequisite for exceptional preservation and is most typical of environments with standing water. This is the case for the majority of exceptional vertebrate-bearing fossil localities across the late Paleozoic, which include anoxic organic-rich oxbow lakes (e.g., Hook and Baird, 1986; Clarkson et al., 1993), largely anoxic graben lake deposits, and shallow brackish lagoons (e.g., Clements et al., 2019). Aside from sites of exceptional preservation, small vertebrate skeletons may be concentrated but remain relatively undisturbed in very specific circumstances, such as within fissure fill deposits such as the Fort Sill locality (MacDougall et al., 2017) and the classic Joggins lycopod stump localities (Falcon-Lang et al., 2006). These latter types of localities are exceedingly rare and preserve a unique fauna, but the deposition in anoxic lacustrine or estuarine systems tends to be repeated where similar environmental conditions are present (Baird et al., 1985). This means that trends in both regional paleoenvironment and global paleoclimate may bias discovery probability in a given region or interval.

What does this mean for the probability of discovery of small-bodied tetrapods across the Late Paleozoic? Although this has not been investigated for the entire Paleozoic tetrapod record, regional trends have been investigated for the interval spanning from 315 to 272 Ma and found that the vertebrate record samples these sorts of environments well only in the late Carboniferous, and this sampling of these environments reflects regional variation in climate change across the late Paleozoic (Pardo et al., 2019). Such environments are poorly sampled outside of Europe and North America in this interval, and in fact are essentially completely unsampled in the "classic" Upper Permian sequence of South Africa until the earliest Triassic. This presents a dual challenge in discovering the earliest representatives of the amniote and amphibian crown groups. First, the earliest amniotes were certainly highly terrestrial (Carroll, 1964; Laurin and de Buffrénil, 2016), and it appears that the earliest amphibians may have been as well (Pardo et al., 2019), and therefore lived in habitats that may have been spatially separated from ideal preservational environments. This means that early amniotes and amphibians are likely rare even among rare vertebrate fossils and will likely not be seen in localities without extensive worker effort. Second, the probability of discovery is likely limited by the abundance of these sorts of localities. Within the Lower Carboniferous, only three localities contain this level or preservation: the nearshore marine Wardie Shale (Trojan et al., 2015) and Cheese Bay Shrimp Bed (Hesselbo and Trewin, 1984), and the thermally altered lake deposits at East Kirkton (Clarkson et al., 1993). Of these, the only locality that has yielded more than a single tetrapod fossil is East Kirkton. East Kirkton also represents the only definitive Lower Carboniferous occurrences of crown tetrapods within recent phylogenetic analyses (Milner and Sequeira, 1993).

Spatiotemporal Heterogeneity

In addition to the sampling biases imposed by small body size, there are several broader patterns of the early tetrapod fossil record that may also preferentially obscure the early records of crown tetrapods, crown amphibians, and crown amniotes. In particular, the early tetrapod record is itself relatively heterogeneous in both space and time (**Figure 4**). The fossil record of tetrapods from the Late Devonian until the Middle Permian is almost entirely restricted to localities from North America and Europe (Milner, 1993). At the time, this represented a single continental landmass restricted to within 10 degrees of the equator (**Figure 4A**). A few localities exist outside of this narrow equatorial band in the Devonian, Carboniferous, and Permian, but taxonomic diversity and worker effort remain far lower in these regions than in Euramerica, particularly in Gondwana, which had not fully joined the Pangaean supercontinent until the late Carboniferous to early Permian (Ziegler et al., 1979; Stampfli et al., 2013). A robust record outside of Euramerica does not appear until the latter part of the Middle Permian (**Figures 4A,B**), at which point concurrent well-sampled records appear in the Karoo Basin of South Africa (Rubidge et al., 2013), the Paraná Basin of southern Brazil (Dias-da-Silva, 2012), and a series of basins across Russia and China (Olroyd and Sidor, 2017). These faunas preserve very different vertebrate communities dominated by diverse and abundant derived mammal-line synapsids not observed in Euramerica.

This would not be a substantial problem if early tetrapods did not show any substantial biogeographic patterns. However, where faunas outside of the Euramerican transect are known, they contrast in some important ways with contemporary Euramerican faunas, preserving both unique lineages and extremely early records of Late Permian taxa (Milner and Panchen, 1973; Cisneros et al., 2015). Carboniferous-Permian faunas from central Asia are dominated by seymouriamorphs (interpreted as stem amniotes under most phylogenetic hypotheses) but completely lack representatives of classic Euramerican taxa until well into the Middle Permian (Reisz and Laurin, 2001). Interestingly, central Asia appears to be the epicenter of another putative stem-amniote lineage, the Chroniosuchia, which appears to be completely absent from the Carboniferous-Permian transition of Euramerica, as well as later sequences across Gondwana (Golubev, 1998; Witzmann et al., 2008). Conversely, new localities from the Carboniferous-Permian transition of Gondwana seem to preserve a tetrapod community that is roughly similar to

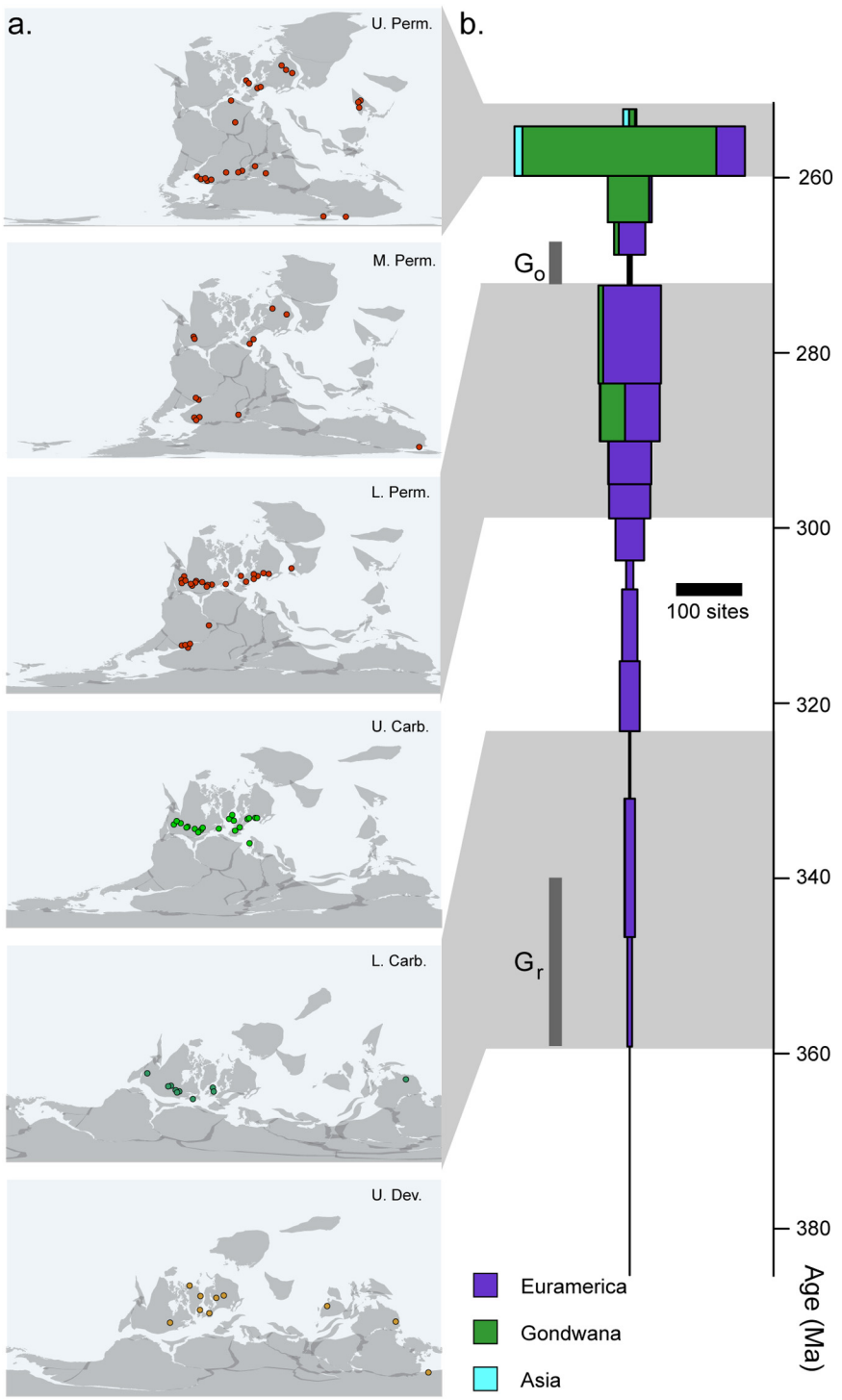


FIGURE 4 | Completeness of the early tetrapod record. **(A)** Geographic distribution of tetrapod fossil localities across the late Palaeozoic. **(B)** Number of tetrapod-bearing localities recorded in the Paleobiology Database (http://paleobioDB.org) by stage, showing heterogeneity in regional and temporal sampling.

Euramerican faunas but include unique Permo-Triassic-like components, including rhinesuchid stereospondyls and advanced lungfishes (Cisneros et al., 2015), further hinting at important biogeographic patterns, either regional endemism or differences in paleoenvironments. These biogeographic patterns are not restricted to the fossil record; distribution of extant amphibians suggests a clear Gondwanan origin of crown-group caecilians and Laurasian origin of salamanders, with modern representatives

essentially restricted to these regions (recently summarized in Pyron, 2014). Given this substantial provincialization, it would seem at least possible that the origin of some major tetrapod clades may have occurred in a biogeographical province not currently represented in the tetrapod record. The most obvious candidate group would be amniotes, which are already highly diverse at the time of their first appearance and are not preceded by an unambiguous stem group. An origin of amniotes in a Central Asian epicenter would appear plausible and has been suggested at least obliquely in qualitative studies of the tetrapod record at this time (Milner and Panchen, 1973).

Furthermore, the early tetrapod record is punctuated by several key intervals of minimal sampling of the tetrapod record (**Figure 4B**). Two of these are particularly noteworthy: an 18-million-year interval in the lower Carboniferous, spanning from the Devonian-Carboniferous boundary (358.9 Ma) until the middle Viséan (∼330.9 Ma), and a second within the first half of the middle Permian (272–265 Ma). The former is generally referred to as Romer's Gap and likely coincides with the origin of the tetrapod crown group, whereas the second, referred to as Olson's Gap, coincides with a major faunal turnover between Carboniferous-Permian transition faunas dominated by archaic tetrapods and early amniotes and Upper Permian faunas dominated by diverse therapsid-grade stem mammals, but also spans an interval that may represent the assembly of distinct lissamphibian body plans (Marjanović and Laurin, 2007; Anderson et al., 2008; Pardo et al., 2017a). It has been suggested previously that this transition, at least among synapsids, represents a physiological shift in response to the rapidly changing environment (Kemp, 2006). Although some work has been done in recent years to bridge these sampling gaps (Smithson et al., 2012; Anderson et al., 2015, Clack et al., 2017), these remain relatively unexplored intervals, and it is not possible to assess at this time whether the absence of identified fossils of key informative taxa (Mississippian crown-group tetrapods, middle Permian crown-group lissamphibians) represents a real absence from these faunas. Further intervals are also substantially undersampled in addition to these historical "gaps." This extends throughout the Lower Carboniferous (**Figure 4B**), where sampling effort is not only very poor but is extremely geographically restricted (**Figure 4A**). Given that this interval appears to contain, at the very least, the origin of both crown-group tetrapods and crown-group amniotes, confidently applying limits to age estimates for these nodes is likely impossible.

DISCUSSION AND RECOMMENDATIONS

Attempts to establish *a priori* constraints for major tetrapod clade ages must contend with two parallel problems: there is little agreement on the inclusiveness of these clades, and the early tetrapod record is so unevenly sampled that we cannot assume representative sampling of early members of these clades. The result is that two of the three nodes assessed here exhibit substantial variation in *a priori* calibration ages based on phylogenetic hypothesis, with a range of credible estimates spanning over 30 million years for *a priori* calibrations of the tetrapod crown and over 70 million years for *a priori* calibrations of the lissamphibian crown. Recent work has suggested that the earliest representatives of the tetrapod crown group may be substantially younger than previously thought, whereas new fossils and hypotheses may support substantially older calibration ages for the lissamphibian crown group than previously appreciated. These depend on three major points of phylogenetic disagreement (amphibian origins within early tetrapods, caecilian origins within total-group amphibians, and delimitation of the tetrapod crown group) that will likely remain under debate for some time in the future, but workers calibrating deep nodes in the tetrapod tree should be prepared to take these ages into account.

Conversely, the *a priori* constraints on the age of the first representatives of the amniote crown are relatively robust to phylogenetic disagreement. However, the earliest amniotes appear after a long interval of poor sampling, and early members of the amniote total group show extremely poor stratigraphic concordance, with members of the amniote stem appearing millions of years later than the earliest crown-group amniotes. One explanation for this problem is that the faunas in which amniotes originated are unsampled within the early Carboniferous. This can be attributed to multiple factors: (1) that early Carboniferous localities heavily sample aquatic habitats, but only poorly sample dryland terrestrial environments (e.g., Pardo et al., 2019); (2) that amniotes originated in a biogeographic region outside of and with limited connectivity to Euramerica prior to the Late Carboniferous; (3) or a combination of both explanations. Such a hypothesis would not necessarily be limited to amniotes; crown tetrapods in general seem to have appeared abruptly at the end of the Early Carboniferous within a relatively brief 20-million-year interval, with high levels of terrestriality seen across the tetrapod crown group in general (Pardo et al., 2019). One suggested location for this biogeographic province would be the Kazakh plate that now forms much of central Asia and that is home to a uniquely diverse putative stem amniote assemblage in the early Permian (Milner and Panchen, 1973), but this evidence remains highly circumstantial, given that no early amniotes are known from this province and the seymouriamorph-dominated assemblages appear to be younger than the earliest amniote-dominated assemblages, such as Joggins, from Euramerica. There is little direct evidence for any such phylogeographic structure of Carboniferous tetrapod assemblages without new sampling from the Carboniferous of Gondwana and Asia, as well as more aggressive sampling within the interval roughly between 320 and 340 Ma.

This problem might be resolvable if molecular clock estimates converge on a tight estimate of the origins of major tetrapod clades with tight correspondence to a subset of hypotheses. This has been argued from both the molecular (San Mauro, 2010) and paleontological (Marjanović and Laurin, 2007) perspectives. However, we find no such tight correspondence. In fact, the dispersion of molecular clock estimates broadly compares with the dispersion of possible calibration dates, in that the estimated

ages of the tetrapod and amphibian crown groups are difficult to constrain, whereas the amniote crown group is more tightly constrained. As most molecular clock analyses have used the node calibrations of Benton and Donoghue (2006) and Benton et al. (2015) for the tetrapod and amphibian crown groups, it seems likely that the uncertainty is a function of the poor early tetrapod (and amphibian) fossil record, which interacts with variation in taxonomic and molecular sampling and model parameterization to produce highly volatile estimates.

Recommendations for best practices in calibrating nodes in molecular clock studies have been previously made by Parham et al. (2012), with a focus on ensuring that calibration ages are replicable by tying the age to a specimen and stratigraphic horizon. However, these recommendations generally do not provide guidance for dealing with the problems we have identified here in calibrating nodes in the Palaeozoic and earlier. With this in mind, we urge that those completing studies calibrating deep tetrapod nodes, as well as other deep nodes, to keep the following in mind:

(1) Compendia of node age calibrations, such as those of Benton and Donoghue (2006) and Benton et al. (2015) may misrepresent confidence in node age calibrations from the Palaeozoic by understating disagreement between specialists on the underlying phylogeny and even anatomy. This is particularly problematic for Palaeozoic and pre-Palaeozoic calibrations where anatomical evidence from modern representatives of clades may be scarce and where the record may be generally poor.
(2) Molecular clock studies relying on deep tetrapod node calibrations should be cognizant of disagreements in phylogenetic analyses and should try as much as possible to incorporate this uncertainty where possible. Because possible ages of the tetrapod and lissamphibian crown groups vary so much, depending on specific phylogenetic hypotheses, we strongly recommend conducting multiple independent calibrations rather than adjusting hard minima and soft maxima to include the full range.
(3) In cases where a single tree and a single set of node calibrations are used, authors must explicitly state and defend the phylogenetic hypothesis used to generate those calibrations in terms of confidence in the underlying tree and its associated hypotheses of body plan evolution. Some datasets may be easier to adopt into total evidence approaches, but differences in total number of characters, total number of fossil operational taxonomic units, or degree of taxonomic overlap with molecular datasets do not necessarily reflect confidence in the underlying topology among specialists.
(4) Application of more precise calibration approaches (e.g., tip dating and fossilized birth-death models) cannot be considered a replacement for satisfactorily resolving phylogenetic uncertainty in the origin of the tetrapod and amphibian crown groups.
(5) Tip-dating approaches should not be used as an independent assessment of the quality of priors, including node calibration priors or tree priors. Regardless of the arguments for or against the use of tip-dating methods for assessing quality of priors, the early tetrapod record is highly heterogeneous both in terms of the observed pattern of preservation and the taxonomic expectation of preservation. It is therefore likely that the tetrapod record will violate certain assumptions of tip-dating approaches unless appropriately parameterized.
(6) Early tetrapod workers need to bring their attention to undersampled intervals and regions. Establishing a tetrapod fossil record from the Carboniferous and Early Permian of Asia and Gondwana is of particular importance. Terrestrial rocks are known from these regions, in some cases preserving fossils of other vertebrate groups (actinopterygians, chondrichthyans, and dipnoans) and plants, but tetrapods from these rocks are essentially unknown.

These best practices can be applied more generally to efforts to calibrate nodes prior to the end of the Palaeozoic, as many of the same principles apply to phylogenetic problems among Palaeozoic organisms more generally (difficulty relating extant phylogenetic patterns in anatomy to earliest fossil relatives, preservational biases, temporospatial megabiases, etc.). We do caution that the particular problems we identify here with relating the early tetrapod record to the origin of major tetrapod clades may not directly correspond to problems in other groups, although conceptual similarities almost certainly exist. Workers attempting to calibrate these nodes should exercise caution and seek direct consultation with experts on relevant parts of the fossil record.

AUTHOR CONTRIBUTIONS

JP led the project and created the images. KL generated new data on nectrideans. JA underwrote the project and provided CT access. All authors wrote the manuscript.

FUNDING

This research was supported in part by NSERC Discovery Grant 17-04821 awarded to JA.

ACKNOWLEDGMENTS

We thank C. Mansky, T. Fedak, and B. Simpson for access to specimens. We also thank A. Milner, J. Cisneros, C. Marsicano, A. Huttenlocker, A. Mann, B. Gee, D. Marjanović, P. Ahlberg, R. Hook, W. DiMichele, P. Kroehler, and the late J. Clack for helpful discussions, which informed the ideas presented here. We especially thank M. Laurin, R. Warnock, and G. Didier for their invitation to participate in this volume. M. Laurin and two reviewers provided insightful comments which aided us in preparing this manuscript.

REFERENCES

Ahlberg, P. E., and Clack, J. A. (1998). Lower jaws, lower tetrapods–a review based on the Devonian genus Acanthostega. *Earth Environ. Sci. Trans Royal Soc. Edinburgh* 89, 11–46. doi: 10.1017/s0263593300002340

Anderson, J. S. (2008). Focal review: the origin (s) of modern amphibians. *Evol. Biol.* 35, 231–247. doi: 10.1007/s11692-008-9044-5

Anderson, J. S., Carroll, R. L., and Rowe, T. B. (2003). New information on *Lethiscus stocki* (Tetrapoda: Lepospondyli: Aistopoda) from high-resolution computed tomography and a phylogenetic analysis of Aistopoda. *Can. J. Earth Sci.* 40, 1071–1083. doi: 10.1139/e03-023

Anderson, J. S., Reisz, R. R., Scott, D., Fröbisch, N. B., and Sumida, S. S. (2008). A stem batrachian from the early permian of texas and the origin of frogs and salamanders. *Nature* 453: 515–518. doi: 10.1038/nature06865

Anderson, J. S., Smithson, T., Mansky, C. F., Meyer, T., and Clack, J. (2015). A diverse tetrapod fauna at the base of 'Romer's Gap'. *PLoS One* 10:e0125446. doi: 10.1371/journal.pone.0125446

Ascarrunz, E., Rage, J. C., Legreneur, P., and Laurin, M. (2016). Triadobatrachus massinoti, the earliest known lissamphibian (Vertebrata: Tetrapoda) re-examined by μCT scan, and the evolution of trunk length in batrachians. *Contributions to Zool.* 85, 201–234. doi: 10.1163/18759866-08502004

Baird, G. C., Sroka, S. D., Shabica, C. W., and Beard, T. L. (1985). Mazon Creek-type fossil assemblages in the US midcontinent Pennsylvanian: their recurrent character and palaeoenvironmental significance. *Phil. Trans. Royal Soc. London. B, Biol. Sci.* 311, 87–99. doi: 10.1098/rstb.1985.0141

Behrensmeyer, A. K., Western, D., and Boaz, D. E. D. (1979). New perspectives in vertebrate paleoecology from a recent bone assemblage. *Paleobiology* 5, 12–21. doi: 10.1017/s0094837300006254

Benson, R. B. (2012). Interrelationships of basal synapsids: cranial and postcranial morphological partitions suggest different topologies. *J. Systematic Palaeontol.* 10, 601–624. doi: 10.1080/14772019.2011.631042

Benton, M. J. (1984). Consensus on archosaurs. *Nature* 312:599. doi: 10.1038/312599a0

Benton, M. J., and Donoghue, P. C. (2006). Paleontological evidence to date the tree of life. *Mol. Biol. Evol.* 24, 26–53. doi: 10.1093/molbev/msl150

Benton, M. J., Donoghue, P. C., Asher, R. J., Friedman, M., Near, T. J., and Vinther, J. (2015). Constraints on the timescale of animal evolutionary history. *Palaeontol. Electron.* 18, 1–116. doi: 10.1006/anbe.1999.1287

Berman, D. S. (2000). Origin and early evolution of the amniote occiput. *J. Paleontol.* 74, 938–956. doi: 10.1666/0022-3360(2000)074<0938:oaeeot>2.0.co;2

Bever, G. S., Lyson, T. R., Field, D. J., and Bhullar, B. A. S. (2015). Evolutionary origin of the turtle skull. *Nature* 525, 239–242. doi: 10.1038/nature14900

Blackburn, D. C., Roberts, E. M., and Stevens, N. J. (2015). The earliest record of the endemic African frog family ptychadenidae from the oligocene nsungwe formation of Tanzania. *J. Vertebrate Paleontol.* 35:e907174. doi: 10.1080/02724634.2014.907174

Bolt, J. R., and Lombard, R. E. (2001). The mandible of the primitive tetrapod greererpeton, and the early evolution of the tetrapod lower jaw. *J. Paleontol.* 75, 1016–1042. doi: 10.1017/s0022336000039913

Bolt, J. R., and Lombard, R. E. (2010). *Deltaherpeton hiemstrae*, a new colosteid tetrapod from the Mississippian of Iowa. *J. Paleontol.* 84, 1135–1151. doi: 10.1666/10-020.1

Bolt, J. R., and Lombard, R. E. (2018). Palate and braincase of whatcheeria deltae Lombard & Bolt, 1995. *Earth Environ. Sci. Trans. Royal Soc. Edinburgh* 109, 1–24.

Brocklehurst, N., Reisz, R. R., Fernandez, V., and Fröbisch, J. (2016). A re-description of 'Mycterosaurus' smithae, an Early Permian eothyridid, and its impact on the phylogeny of pelycosaurian-grade synapsids. *PLoS One* 11:e0156810. doi: 10.1371/journal.pone.0156810

Cannatella, D. (2015). Xenopus in space and time: fossils, node calibrations, tip-dating, and paleobiogeography. *Cytogenet. Genome Res.* 145, 283–301. doi: 10.1159/000438910

Carroll, R. (1966). "Microsaurs from the westphalian b of joggins, nova scotia," in *Proceedings of the Linnean Society of London*, (Oxford: Oxford University Press), 63–97. doi: 10.1111/j.1095-8312.1966.tb00952.x

Carroll, R. L. (1964). The earliest reptiles. *Zool. J. Linnean Soc.* 45, 61–83. doi: 10.1111/j.1096-3642.1964.tb00488.x

Carroll, R. L. (1970). The ancestry of reptiles. *Phil. Trans. Royal Soc. of London. B, Biol. Sci.* 257, 267–308. doi: 10.1098/rstb.1970.0026

Carroll, R. L. (1982). Early evolution of reptiles. *Annu. Rev. Ecol. Syst.* 13, 87–109. doi: 10.1146/annurev.es.13.110182.000511

Carroll, R. L. (2001). The origin and early radiation of terrestrial vertebrates. *J. Paleontol.* 75, 1202–1213. doi: 10.1666/0022-3360(2001)075<1202:toaero>2.0.co;2

Carroll, R. L. (2007). The Palaeozoic ancestry of salamanders, frogs and caecilians. *Zool. J. Linnean Soc.* 150(Suppl._1), 1–140. doi: 10.1111/j.1096-3642.2007.00246.x

Carroll, R. L., and Gaskill, P. (1978). *The Order Microsauria*, Vol. 126. Philadelphia, PA: American Philosophical Society, 211.

Chaboureau, A. C., Sepulchre, P., Donnadieu, Y., and Franc, A. (2014). Tectonic-driven climate change and the diversification of angiosperms. *Proc. Natl. Acad. Sci. U.S.A.* 111, 14066–14070. doi: 10.1073/pnas.1324002111

Chen, M. Y., Liang, D., and Zhang, P. (2015). Selecting question-specific genes to reduce incongruence in phylogenomics: a case study of jawed vertebrate backbone phylogeny. *Systematic Biol.* 64, 1104–1120. doi: 10.1093/sysbio/syv059

Chiari, Y., Cahais, V., Galtier, N., and Delsuc, F. (2012). Phylogenomic analyses support the position of turtles as the sister group of birds and crocodiles (Archosauria). *BMC Biol.* 10:65. doi: 10.1186/1741-7007-10-65

Cisneros, J. C., Marsicano, C., Angielczyk, K. D., Smith, R. M., Richter, M., Fröbisch, J., et al. (2015). New Permian fauna from tropical Gondwana. *Nat. Commun.* 6:8676.

Clack, J. A. (1998). A new early carboniferous tetrapod with a mélange of crown-group characters. *Nature* 394, 66–69. doi: 10.1038/27895

Clack, J. A. (2011a). A Carboniferous embolomere tail with supraneural radials. *J. Vertebrate Paleontol.* 31, 1150–1153. doi: 10.1080/02724634.2011.595467

Clack, J. A. (2011b). A new microsaur from the Early Carboniferous (Viséan) of East Kirkton. Scotland, showing soft tissue evidence. *Special Papers Paleontol. Stud. Fossil Tetrapods* 86, 45–55.

Clack, J. A., Bennett, C. E., Carpenter, D. K., Davies, S. J., Fraser, N. C., Kearsey, T. I., et al. (2017). Phylogenetic and environmental context of a Tournaisian tetrapod fauna. *Nat. Ecol. Evol.* 1:0002. doi: 10.1038/s41559-016-0002

Clack, J. A., Porro, L. B., and Bennett, C. E. (2018). A Crassigyrinus-like jaw from the Tournaisian (Early Mississippian) of Scotland. *Earth Environ. Sci. Trans. Royal Soc. Edinburgh* 108, 37–46. doi: 10.1017/s1755691018000087

Clarkson, E. N. K., Milner, A. R., and Coates, M. I. (1993). Palaeoecology of the Viséan of East Kirkton. West Lothian, Scotland. *Earth Environ. Sci. Trans. Royal Soc. Edinburgh* 84, 417–425. doi: 10.1017/s0263593300006210

Clements, T., Purnell, M., and Gabbott, S. (2019). The mazon creek lagerstätte: a diverse late Paleozoic ecosystem entombed within siderite concretions. *J. Geol. Soc.* 176, 1–11. doi: 10.1144/jgs2018-088

Coates, M. I., and Clack, J. A. (1991). Fish-like gills and breathing in the earliest known tetrapod. *Nature* 352, 234–236. doi: 10.1038/352234a0

Coates, M. I., Ruta, M., and Friedman, M. (2008). Ever since Owen: changing perspectives on the early evolution of tetrapods. *Ann. Rev. Ecol. Evol. Systemat.* 39, 571–592. doi: 10.1146/annurev.ecolsys.38.091206.095546

Daeschler, E. B., Clack, J. A., and Shubin, N. H. (2009). Late Devonian tetrapod remains from Red Hill, Pennsylvania, USA: how much diversity? *Acta Zoologica* 90, 306–317. doi: 10.1111/j.1463-6395.2008.00361.x

Dias-da-Silva, S. (2012). Middle–Late Permian tetrapods from the Rio do Rasto Formation, southern Brazil: a biostratigraphic reassessment. *Lethaia* 45, 109–120. doi: 10.1111/j.1502-3931.2011.00263.x

Didier, G., Chabrol, O., and Laurin, M. (2019). Parsimony-based test for identifying changes in evolutionary trends for quantitative characters: implications for the origin of the amniotic egg. *Cladistics* 35, 576–599. doi: 10.1111/cla.12371

Dilkes, D. (1998). The early Triassic rhynchosaur Mesosuchus browni and the interrelationships of basal archosauromorph reptiles. *Trans. Royal Soc.* 353, 501–541. doi: 10.1098/rstb.1998.0225

Duchêne, S., Lanfear, R., and Ho, S. Y. (2014). The impact of calibration and clock-model choice on molecular estimates of divergence times. *Mol. Phylogenet. Evol.* 78, 277–289. doi: 10.1016/j.ympev.2014.05.032

Evans, S. E. (1984). The classification of the Lepidosauria. *Zool. J. Linnean Soc.* 82, 87–100. doi: 10.1111/j.1096-3642.1984.tb00537.x

Evans, S. E., Lally, C., Chure, D. C., Elder, A., and Maisano, J. A. (2005). A late Jurassic salamander (Amphibia: Caudata) from the Morrison formation of

North America. *Zool. J. Linnean Soc.* 143, 599–616. doi: 10.1111/j.1096-3642.2005.00159.x

Evans, S. E., and Sigogneau-Russell, D. (2001). A stem-group caecilian (Lissamphibia: Gymnophiona) from the Lower Cretaceous of North Africa. *Palaeontology* 44, 259–273. doi: 10.1111/1475-4983.00179

Ezcurra, M. D. (2016). The phylogenetic relationships of basal archosauromorphs, with an emphasis on the systematics of proterosuchian archosauriforms. *PeerJ* 4:e1778. doi: 10.7717/peerj.1778

Falcon-Lang, H. J., Benton, M. J., Braddy, S. J., and Davies, S. J. (2006). The Pennsylvanian tropical biome reconstructed from the joggins formation of Nova Scotia, Canada. *J. Geol. Soc.* 163, 561–576. doi: 10.1144/0016-764905-063

Feng, Y. J., Blackburn, D. C., Liang, D., Hillis, D. M., Wake, D. B., Cannatella, D. C., et al. (2017). Phylogenomics reveals rapid, simultaneous diversification of three major clades of Gondwanan frogs at the Cretaceous–Paleogene boundary. *Proc. Natl. Acad. Sci. U.S.A.* 114, E5864–E5870.

Field, D. J., Bercovici, A., Berv, J. S., Dunn, R., Fastovsky, D. E., Lyson, T. R., et al. (2018). Early evolution of modern birds structured by global forest collapse at the end-cretaceous mass extinction. *Curr. Biol.* 28, 1825–1831. doi: 10.1016/j.cub.2018.04.062

Field, D. J., Gauthier, J. A., King, B. L., Pisani, D., Lyson, T. R., and Peterson, K. J. (2014). Toward consilience in reptile phylogeny: miRNAs support an archosaur, not lepidosaur, affinity for turtles. *Evol. Dev.* 16, 189–196. doi: 10.1111/ede.12081

Ford, D. P., and Benson, R. B. (2018). A redescription of Orovenator mayorum (Sauropsida, Diapsida) using high-resolution μ CT, and the consequences for early amniote phylogeny. *Papers Palaeontol.* 5, 197–239. doi: 10.1002/spp2.1236

Ford, D. P., and Benson, R. B. (2020). The phylogeny of early amniotes and the affinities of Parareptilia and Varanopidae. *Nat. Ecol. Evol.* 4, 57–65. doi: 10.1038/s41559-019-1047-3

Fröbisch, N. B., Carroll, R. L., and Schoch, R. R. (2007). Limb ossification in the Paleozoic branchiosaurid Apateon (Temnospondyli) and the early evolution of preaxial dominance in tetrapod limb development. *Evol. Dev.* 9, 69–75. doi: 10.1111/j.1525-142x.2006.00138.x

Fröbisch, N. B., and Schoch, R. R. (2009). Testing the impact of miniaturization on phylogeny: paleozoic dissorophoid amphibians. *Systemat. Biol.* 58, 312–327. doi: 10.1093/sysbio/syp029

Fröbisch, N. B., and Shubin, N. H. (2011). Salamander limb development: integrating genes, morphology, and fossils. *Dev. Dyn.* 240, 1087–1099. doi: 10.1002/dvdy.22629

Gao, K. Q., and Shubin, N. H. (2003). Earliest known crown-group salamanders. *Nature* 422, 424–428. doi: 10.1038/nature01491

Gauthier, J., Kluge, A. G., and Rowe, T. (1988). Amniote phylogeny and the importance of fossils. *Cladistics* 4, 105–209. doi: 10.1111/j.1096-0031.1988.tb00514.x

Gee, B. M., Bevitt, J. J., Garbe, U., and Reisz, R. R. (2019). New material of the 'microsaur' Llistrofus from the cave deposits of Richards Spur, Oklahoma and the paleoecology of the Hapsidopareiidae. *PeerJ* 7:e6327. doi: 10.7717/peerj.6327

Gee, B. M., Bevitt, J. J., and Reisz, R. R. (2020). Computed tomographic analysis of the cranium of the early Permian recumbirostran 'microsaur' Euryodus dalyae reveals new details of the braincase and mandible. *Papers Palaeontol.* doi: 10.1002/spp2.1304

Golubev, V. K. (1998). Revision of the Late Permian Chroniosuchians (Amphibia, Anthracosauromorpha) from Eastern Europe. *Paleontol. J.* 32, 68–77.

Gómez, R. O., Báez, A. M., and Muzzopappa, P. (2011). A new helmeted frog (Anura: Calyptocephalellidae) from an Eocene subtropical lake in northwestern Patagonia, Argentina. *J. Vertebrate Paleontol.* 31, 50–59. doi: 10.1080/02724634.2011.539654

Graur, D., and Martin, W. (2004). Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision. *Trends Genet.* 20, 80–86. doi: 10.1016/j.tig.2003.12.003

Heath, T. A., Huelsenbeck, J. P., and Stadler, T. (2014). The fossilized birth–death process for coherent calibration of divergence-time estimates. *Proc. Natl. Acad. Sci. U.S.A.* 111, E2957–E2966.

Hedges, S. B., Tao, Q., Walker, M., and Kumar, S. (2018). Accurate timetrees require accurate calibrations. *Proc. Natl. Acad. Sci. U.S.A.* 115, E9510–E9511.

Hedman, M. M. (2010). Constraints on clade ages from fossil outgroups. *Paleobiology* 36, 16–31. doi: 10.1666/0094-8373-36.1.16

Hendrikse, J. L., Parsons, T. E., and Hallgrímsson, B. (2007). Evolvability as the proper focus of evolutionary developmental biology. *Evol. Dev.* 9, 393–401. doi: 10.1111/j.1525-142x.2007.00176.x

Hesselbo, S. P., and Trewin, N. H. (1984). Deposition, diagenesis and structures of the cheese bay shrimp bed, lower carboniferous, east lothian. *Scottish J. Geol.* 20, 281–296. doi: 10.1144/sjg20030281

Hime, P. M., Lemmon, A. R., Lemmon, E. C. M., Prendini, E., Brown, J. M., Thomson, R. C., et al. (2020). Phylogenomics reveals ancient gene tree discordance in the amphibian tree of life. *Systematic Biol.* Online ahead of print.

Hook, R. W., and Baird, D. (1986). The diamond coal mine of Linton, Ohio, and its Pennsylvanian-age vertebrates. *J. Vertebrate Paleontol.* 6, 174–190. doi: 10.1080/02724634.1986.10011609

Irisarri, I., Baurain, D., Brinkmann, H., Delsuc, F., Sire, J. Y., Kupfer, A., et al. (2017). Phylotranscriptomic consolidation of the jawed vertebrate timetree. *Nat. Ecol. Evol.* 1, 1370–1378. doi: 10.1038/s41559-017-0240-5

Jenkins, F. A., Walsh, D. M., and Carroll, R. L. (2007). Anatomy of eocaecilia micropodia, a limbed caecilian of the early jurassic. *Bull. Museum Comparat. Zool.* 158, 285–366. doi: 10.3099/0027-4100(2007)158[285:aoemal]2.0.co;2

Kemp, T. S. (2006). The origin and early radiation of the therapsid mammal-like reptiles: a palaeobiological hypothesis. *J. Evol. Biol.* 19, 1231–1247. doi: 10.1111/j.1420-9101.2005.01076.x

Kemp, T. S. (2007). Acoustic transformer function of the postdentary bones and quadrate of a nonmammalian cynodont. *J. Vertebrate Paleontol.* 27, 431–441. doi: 10.1671/0272-4634(2007)27[431:atfotp]2.0.co;2

King, B., and Beck, R. M. (2019). Bayesian tip-dated phylogenetics: topological effects, stratigraphic fit and the early evolution of mammals. *bioRxiv* doi: 10.1101/533885

Klembara, J., Hain, M., Ruta, M., Berman, D. S., Pierce, S. E., and Henrici, A. C. (2020). Inner ear morphology of diadectomorphs and seymouriamorphs (Tetrapoda) uncovered by high-resolution x-ray microcomputed tomography, and the origin of the amniote crown group. *Palaeontology* 63, 131–154. doi: 10.1111/pala.12448

Laurin, M. (2001). L'utilisation de la taxonomie phylogénétique en paléontologie: avantages et inconvénients. *Biosystema* 19, 197–211.

Laurin, M. (2004). The evolution of body size, Cope's rule and the origin of amniotes. *Systematic Biol.* 53, 594–622. doi: 10.1080/10635150490445706

Laurin, M., and de Buffrénil, V. (2016). Microstructural features of the femur in early ophiacodontids: a reappraisal of ancestral habitat use and lifestyle of amniotes. *Comptes Rendus Palevol* 15, 115–127. doi: 10.1016/j.crpv.2015.01.001

Laurin, M., Girondot, M., and de Ricqlès, A. (2000). Early tetrapod evolution. *Trends Ecol. Evol.* 15, 118–123. doi: 10.1016/s0169-5347(99)01780-2

Laurin, M., Lapauze, O., and Marjanović, D. (2019). What do ossification sequences tell us about the origin of extant amphibians? PCI Paleontology. *bioRxiv* doi: 10.1101/352609v4

Laurin, M., and Piñeiro, G. H. (2017). A reassessment of the taxonomic position of mesosaurs, and a surprising phylogeny of early amniotes. *Front. Earth Sci.* 5:88. doi: 10.3389/feart.2017.00088

Laurin, M., and Reisz, R. R. (1995). A reevaluation of early amniote phylogeny. *Zool. J. Linnean Soc.* 113, 165–223. doi: 10.1111/j.1096-3642.1995.tb00932.x

Lebedev, O. A., and Coates, M. I. (1995). The postcranial skeleton of the Devonian tetrapod Tulerpeton curtum Lebedev. *Zool. J. Linnean Soc.* 114, 307–348. doi: 10.1111/j.1096-3642.1995.tb00119.x

Lee, M. S., and Yates, A. M. (2018). Tip-dating and homoplasy: reconciling the shallow molecular divergences of modern gharials with their long fossil record. *Proc. Royal Soc. B: Biol. Sci. U.S.A.* 285:20181071. doi: 10.1098/rspb.2018.1071

Liu, L., and Stockli, D. F. (2020). U-Pb ages of detrital zircons in lower Permian sandstone and siltstone of the Permian Basin, west Texas, USA: evidence of dominant Gondwanan and peri-Gondwanan sediment input to Laurentia. *Bulletin* 132, 245–262. doi: 10.1130/b35119.1

Lombard, R. E., and Bolt, J. R. (1995). A new primitive tetrapod, whatcheeria deltae, from the Lower Carboniferous of Iowa. *Palaeontology* 38, 471–494.

Lombard, R. E., and Bolt, J. R. (1999). A microsaur from the mississippian of illinois and a standard format for morphological characters. *J. Paleontol.* 73, 908–923. doi: 10.1017/s0022336000040749

MacDougall, M. J., Modesto, S. P., Brocklehurst, N., Verrière, A., Reisz, R. R., and Fröbisch, J. (2018). Response: a reassessment of the taxonomic position of mesosaurs, and a surprising phylogeny of early amniotes. *Front. Earth Sci.* 6:99. doi: 10.3389/feart.2018.00099

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 Palaeoclimatol. Palaeoecol.* 475, 1–11. doi: 10.1016/j.palaeo.2017.02.019

Maddin, H. C., and Anderson, J. S. (2012). Evolution of the amphibian ear with implications for lissamphibian phylogeny: insight gained from the caecilian inner ear. *Fieldiana Life Earth Sci.* 2012, 59–76. doi: 10.3158/2158-5520-5.1.59

Maddin, H. C., Jenkins, F. A. Jr., and Anderson, J. S. (2012). The braincase of Eocaecilia micropodia (Lissamphibia. Gymnophiona) and the origin of caecilians. *PLoS One* 7:e50743. doi: 10.1371/journal.pone.0050743

Maddin, H. C., Mann, A., and Hebert, B. (2020). Varanopid from the Carboniferous of Nova Scotia reveals evidence of parental care in amniotes. *Nat. Ecol. Evol.* 4, 50–56. doi: 10.1038/s41559-019-1030-z

Mann, A., Gee, B. M., Pardo, J. D., Marjanović, D., Adams, G. R., Calthorpe, A. S., et al. (2020). Reassessment of historic 'microsaurs' from Joggins, Nova Scotia, reveals hidden diversity in the earliest amniote ecosystem. *Papers Palaeontol.* doi: 10.1002/spp2.1316

Mann, A., McDaniel, E. J., McColville, E. R., and Maddin, H. C. (2019). *Carbonodraco lundi* gen et sp. nov., the oldest parareptile, from Linton, Ohio, and new insights into the early radiation of reptiles. *R. Soc. Open Sci.* 6:191191. doi: 10.1098/rsos.191191

Marjanović, D., and Laurin, M. (2007). Fossils, molecules, divergence times, and the origin of lissamphibians. *Systematic Biol.* 56, 369–388. doi: 10.1080/10635150701397635

Marjanović, D., and Laurin, M. (2013). The origin (s) of extant amphibians: a review with emphasis on the "lepospondyl hypothesis". *Geodiversitas* 35, 207–273. doi: 10.5252/g2013n1a8

Marjanović, D., and Laurin, M. (2019). Phylogeny of Paleozoic limbed vertebrates reassessed through revision and expansion of the largest published relevant data matrix. *PeerJ* 6:e5565. doi: 10.7717/peerj.5565

Matzke, N. J., and Irmis, R. B. (2018). Including autapomorphies is important for paleontological tip-dating with clocklike data, but not with non-clock data. *PeerJ* 6:e4553. doi: 10.7717/peerj.4553

Milner, A. R. (1980). The temnospondyl amphibian dendrerpeton from the upper carboniferous of Ireland. *Palaeontology* 23, 125–141.

Milner, A. R. (1982). Small temnospondyl amphibians from the Middle Pennsylvanian of Illinois. *Palaeontology* 25, 635–664.

Milner, A. R. (1993). "Biogeography of Palaeozoic Tetrapods" in *Palaeozoic Vertebrate Biostratigraphy and Biogeography* J.A. Long (ed.). (London: Belhaven Press), 324–353

Milner, A. R., and Panchen, A. L. (1973). "Geographical variation in the tetrapod faunas of the Upper Carboniferous and Lower Permian," in *Implications of Continental Drift to the Earth Sciences*, Vol. 1, eds D. H. Tarling and S. K. Runcorn (London: Academic Press), 353–368.

Milner, A. R., and Sequeira, S. E. K. (1993). The temnospondyl amphibians from the viséan of east kirkton, west lothian, Scotland. *Earth Environ. Sci. Trans. Royal Soc. Edinburgh* 84, 331–361. doi: 10.1017/s0263593300006155

Modesto, S. P., Scott, D. M., MacDougall, M. J., Sues, H. D., Evans, D. C., and Reisz, R. R. (2015). The oldest parareptile and the early diversification of reptiles. *Proc. Royal Soc. B: Biol. Sci.* 282:20141912. doi: 10.1098/rspb.2014.1912

Morris, J. L., Puttick, M. N., Clark, J. W., Edwards, D., Kenrick, P., Pressel, S., et al. (2018). Reply to Hedges et al.: accurate timetrees do indeed require accurate calibrations. *Proc. Natl. Acad. Sci.* 115, E9512–E9513.

Müller, J., and Reisz, R. R. (2005). Four well-constrained calibration points from the vertebrate fossil record for molecular clock estimates. *BioEssays* 27, 1069–1075. doi: 10.1002/bies.20286

Müller, J., and Reisz, R. R. (2006). The phylogeny of early eureptiles: comparing parsimony and bayesian approaches in the investigation of a basal fossil clade. *Syst. Biol.* 55, 503–511. doi: 10.1080/10635150600755396

Nasbitt, S. J. (2011). The early evolution of archosaurs: relationships and the origin of major clades. *Bull. Am. Museum Nat. History* 352, 1–292. doi: 10.1206/352.1

Near, T. J., and Sanderson, M. J. (2004). Assessing the quality of molecular divergence time estimates by fossil calibrations and fossil–based model selection. *Phil. Trans. Royal Soc. London. Series B: Biol. Sci.* 359, 1477–1483. doi: 10.1098/rstb.2004.1523

Nelson, W. J., Hook, R. W., and Chaney, D. S. (2013a). Lithostratigraphy of the lower Permian (Leonardian) clear fork formation of north-central Texas. *New Mexico Museum Natural History Sci. Bull.* 60, 286–311.

Nelson, W. J., Hook, R. W., and Elrick, S. (2013b). Subsurface nomenclature in the Permian basin (Texas-New Mexico): lithostratigrapic chaos or fixable problem? *New Mexico Museum Natural History Sci. Bull.* 60, 312–313.

Olive, S., Ahlberg, P. E., Pernègre, V. N., Poty, É, Steurbaut, É, and Clément, G. (2016). New discoveries of tetrapods (ichthyostegid-like and whatcheeriid-like) in the Famennian (Late Devonian) localities of Strud and Becco (Belgium). *Palaeontology* 59, 827–840. doi: 10.1111/pala.12261

Olori, J. C. (2013). Ontogenetic sequence reconstruction and sequence polymorphism in extinct taxa: an example using early tetrapods (Tetrapoda: Lepospondyli). *Paleobiology* 39, 400–428. doi: 10.1666/12031

Olroyd, S. L., and Sidor, C. A. (2017). A review of the Guadalupian (middle Permian) global tetrapod fossil record. *Earth-Sci. Rev.* 171, 583–597. doi: 10.1016/j.earscirev.2017.07.001

Pardo, J. D., and Anderson, J. S. (2016). Cranial morphology of the Carboniferous-Permian tetrapod *Brachydectes newberryi* (Lepospondyli, Lysorophia): new data from μCT. *PLoS One* 11:e0161823. doi: 10.1371/journal.pone.0161823

Pardo, J. D., Holmes, R., and Anderson, J. S. (2018). An enigmatic braincase from Five Points, Ohio (Westphalian D) further supports a stem tetrapod position for aïstopods. *Earth Environ. Sci. Trans. Royal Soc. Edinburgh* 109, 255–264. doi: 10.1017/s1755691018000567

Pardo, J. D., and Mann, A. (2018). A basal aïstopod from the earliest pennsylvanian of Canada, and the antiquity of the first limbless tetrapod lineage. *Royal Soc. Open Sci.* 5:181056. doi: 10.1098/rsos.181056

Pardo, J. D., Small, B. J., and Huttenlocker, A. K. (2017a). Stem caecilian from the Triassic of Colorado sheds light on the origins of Lissamphibia. *Proc. Natl. Acad. Sci. U.S.A.* 114, E5389–E5395.

Pardo, J. D., Szostakiwskyj, M., Ahlberg, P. E., and Anderson, J. S. (2017b). Hidden morphological diversity among early tetrapods. *Nature* 546, 642–645. doi: 10.1038/nature22966

Pardo, J. D., Small, B. J., Milner, A. R., and Huttenlocker, A. K. (2019). Carboniferous–Permian climate change constrained early land vertebrate radiations. *Nat. Ecol. Evol.* 3, 200–206. doi: 10.1038/s41559-018-0776-z

Pardo, J. D., Szostakiwskyj, M., and Anderson, J. S. (2015). Cranial morphology of the brachystelechid 'microsaur' Quasicaecilia texana Carroll provides new insights into the diversity and evolution of braincase morphology in recumbirostran 'microsaurs'. *PLoS One* 10:e0130359. doi: 10.1371/journal.pone.0130359

Parham, J. F., Donoghue, P. C. J., Bell, C. J., Calway, T. D., Head, J. J., Holroyd, P. A., et al. (2012). Best practices for justifying fossil calibrations. *Syst. Biol.* 61, 346–359. doi: 10.1093/sysbio/syr107

Parham, J. F., and Irmis, R. B. (2008). Caveats on the use of fossil calibrations for molecular dating: a comment on Near et al. *Am. Nat.* 171, 132–136. doi: 10.1086/524198

Patterson, C. (1981). Significance of fossils in determining evolutionary relationships. *Ann. Rev. Ecol. Systematics* 12, 195–223. doi: 10.1146/annurev.es.12.110181.001211

Pérez-Ben, C. M., Schoch, R. R., and Báez, A. M. (2018). Miniaturization and morphological evolution in Paleozoic relatives of living amphibians: a quantitative approach. *Paleobiology* 44, 58–75. doi: 10.1017/pab.2017.22

Pyron, R. A. (2014). Biogeographic analysis reveals ancient continental vicariance and recent oceanic dispersal in amphibians. *Systematic Biol.* 63, 779–797. doi: 10.1093/sysbio/syu042

Rage, J. C., and Roček, Z. (1989). Redescription of Triadobatrachus massinoti (Piveteau, 1936) an anuran amphibian from the early Triassic. *Palaeontographica A* 206, 1–16.

Raup, D. M., and Sepkoski, J. J. (1982). Mass extinctions in the marine fossil record. *Science* 215, 1501–1503. doi: 10.1126/science.215.4539.1501

Reisz, R. (1972). Pelycosaurian reptiles from the middle pennsylvanian of North America. *Bull. Museum Comparat. Zool. Harvard College* 144, 27–61.

Reisz, R. R. (1997). The origin and early evolutionary history of amniotes. *Trends Ecol. Evol.* 12, 218–222. doi: 10.1016/s0169-5347(97)01060-4

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. *Geol. Soc. Am. Bull.* 113, 1229–1233. doi: 10.1130/0016-7606(2001)113<1229:trmfve>2.0.co;2

Reisz, R. R., Modesto, S. P., and Scott, D. (2011). A new Early Permian reptile and its significance in early diapsid evolution. *Proc. R. Soc. B* 278, 3731–3737. doi: 10.1098/rspb.2011.0439

Rubidge, B. S., Erwin, D. H., Ramezani, J., Bowring, S. A., and de Klerk, W. J. (2013). High-precision temporal calibration of Late Permian vertebrate biostratigraphy: U-Pb zircon constraints from the Karoo Supergroup, South Africa. *Geology* 41, 363–366. doi: 10.1130/g33622.1

Ruta, M., and Clack, J. A. (2006). A review of silvanerpeton miripedes, a stem amniote from the lower carboniferous of east kirkton, West Lothian, Scotland. *Trans. Royal Soc. Edinburgh: Earth Sci.* 97, 31–63. doi: 10.1017/s0263593306000034

Ruta, M., and Coates, M. I. (2007). Dates, nodes and character conflict: addressing the lissamphibian origin problem. *J. Systematic Palaeontol.* 5, 69–122. doi: 10.1017/s1477201906002008

Ruta, M., Jeffery, J., and Coates, M. (2003). A supertree of early tetrapods. *Proc. Biol. Sci.* 270, 2507–2516. doi: 10.1098/rspb.2003.2524

Sallan, L. C., and Coates, M. I. (2010). End-Devonian extinction and a bottleneck in the early evolution of modern jawed vertebrates. *Proc. Natl. Acad. Sci. U.S.A.* 107, 10131–10135. doi: 10.1073/pnas.0914000107

San Mauro, D. (2010). A multilocus timescale for the origin of extant amphibians. *Mol. Phylogenet. Evol.* 56, 554–561. doi: 10.1016/j.ympev.2010.04.019

San Mauro, D., Vences, M., Alcobendas, M., Zardoya, R., and Meyer, A. (2005). Initial diversification of living amphibians predated the breakup of Pangaea. *Am. Nat.* 165, 590–599. doi: 10.1086/429523

Schluter, D. (2000). *The Ecology of Adaptive Radiation*. Oxford: Oxford University Press, 296.

Schoch, R. R., and Sues, H. D. (2015). A middle triassic stem-turtle and the evolution of the turtle body plan. *Nature* 523, 584. doi: 10.1038/nature14472

Schoch, R. R., Werneburg, R., and Voigt, S. (2020). A triassic stem-salamander from kyrgyzstan and the origin of salamanders. *Proc. Natl. Acad. Sci. U.S.A.* 117, 11584–11588. doi: 10.1073/pnas.2001424117

Sidor, C. A., and Hopson, J. A. (1998). Ghost lineages and "mammalness": assessing the temporal pattern of character acquisition in the Synapsida. *Paleobiology* 24, 254–273.

Sigurdsen, T., and Bolt, J. R. (2009). The lissamphibian humerus and elbow joint, and the origins of modern amphibians. *J. Morphol.* 270, 1443–1453. doi: 10.1002/jmor.10769

Sigurdsen, T., and Green, D. M. (2011). The origin of modern amphibians: a re-evaluation. *Zool. J. Linnean Soc.* 162, 457–469. doi: 10.1111/j.1096-3642.2010.00683.x

Simões, T. R., Caldwell, M. W., Tałanda, M., Bernardi, M., Palci, A., Vernygora, O., et al. (2018). The origin of squamates revealed by a Middle Triassic lizard from the Italian Alps. *Nature* 557, 706–709. doi: 10.1038/s41586-018-0093-3

Smithson, T. R. (1985). The morphology and relationships of the Carboniferous amphibian Eoherpeton watsoni Panchen. *Zool. J. Linnean Soc.* 85, 317–410. doi: 10.1111/j.1096-3642.1985.tb01517.x

Smithson, T. R., Carroll, R. L., Panchen, A. L., and Andrews, S. M. (1993). Westlothiana lizziae from the Viséan of East Kirkton, West Lothian, Scotland, and the amniote stem. *Earth Environ. Sci. Trans. Royal Soc. Edinburgh* 84, 383–412. doi: 10.1017/s0263593300006192

Smithson, T. R., Wood, S. P., Marshall, J. E., and Clack, J. A. (2012). Earliest Carboniferous tetrapod and arthropod faunas from Scotland populate Romer's Gap. *Proc. Natl. Acad. Sci. U.S.A.* 109, 4532–4537. doi: 10.1073/pnas.1117332109

Stadler, T., and Yang, Z. (2013). Dating phylogenies with sequentially sampled tips. *Systematic Biol.* 62, 674–688. doi: 10.1093/sysbio/syt030

Stampfli, G. M., Hochard, C., Vérard, C., and Wilhem, C. (2013). The formation of Pangea. *Tectonophysics* 593, 1–19. doi: 10.1016/j.tecto.2013.02.037

Sues, H. D. (2019). Authorship and date of publication of the name Tetrapoda. *J. Vertebrate Paleontol.* 39:e1564758. doi: 10.1080/02724634.2019.1564758

Szostakiwskyj, M., Pardo, J. D., and Anderson, J. S. (2015). Micro-CT study of Rhynchonkos stovalli (Lepospondyli, Recumbirostra), with description of two new genera. *PLoS One* 10:e0127307. doi: 10.1371/journal.pone.0127307

Takezaki, N., and Nishihara, H. (2017). Support for lungfish as the closest relative of tetrapods by using slowly evolving ray-finned fish as the outgroup. *Genome Biol. Evol.* 9, 93–101.

Trojan, A., Bojanowski, M. J., Gola, M., Grafka, O., Marynowski, L., and Clarkson, E. N. (2015). Organic geochemical characteristics of the Mississippian black shales from Wardie, Scotland. *Earth Environ. Sci. Trans. Royal Soc. Edinburgh* 106, 55–65. doi: 10.1017/s1755691015000225

Vallin, G., and Laurin, M. (2004). Cranial morphology and affinities of Microbrachis, and a reappraisal of the phylogeny and lifestyle of the first amphibians. *J. Vertebrate Paleontol.* 24, 56–72. doi: 10.1671/5.1

Ward, P., Labandeira, C., Laurin, M., and Berner, R. A. (2006). Confirmation of Romer's Gap as a low oxygen interval constraining the timing of initial arthropod and vertebrate terrestrialization. *Proc. Natl. Acad. Sci. U.S.A.* 103, 16818–16822. doi: 10.1073/pnas.0607824103

Warren, A. (2007). New data on Ossinodus pueri, a stem tetrapod from the Early Carboniferous of Australia. *J. Vertebrate Paleontol.* 27, 850–862. doi: 10.1671/0272-4634(2007)27[850:ndoopa]2.0.co;2

Witzmann, F., Schoch, R. R., and Maisch, M. (2008). A relict basal tetrapod from Germany: first evidence of a Triassic chroniosuchian outside Russia. *Sci. Nat.* 95, 67–72. doi: 10.1007/s00114-007-0291-6

Witzmann, F., and Werneburg, I. (2017). The palatal interpterygoid vacuities of temnospondyls and the implications for the associated eye and jaw musculature. *Anat. Rec.* 300, 1240–1269. doi: 10.1002/ar.23582

Zhu, M., and Yu, X. (2002). A primitive fish close to the common ancestor of tetrapods and lungfish. *Nature* 418, 767–770. doi: 10.1038/nature00871

Ziegler, A. M., Scotese, C. R., McKerrow, W. S., Johnson, M. E., and Bambach, R. K. (1979). Paleozoic paleogeography. *Ann. Rev. Earth Planetary Sci.* 7, 473–502.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Pardo, Lennie and Anderson. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genetics

HYPOTHESIS AND THEORY
published: 12 May 2021
doi: 10.3389/fgene.2021.521693

The Making of Calibration Sausage Exemplified by Recalibrating the Transcriptomic Timetree of Jawed Vertebrates

*David Marjanović**

Department of Evolutionary Morphology, Science Programme "Evolution and Geoprocesses", Museum für Naturkunde – Leibniz Institute for Evolutionary and Biodiversity Research, Berlin, Germany

Molecular divergence dating has the potential to overcome the incompleteness of the fossil record in inferring when cladogenetic events (splits, divergences) happened, but needs to be calibrated by the fossil record. Ideally but unrealistically, this would require practitioners to be specialists in molecular evolution, in the phylogeny and the fossil record of all sampled taxa, and in the chronostratigraphy of the sites the fossils were found in. Paleontologists have therefore tried to help by publishing compendia of recommended calibrations, and molecular biologists unfamiliar with the fossil record have made heavy use of such works (in addition to using scattered primary sources and copying from each other). Using a recent example of a large node-dated timetree inferred from molecular data, I reevaluate all 30 calibrations in detail, present the current state of knowledge on them with its various uncertainties, rerun the dating analysis, and conclude that calibration dates cannot be taken from published compendia or other secondary or tertiary sources without risking strong distortions to the results, because all such sources become outdated faster than they are published: 50 of the (primary) sources I cite to constrain calibrations were published in 2019, half of the total of 280 after mid-2016, and 90% after mid-2005. It follows that the present work cannot serve as such a compendium either; in the slightly longer term, it can only highlight known and overlooked problems. Future authors will need to solve each of these problems anew through a thorough search of the primary paleobiological and chronostratigraphic literature on each calibration date every time they infer a new timetree, and that literature is not optimized for that task, but largely has other objectives.

Keywords: timetree, calibration, fossil record, gnathostomata, vertebrata, stemward slippage, divergence date

OPEN ACCESS

Edited by:
Denis Baurain,
University of Liège, Belgium

Reviewed by:
Matt Phillips,
Queensland University of Technology, Australia
Per Erik Ahlberg,
Uppsala University, Sweden

***Correspondence:**
David Marjanović
david.marjanovic@gmx.at
orcid.org/0000-0001-9720-7726

Specialty section:
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Genetics

Received: *16 January 2020*
Accepted: *22 March 2021*
Published: *12 May 2021*

Citation:
Marjanović D (2021) The Making of Calibration Sausage Exemplified by Recalibrating the Transcriptomic Timetree of Jawed Vertebrates. Front. Genet. 12:521693. doi: 10.3389/fgene.2021.521693

INTRODUCTION

This work is not intended as a review of the theory or practice of node (or tip) dating with calibration dates (or tip dates) inferred from the fossil record; as the most recent reviews of methods and sources of error, I recommend those by Barido-Sottani et al. (2019), Barido-Sottani et al. (2020), Marshall (2019), Matschiner (2019), Guindon (2020), Pardo et al. (2020), Powell et al. (2020), and, with caveats of which I will address two (see section "Materials and Methods": Calibrations:

Node 152 – Placentalia), Springer et al. (2019). Neither is it intended as a review of the history of the dates assigned to certain calibrations; as an example of a recent detailed review of three commonly used calibrations, I recommend Pardo et al. (2020). Although I discuss wider implications, the scope of this work is narrow: to evaluate each of the 30 calibrations used in the largest vertebrate timetree yet published, that by Irisarri et al. (2017), and the total impact of the errors therein on the results (using the same node-dating method they used, which I do not evaluate beyond mentioning potential general points of criticism).

Irisarri et al. (2017) inferred a set of timetrees from the transcriptomes of 100 species of gnathostomes (jawed vertebrates) and combinations of up to 30 calibrations from the fossil record. On the unnumbered ninth page of their supplementary information, they described their calibration dates as "five well-accepted fossil calibrations plus a prior on the root" and "24 additional well-established calibration points with solid paleontological evidence." For many of the calibrations, these optimistic assessments are not tenable. I have tried to present, and use, the current state of knowledge on each of these calibrations.

In doing so, the present work naturally resembles the compendia of suggested calibrations that paleontologists have occasionally compiled with the intent to provide a handy reference for molecular biologists who wish to date divergences [e.g., Müller and Reisz, 2007; Benton et al., 2015, and six other articles in *Palaeontologia Electronica* 18(1); Wolfe et al., 2016; Morris et al., 2018]; Irisarri et al. (2017) took 7 of their 30 calibrations from the compendium in Benton and Donoghue (2007: table 1) alone—without citing the enlarged update by Benton et al. (2015)—compared to six taken from the primary literature. However, I will show that all such compendia are doomed to be (partially) outdated almost as fast as they are published in the best case, and faster than they are published in the average case. Soon, therefore, the present work will no longer be reliable as such a compendium either; rather, it is intended to show readers where the known uncertainties and disagreements lie, and thus what anybody who wants to use a particular calibration should probably search the most recent literature for. This is why I do not generally begin my discussion of a calibration by presenting my conclusions on what the best, or least bad, minimum and maximum ages of the calibration may be (They are, however, presented without further ornament in **Table 1**.) Instead, I walk the reader through a sometimes meandering discovery process, demonstrating how this knowledge was arrived at and how it may soon change—how the sausage was made and how it may spoil.

Some works used as compendia in this sense are not even compiled by paleontologists: molecular biologists often copy from each other. Irisarri et al. (2017) took four of their calibrations from table 1 of Noonan and Chippindale (2006), a work that contains a phylogenetic and divergence-date analysis of molecular data and cites severely outdated paleontological primary and secondary literature (from 1981 to 2003) as its sources.

A continually updated online compendium could largely avoid the problem that knowledge has a half-life. There has been one attempt to create one, the Fossil Calibration Database (Ksepka et al., 2015[1]; not counting separately its predecessor, called Date a Clade, which is no longer online and apparently merely presented Table 1 of Benton and Donoghue, 2007). It appears to have run out of funding long ago and has not been updated since February 2, 2018, the day on which three of the numerous calibrations proposed in Wolfe et al. (2016) were added; other calibrations from the same source were added on January 30 and 31, 2018 (one each) and December 22, 2017 (three), and no other updates were made on those days. I cannot resist pointing out that this is one of many cases where funding menial labor in the sciences—reading and interpreting papers, evaluating the contradictions between them, and entering the interpretations in a database, a task that cannot be automated—would go a long way toward improving the quality of a large number of publications, but is unlikely to be granted because it is not likely to result in a single flashy publication or in an immediately marketable application directly, even though precise and accurate timetrees are an essential component of our understanding of the model organisms used in biomedical research.

A continually updated online database aiming to represent the entire fossil record exists and is currently being funded: the Paleobiology Database, accessible through two different interfaces at http://www.pbdb.org and https://paleobiodb.org. Among many other things, it aims to contain the oldest currently known record of every taxon and would thus be useful as a source for calibrations. However, the warnings by Parham et al. (2011) still apply: the quality of the Paleobiology Database is quite heterogeneous. While some entries are written by the current top experts in the respective fields, others copy decades-old primary descriptive literature uncritically, often leading to severely outdated taxonomic, let alone phylogenetic placements (in all but the most recent literature that is not the same), not to mention misunderstandings based on the convoluted history of taxonomic nomenclature. It is not uncommon for two entries to contradict each other. Finally, despite the hundreds of contributors, our current knowledge of the fossil record is so vast that the database remains incomplete (again, of course, differently so for different taxa). Like Irisarri et al. (2017), I have not used the Paleobiology Database or the Fossil Calibration Database; I have relied on the primary literature.

Nomenclature

After the publication of the *International Code of Phylogenetic Nomenclature (PhyloCode)* (Cantino and de Queiroz, 2020) and its companion volume *Phylonyms* (de Queiroz et al., 2020), the registration database for phylogenetic nomenclature—*RegNum* (Cellinese and Dell, 2020)—went online on June 8, 2020; regulated phylogenetic nomenclature is therefore operational. In an effort to promote uniformity and stability in nomenclature, I have used the names and definitions from *Phylonyms*, Ezcurra et al. (2020: online methods) and Joyce et al. (2021) here; wherever applicable, all of them are followed by "[PN]" at least at the first mention (this includes vernacularized forms like "gnathostome") to avoid confusion with earlier uses of

[1]https://fossilcalibrations.org

TABLE 1 | The first four columns of Irisarri et al. (2017: supplementary table 8), here expanded to five, followed by the ages used here for the same calibrations and the differences (Δ).

Node number in Irisarri et al. (2017: supplementary table 8 and supplementary figure 19)	Description of cladogenesis	The sampled terminal taxa that stem from this node are:	Minimum age in Irisarri et al. (2017)	Maximum age in Irisarri et al. (2017)	Minimum age used here	Maximum age used here	Δ minimum ages	Δ maximum ages
100	Root node = Gnathostomata: total group including Chondrichthyes – Pan-Osteichthyes	Entire sample	421.75	462.5	**465***	**475**	+43.25	+12.5
102	Osteichthyes: Pan-Actinopterygii – Sarcopterygii	Entire sample except Chondrichthyes	416	439	(420*)	(475)	+4	+36
104	Dipnomorpha – Tetrapodomorpha	Dipnoi – Tetrapoda	408	419	**420***	(475)	+12	+56
105	Tetrapoda: Amphibia – Pan-Amniota	Lissamphibia – Amniota	330.4	350.1	**335*** (or **350***)	365	+4.6 (or +19.6)	+14.9
106	Amniota: Pan-Mammalia – Sauropsida	Mammalia – Reptilia	288	338	**318***	(365)	+30	+27
107	Reptilia: Pan-Lepidosauria – total group of Archelosauria	Lepidosauria – Testudines, Crocodylia, Aves	252	257	(263*)	(365)	+11	+108
108	Archelosauria: Pan-Testudines – Pan-Archosauria	Testudines – Crocodylia, Aves	(243)	(257)	**263***	(365)	+20	+108
109	Archosauria: Crocodylotarsi – Pan-Aves	Crocodylia – Aves	243	251	**248***	252	+5	+1
111	Alligatoridae: Alligatorinae – Caimaninae	*Alligator – Caiman*	66	75	65*	200*	−1	+125
113	Neognathae: Galloanserae – Neoaves	*Anas*, *Gallus*, *Meleagris – Taeniopygia*	66	86.5	71	115	+5	+28.5
117	Testudines: Pan-Pleurodira – Pan-Cryptodira	*Phrynos*, *Pelusios* – all other sampled turtles	210	(257)	**158***	**185**	−52	−72
124	Pleurodira: Pan-Chelidae – Pan-Pelomedusoides	*Phrynops – Pelusios*	25	(257)	**125***	(185)	+100	−122
125	Lepidosauria: Rhynchocephalia – Pan-Squamata	*Sphenodon* – Squamata	238	(257)	**244***	**290**	+6	+33
129	Toxicofera: Pan-Serpentes – Anguimorpha + Pan-Iguania	Snakes – their sister-group	"148" (165)	(257)	130*	(290)	"−18" (−35)	+33
131	Iguania: Pan-Acrodonta – Pan-Iguanidae	*Pogona*, *Chamaeleo* – *Iguana*, *Basiliscus*, *Sceloporus*, *Anolis*	165	230	**72***	(290)	−93	+60

(Continued)

TABLE 1 | Continued

Node number in Irisarri et al. (2017: supplementary table 8 and supplementary figure 19)	Description of cladogenesis	The sampled terminal taxa that stem from this node are:	Minimum age in Irisarri et al. (2017)	Maximum age in Irisarri et al. (2017)	Minimum age used here	Maximum age used here	Δ minimum ages	Δ maximum ages
132	Iguanidae: Iguaninae + Corytophanidae – Phrynosomatidae + Dactyloidae	*Iguana*, *Basiliscus* – *Sceloporus*, *Anolis*	125	180	**53***	(290)	−72	+110
150	Mammalia (Pan-Monotremata – Theriimorpha)	*Ornithorhynchus* – Theria	162.5	191.4	**179***	233*	+16.5	+41.6
151	Theria: Metatheria – Eutheria	Marsupialia – Placentalia	124.6	138.4	**126***	160	+1.4	+21.6
152	Placentalia: Atlantogenata – Boreo(eu)theria	*Loxodonta*, *Dasypus* – *Felis*, *Canis*, *Homo*, *Mus*	95.3	113	(66*)	**72***	−29.3	−41
153	Boreo(eu)theria: Laurasiatheria – Euarchontoglires/Supraprimates	*Felis*, *Canis* – *Homo*, *Mus*	(61.5)	(113)	**66***	(72*)	+4.5	−41
154	Carnivora: Pan-Feliformia – Pan-Caniformia	*Felis* – *Canis*	42.8	63.8	**38***	**56***	−4.8	−7.8
155	Euarchontoglires/Supraprimates: Gliriformes – Primatomorpha	*Mus* – *Homo*	61.5	100.5	**65***	(72*)	+3.5	−28.5
157	Marsupialia: Didelphimorphia – Paucituberculata + Australidelphia	*Monodelphis* – *Macropus*, *Sarcophilus*	61.5	71.2	**55***	**68***	−6.5	−3.2
160	Batrachia: Urodela – Salientia	Caudata – Anura	249	(350.1)	**249***	**290**	0	−60.1
169	Crown group of Cryptobranchoidea: Hynobiidae – Pancryptobrancha	*Hynobius* – *Andrias*	145.5	(350.1)	101*	(290)	−44.5	−60.1
170	Lalagobatrachia/Bombinanura: total group of Bombinatoroidea/Costata – total group of Pipanura	*Bombina*, *Discoglossus* – all other sampled frogs	161.2	(350.1)	(153*)	(290)	−8.2	−60.1
171	Pipanura: total group of Pipoidea/Xenoanura – total group of Acosmanura	*Pipa*, *Hymenochirus*, *Silurana* – their sister-group	145.5	(350.1)	**153***	(290)	+7.5	−60.1
178	Pipidae: Pipinomorpha – Xenopodinomorpha	*Pipa* – *Silurana*, *Hymenochirus*	86	(350.1)	**84***	199*	−2	−151.1
187	Crown group of Chondrichthyes (Holocephali – Elasmobranchii)	*Callorhinchus* – Elasmobranchii	410	"495" (462.5)	**385***	(475)	−25	"-20" (+12.5)

(Continued)

TABLE 1 | Continued

Node number in Irisarri et al. (2017: supplementary table 8 and supplementary figure 19)	Description of cladogenesis	The sampled terminal taxa that stem from this node are:	Minimum age in Irisarri et al. (2017)	Maximum age in Irisarri et al. (2017)	Minimum age used here	Maximum age used here	Δ minimum ages	Δ maximum ages
188	Crown group of Elasmobranchii (Selachimorpha – Batomorpha)	Sharks – rays	190	(462.5)	201*	395*	+11	−67.5
192	Batoidea (Rajiformes – rays)	*Neotrygon* – *Raja*, *Leucoraja*	176	(462.5)	**184***	**201***	+8	−261.5
195	Neopterygii (Holosteomorpha – Pan-Teleostei)	*Lepisosteus*, *Amia* – *Takifugu*, *Danio*	345	392	**249***	299	−96	−93

Boldface is a rough indicator of my confidence. Hard bounds are marked with an asterisk (though all nodes had to be treated as hard in my analysis, see text). Dates in parentheses were not specified in the analysis; the node was constrained in practice by the given constraint on a preceding (for maximum ages) or following node (for minimum ages) elsewhere in this table—see ***Figure 1*** *for which nodes precede each other. The two dates in quotation marks were specified by* Irisarri et al. (2017)*, but had no effect because they were in practice constrained by the dates specified for other nodes. Dashes in the second and third column separate the two branches stemming from the node in question. Depending on the node, see the text or the* ***Supplementary Material*** *for discussion and references.*

the same names for different clades. I have not, however, followed the *ICPN*'s Recommendation 6.1A to set all taxonomic names in italics.

The definitions of these names, their registration numbers (which establish priority among the combinations of name and definition), and the exact chapter citations can be found in *RegNum*, which is freely accessible[2].

ICPN-regulated names have not been created or converted according to a single overarching scheme. As a result, for example, the name Osteichthyes has been defined as applying to a crown group, and the corresponding total group has been named Pan-Osteichthyes, but the name Chondrichthyes has not been defined and could end up as the name for a crown group, a total group, or neither (indeed, current common usage by paleontologists is neither). This has required some awkward circumlocutions. Following Recommendation 9B of the *ICPN*, I have not coined any new names or definitions in the present work.

The shapes and definitions of most other taxonomic names used here do not currently compete for homonymy or synonymy under any code of nomenclature. (The *ICPN* is not retroactive, and the rank-based *International Code of Zoological Nomenclature* [International Commission on Zoological Nomenclature, 1999] does not regulate the priority of names at ranks above the family group.) In such cases, I have followed current usage where that is trivial; I occasionally mention synonyms where that seems necessary.

The usage of "stem" and "crown" requires a comment. The crown group of a clade consists of the last common ancestor of all extant members of that clade, plus all its descendants. The rest of the clade in question is its stem group. For example, *Gallus* is a crown-group dinosaur, and *Triceratops* is a stem-group dinosaur. In a development that seems not to have been foreseen by the first two or so generations of phylogeneticists that established the terminology—for example, the zoology textbook by Ax (1987) exclusively named total groups, i.e., halves of crown groups!—many clades with defined names are now identical to their crown groups (in other words, they are crown clades); they do not contain any part of their stem. Aves [PN] is an example; although *Triceratops* is a stem-dinosaur [PN], a stem-dinosauromorph [PN], and a stem-ornithodiran [PN] among other things, it is not a stem-bird or stem-avian because by definition there is no such thing. It is instead a stem-pan-avian [PN], i.e., a stem-group member of Pan-Aves [PN] (Ezcurra et al., 2020: online methods). If no name is available for a suitable larger group, I have resorted to the circumlocution that *Triceratops*, for instance, is "on the bird stem" or "in the avian total group" (expressing that it is closer to Aves than to any mutually exclusive crown group).

MATERIALS AND METHODS

Although I have followed the spirit of the guidelines developed by Parham et al. (2011) for how best to justify or evaluate a proposed calibration, I have not consistently followed their letter. Most

[2]https://www.phyloregnum.org/

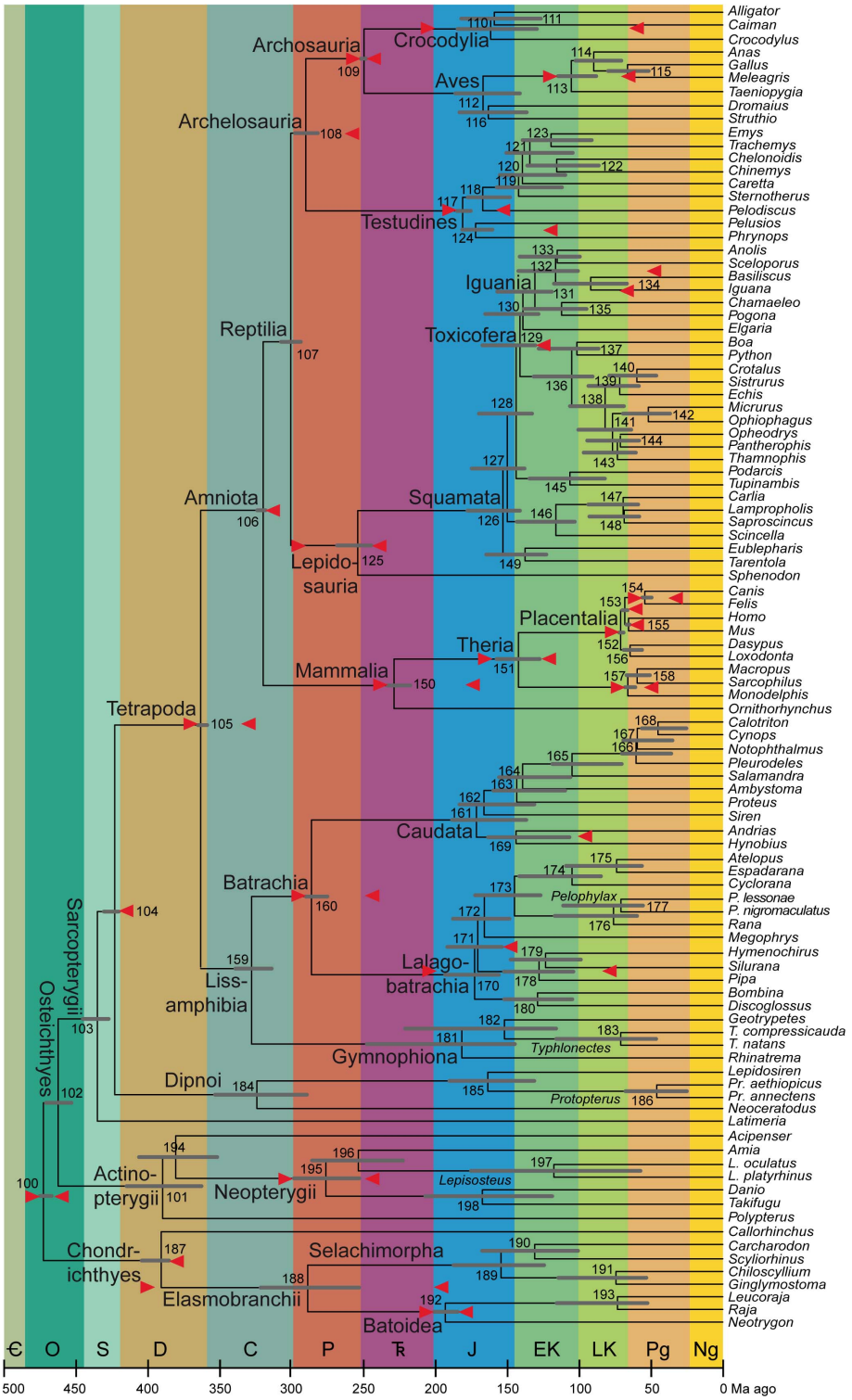


FIGURE 1 | Average timetree resulting from application of the calibrations described here (mostly in the **Supplementary Material**). As in **Table 2** and in Irisarri et al. (2017: figure 3), the bars on the nodes are the superimposed 95% credibility intervals from the two runs in PhyloBayes. The calibrations are shown as red arrows horizontally in line with the nodes they apply to; note that the arrow that is almost aligned with the branch of Lalagobatrachia and the one that is almost aligned with the terminal branch for *Silurana* are the maximum and minimum ages of Node 178 (Pipidae), the one on *Iguana* to Node 131 (Iguania), and the one on *Pelodiscus* to Node 117 (Testudines). The abbreviated genus names are spelled out as clade names on their common branches; where only one species per genus is sampled,

(Continued)

FIGURE 1 | Continued
see Irisarri et al. (2017) for full species names. To the extent possible, clade names with minimum-clade (node-based) definitions are placed close to those nodes, while names with maximum-clade (branch-based) definitions are shown close to the origin of that branch (i.e., the preceding node if sampled) and undefined names stay in the middle. Period/epoch symbols from oldest to youngest: Cambrian (cut off at 500 Ma), Ordovician, Silurian, Devonian, Carboniferous, Permian, Triassic, Jurassic, Early Cretaceous, Late Cretaceous, Paleogene, and Neogene including Quaternary (which comprises the last 2.58 Ma and is not shown separately). Timescale (including colors) from the International Chronostratigraphic Chart, version 2020/03 (Cohen K. M. et al., 2020). Node numbers, also used in the text and the tables, are from Irisarri et al. (2017).

notably, the specimen numbers of the fossils that I largely refer to by genus names can all be found in the directly cited primary literature, so they are not repeated here.

Hard and Soft Minima and Maxima

Without discussing the matter, Irisarri et al. (2017) stated that they had treated all calibration ages as soft bounds, which, in the software they used, means that "a proportion of 0.05 of the total probability mass is allocated outside the specified bound(s) (which means, 5% on one side, in the cases of the pure lower and pure upper bounds, and 2.5% on each side in the case of a combination of lower and upper bound)" (Lartillot, 2015: manual). This is particularly odd for minimum ages; after all, the probability that a clade is younger than its oldest fossil is not 5% or 2.5%, it is 0%. A few other works have used soft minima as an attempt to account for phylogenetic or chronostratigraphic uncertainty of the specimens chosen as calibrations. I have not used the former approach here (despite two clumsy attempts in the first preprint of this paper—Marjanović, 2019—that were rightly pointed out as incoherent by a reviewer): in the cases of phylogenetic uncertainty discussed below, different fossils that could calibrate the age of a cladogenetic event are commonly tens of millions of years apart, a situation that cannot be smoothed over by using the oldest one as a soft minimum. Soft minima that can be justified by uncertainty over the exact age of a calibrating fossil are very rare nowadays (as already pointed out by Parham et al., 2011); within the scope of this paper, there is only one such case, the minimum age of Neognathae (node 113), which is determined by a specimen that is roughly 70 ± 1 Ma old according to a fairly long chain of inference. I have treated all other minima as hard, and I have not spelled this out below.

As recommended by Parham et al. (2011), minimum ages have generally been chosen in the literature as the youngest possible age of the calibrating specimen(s). This is practically guaranteed to result in ages that are too young for various reasons (Marshall, 2019). To account, if crudely, for non-zero branch lengths and especially for the nested phylogenetic positions of some calibrating specimens, and to counteract "the illusion of precision" (Graur and Martin, 2004: title) spread by calibration ages with five significant digits like 421.75 Ma [the minimum age chosen by Irisarri et al. (2017) for the root node, see below], I have rounded up (stratigraphically down) to the nearest million years, with a few exceptions suggested by mass extinction events.

Maximum ages are by default much more difficult to assign than minimum ages. Absence of proof is not proof of absence; absence of evidence is evidence of absence, but in most cases it is quite weak evidence. Yet, omitting maximum ages altogether and assigning only minimum ages to all calibrations automatically results in much too old divergence dates as nothing stops the 99.9% or 99.99% confidence or credibility intervals for all node ages from avoiding all overlap with the calibrated minimum ages. I have therefore followed Irisarri et al. (2017) and their sources in assigning as many maximum ages as I dare. For this purpose, I have basically followed the recommendations of Parham et al. (2011) and Pardo et al. (2020: 11), which amount to assigning a maximum age whenever we can reasonably expect (after preservation biases, collection biases, collection intensity, paleobiogeography, etc.) to have found evidence of the clade in question if it had been present at the time in question, but have not found any. This has widely been followed in the literature, but various compendia like Benton et al. (2015) have gone beyond this in many cases: in short, the oldest certain fossil provides the minimum age under that approach, while the oldest uncertain fossil of the same clade provides the maximum age. This practice is not defensible; therefore, I assign, in the aggregate, fewer and more distant maximum ages than Irisarri et al. (2017).

Given the limits of our current knowledge of the fossil record, all maximum ages might be expected to be soft bounds. In a few cases discussed below, however, I find that the absence of evidence is so hard to explain away that a hard maximum is justified. This generally concerns unrealistically old maxima that I have chosen because no younger maximum suggests itself. Ultimately, of course, this is subjective.

The choices of hard vs. soft bounds do not seem to make a great difference to the big picture. Due to practical constraints, a set of calibration ages mostly identical to the present ones was analyzed twice, with all bounds treated as soft or as hard, in the first preprint of this work (Marjanović, 2019); the results were quite similar to each other (Marjanović, 2019: figure 1 and table 2). Even so, however, in the run where all bounds were soft, most divergence dates were younger than in the run where all bounds were hard (usually negligibly so, but by 20 Ma in the extreme cases); the mean ages of some calibrated nodes even ended up younger than their minimum ages.

Calibrations

Because this journal imposes a space restriction, most of this section forms the **Supplementary Material**.

In the nine subsections below and the 20 sections of the **Supplementary Material**, I discuss the minimum and maximum ages of all 30 nodes used as calibrations by Irisarri et al. (2017), referring to each by clade names and by the node number assigned by Irisarri et al. (2017: especially supplementary table 8 and supplementary figure 19), also shown in **Figure 1**. The abbreviation Fm stands for Formation;

TABLE 2 | The ages found by Irisarri et al. (2017: supplementary table 9: last three columns) when all calibrations were used (all bounds treated as soft, mean ages averaged over 100 gene-jackknifed runs, extremes absolute over all runs), and the results obtained here with the updated calibrations (all bounds treated as hard, mean ages averaged over two runs with the full dataset, extremes absolute over both runs).

	Irisarri et al. (2017)			Present results		
Node number	**Mean age**	**Younger end of 95% CI**	**Older end of 95% CI**	**Mean age**	**Younger end of 95% CI**	**Older end of 95% CI**
100	460	452	465	472	467	475
101	393	383	403	390	363	415
102	437	431	440	462	453	471
103	426	420	431	435	427	445
104	412	408	418	423	420	4230
105	341	331	350	363	359	365
106	289	283	296	320	318	324
107	257	256	257	301	294	307
108	254	253	256	290	282	298
109	243	242	245	250	248	252
110	120	90	162	162	129	185
111	71	66	75	160	126	182
112	137	111	173	167	141	186
113	83	70	87	105	88	115
114	63	47	73	90	70	102
115	16	8	25	66	52	80
116	92	66	130	163	136	183
117	224	211	234	181	175	185
118	206	184	221	167	148	178
119	168	133	188	142	112	157
120	155	117	176	140	109	155
121	127	90	150	135	104	151
122	95	63	124	116	86	136
123	78	45	107	120	91	139
124	192	167	211	172	160	181
125	239	233	244	254	244	268
126	199	190	208	153	141	178
127	195	185	204	150	138	175
128	187	177	196	144	133	170
129	182	173	192	141	131	167
130	181	172	190	139	128	165
131	166	159	175	131	119	157
132	137	124	151	117	101	142
133	127	111	142	115	99	141
134	130	115	145	92	67	117
135	128	104	143	113	95	138
136	94	72	119	105	91	131
137	88	66	112	102	86	128
138	64	40	91	82	69	106
139	47	26	72	72	59	93
140	11	4	25	60	46	79
141	46	25	72	77	64	100
142	27	13	49	52	37	69
143	39	21	64	74	61	97
144	22	11	42	72	58	94
145	179	167	190	106	82	134
146	156	136	172	116	103	143
147	57	34	77	69	59	93

(Continued)

TABLE 2 | Continued

	Irisarri et al. (2017)			Present results		
Node number	**Mean age**	**Younger end of 95% CI**	**Older end of 95% CI**	**Mean age**	**Younger end of 95% CI**	**Older end of 95% CI**
148	44	24	65	69	58	93
149	165	146	181	138	123	164
150	165	161	172	229	217	233
151	138	136	140	142	128	157
152	94	91	96	71	69	72
153	89	85	92	68	67	70
154	61	53	65	54	50	56
155	79	71	84	66	65	67
156	91	87	94	64	56	69
157	68	62	72	66	61	68
158	50	38	60	59	51	67
159	315	300	328	328	314	339
160	307	290	323	286	275	290
161	202	173	237	170	137	188
162	192	163	226	165	131	183
163	177	146	210	143	110	160
164	168	137	199	139	106	156
165	117	86	143	104	71	119
166	92	62	117	60	36	70
167	77	49	101	59	35	69
168	53	30	74	45	26	56
169	162	134	196	143	107	163
170	201	170	232	173	156	193
171	192	161	224	170	154	192
172	186	154	218	166	148	188
173	155	123	186	147	127	172
174	105	71	140	104	85	141
175	94	62	127	74	57	109
176	70	33	110	75	60	117
177	54	22	89	71	56	111
178	156	119	189	128	104	152
179	144	106	177	123	99	147
180	160	125	194	129	105	152
181	213	162	255	181	145	247
182	155	105	195	152	116	221
183	36	12	65	71	47	116
184	223	165	279	324	289	353
185	78	48	107	164	131	190
186	6	2	15	46	25	68
187	414	402	428	390	385	404
188	293	256	332	289	253	321
189	202	140	269	154	124	187
190	156	92	223	131	101	167
191	98	50	168	74	53	114
192	207	172	262	193	184	201
193	76	42	110	73	53	115
194	380	370	390	380	352	406
195	345	338	352	276	253	298
196	330	319	340	254	222	286
197	55	18	91	118	58	175
198	277	244	297	167	119	207

*All calibration dates are shown in **Table 1**. All ages are rounded to whole Ma. CI, credibility interval.*

ICSC refers to the International Chronostratigraphic Chart v2020/3 (Cohen K. M. et al., 2020); Ma is the quasi-SI symbol for megayear (million years).

Root Node (100): Gnathostomata [PN] (Total Group Including Chondrichthyes – Pan-Osteichthyes [PN])

The cladogenesis that created the total groups of Chondrichthyes and Osteichthyes [PN] was assigned a minimum age of 421.75 Ma, a remarkably precise date close to the Silurian-Devonian boundary, and a maximum age of 462.5 Ma, which is currently (ICSC) thought to lie in the Darriwilian stage of the Middle Ordovician.

The Darriwilian should rather be regarded as the minimum age of this calibration date. While articulated bones and teeth of gnathostomes—both total-group chondrichthyans (Burrow and Young, 1999) and pan-osteichthyans (Choo et al., 2017, and references therein)—are only known from the Ludfordian (Ludlow, late Silurian) upward, a large diversity of scales that are increasingly confidently assigned to stem-chondrichthyans extends all the way down into the early Darriwilian (Sansom et al., 2012; Andreev et al., 2015, 2016a,b; Sansom and Andreev, 2018; Žigaitė-Moro et al., 2018; and references therein). The Darriwilian is currently thought to have begun 467.3 ± 1.1 Ma ago and to have ended 458.4 ± 0.9 Ma ago (ICSC); for the purposes of reducing "the middle part of the Stairway Sandstone" (Sansom et al., 2012, p. 243) to a single number, the age of 465 Ma should be adequate as the minimum age of Gnathostomata.

As a maximum age, I cautiously propose the mid-Floian (Early Ordovician) upper fossiliferous level of the Burgess-like Fezouata Shale; at both levels, gnathostomes are absent among the "over 200 taxa, about half of which are soft-bodied" (Lefebvre et al., 2017, p. 296). Note that the oldest known hard tissues of vertebrates are Floian in age as well (reviewed by Sansom and Andreev, 2018). The Floian began 477.7 ± 1.4 Ma ago and ended 470.0 ± 1.4 Ma ago (ICSC), so I suggest a soft maximum age of 475 Ma for this calibration date.

The minimum and the maximum age proposed here are unexpectedly close together. This may be a sign that one or both is an unduly optimistic assessment of our knowledge of the fossil record—or that the origin of Gnathostomata formed part of the Great Ordovician Biodiversification Event (Sansom et al., 2012; Sansom and Andreev, 2018), which does not seem implausible.

Node 105: Tetrapoda [PN] (Amphibia [PN] – Pan-Amniota [PN])

The divergence between the ancestors of lissamphibians and those of amniotes was assigned a minimum age of 330.4 and a maximum of 350.1 Ma following Benton and Donoghue (2007). Although Pardo et al. (2020) have reviewed the breadth of issues it (many raises far beyond the scope of this work), and although I broadly agree with their conclusions, a few points remain to be addressed or summarized.

For a long time, the oldest tetrapod was thought to be *Lethiscus*, variably supposed to be a stem-amphibian or a stem-pan-amniote (see below), which is mid-Viséan in age (Smithson et al., 2012, and references therein; the Viséan lasted from 346.7 ± 0.4 to 330.9 ± 0.2 Ma ago: ICSC). More likely, *Lethiscus* and the other aïstopods are rather early-branching stem-stegocephalians [PN] (Pardo et al., 2017, 2018; Clack et al., 2019; further discussion in Marjanović and Laurin, 2019). Whether *Casineria* from a geographically (southeastern Scotland) and stratigraphically close site (mid-late Viséan: Paton et al., 1999; Smithson et al., 2012) can replace it in that function depends on two unresolved issues: its own phylogenetic position, for which estimates range from very close to Amniota (within Tetrapoda) into Temnospondyli (Marjanović and Laurin, 2019, and references therein; Clack et al., 2019; Daza et al., 2020: supplementary figure S15), and the controversial phylogenetic position of Lissamphibia [PN] in the stegocephalian tree (Marjanović and Laurin, 2013a, 2019; Danto et al., 2019; Laurin et al., 2019; Daza et al., 2020; Pardo et al., 2020; and references in all five), which determines whether the temnospondyls are tetrapods or quite rootward stem-stegocephalians by determining which node of the otherwise largely stable tree of early stegocephalians bears the name Tetrapoda.

Anderson et al. (2015) reported a number of isolated anthracosaur [PN] (embolomere or eoherpetid) bones from a mid-Tournaisian site (the Tournaisian preceded the Viséan and began at the Devonian/Carboniferous boundary 358.9 ± 0.4 Ma ago: ICSC). Whether these are tetrapods depends on the relative positions of temnospondyls, anthracosaurs and other clades in that region of the tree (Pardo et al., 2018, 2020; Marjanović and Laurin, 2019; Ruta et al., 2020; and references in all four) in addition to the position of Lissamphibia: even if the lissamphibians are temnospondyls, the anthracosaurs may still be stem-stegocephalians.

The same site has also yielded the oldest colosteid remains (Anderson et al., 2015). Colosteidae ("Colosteida" of Pardo et al., 2020) was referred to Temnospondyli throughout the 20th century and found in that position by Marjanović and Laurin (2019) to our great surprise (also in some of the trees by Daza et al., 2020: supplementary figure S15); as pointed out by Pardo et al. (2020), this means it could belong to Tetrapoda. However, ongoing work on enlarging and improving the matrix of Marjanović and Laurin (2019) and Daza et al. (2020) shows that this result was most likely an artifact of the taxon and character sample; similarly, Ruta et al. (2020) found the colosteid they included to be a temnospondyl with weak support in their Bayesian analysis, but to lie rootward of Temnospondyli in their parsimony analyses (unweighted, reweighted, or with implied weighting).

The same site has further yielded tetrapod trackways, some of which are tetradactyl (Smithson et al., 2012, and references therein). Among Paleozoic tetrapods, tetradactyly is only known among "microsaurs" (including lysorophians), scincosaurids, some urocordylids, temnospondyls, and *Colosteus* (but not its close pentadactyl relative *Greererpeton*). (Reports of tetradactyl limbs in diplocaulids have been erroneous: Marjanović and Laurin, 2019; Milner, 2019, and references therein.) *Colosteus* and probably (Clack et al., 2019) the urocordylids are stem-stegocephalians, but both were fully aquatic, thus unlikely to leave trackways; "microsaurs" and probably scincosaurids were tetrapods, and most were amphibious to terrestrial;

temnospondyls spanned the full range of lifestyles, but see above for their phylogenetic position. In short, whether tetradactyl trackways are evidence of tetrapods in the mid-late Tournaisian remains unclear.

The oldest uncontroversial tetrapod is thus *Westlothiana* from close to the end of the Viséan (Marjanović and Laurin, 2019, and references therein, especially Smithson et al., 1994, 2012). Other stegocephalians from the same site and age may or may not be tetrapods: whether the temnospondyl *Balanerpeton* (Milner and Sequeira, 1994; Schoch and Milner, 2014) is one depends on the resolution of the abovementioned controversy about Lissamphibia; likewise, see above on the "anthracosaur-grade" (Marjanović and Laurin, 2019; Ruta et al., 2020) animals *Silvanerpeton* and *Eldeceeon*; *Ophiderpeton kirktonense* is an aïstopod, on which see above; *Kirktonecta* (Clack, 2011) is likely a tetrapod, but needs to be fully prepared or μCT-scanned before a confident assessment can be made.

Thus, the minimum age may be as young as roughly 335 Ma (mid-late Viséan) or as old as roughly 350 Ma (early-middle Tournaisian) depending on two phylogenetic problems.

The few Tournaisian stegocephalian sites discovered so far (Smithson et al., 2012; Anderson et al., 2015; Clack et al., 2016) have not yielded any uncontroversial tetrapods, temnospondyl bones, or temnospondyl footprints; thus, if the temnospondyls are stem-tetrapodomorphs, the ages of these sites (up to roughly 350 Ma) may be useful as a maximum age. However, as stressed by Pardo et al. (2020), they represent a very small region of the Carboniferous globe, so I continue (Marjanović and Laurin, 2019) to caution against this regardless of the phylogenetic issues. Rather, the richer and better studied Famennian (end-Devonian) record, which has not so far yielded close relatives of Tetrapoda but has yielded more rootward stegocephalians and other tetrapodomorphs (Marjanović and Laurin, 2019; Ahlberg and Clack, 2020; and references therein), should be used to place a soft maximum age around very roughly 365 Ma.

Node 106: Amniota [PN] (Pan-Mammalia [PN] – Sauropsida)

The cladogenesis that separated the total group of mammals (also called Synapsida [PN] or Theropsida: Goodrich, 1916) from the total group of diapsids including turtles (Sauropsida: Goodrich, 1916) was assigned a minimum age of 288 Ma (Artinskian, Early Permian) and a maximum age of 338 Ma (Viséan, Early Carboniferous).

This minimum age is rather puzzling. I am not aware of any doubts on the membership of *Hylonomus* in Sauropsida since its redescription by Carroll (1964), except the very vague ones presented by Graur and Martin (2004) and taken from even more outdated literature; none are mentioned in the review by Pardo et al. (2020) either. Because of its late Bashkirian age, this calibration has often been dated to 310 Ma (as discussed by Graur and Martin, 2004). Currently (ICSC), the Bashkirian is thought to have ended 315.2 ± 0.2 and begun 323.2 ± 0.4 Ma ago, and the site (Joggins, Nova Scotia) that has yielded *Hylonomus* has been dated to 317–319 Ma (Carpenter et al., 2015); thus, given the phylogenetic position of *Hylonomus* (Ford and Benson, 2019, and references therein), I suggest a minimum age of 318 Ma for this calibration.

There appears to be pan-mammalian material from the same site (Carroll, 1964; Mann et al., 2020), which has also yielded various "microsaurs" that Pardo et al. (2017) included in Sauropsida (see also Marjanović and Laurin, 2019, and Pardo et al., 2020). I should also emphasize that the next younger sauropsids and pan-mammals (and "microsaurs") older than 288 Ma come from several sites in each following geological stage (Moscovian through Artinskian) and represent a considerable diversity; from the Moscovian alone, four sites of successive ages are known that present more or less complete skeletons of uncontroversial amniotes, namely, sauropsids closely related to Diapsida and *Hylonomus* (*Anthracodromeus*, *Brouffia*, *Cephalerpeton*, *Paleothyris*), the oldest "parareptile" (*Carbonodraco*) as well as what appears to be the sister-group to most other sauropsids (*Coelostegus*), and, on the pan-mammalian side, ophiacodontids (*Echinerpeton*; *Archaeothyris* from two sites). A fifth site preserves the oldest varanopid, a group of amniotes of unclear phylogenetic position (Ford and Benson, 2018, 2019). As reviewed in detail by Pardo et al. (2020), this implies ghost lineages for several other amniote clades that might not have lived in coal swamps; several of these show up in the fossil record of the next and last two stages of the Carboniferous, which ended 298.9 ± 0.15 Ma ago (ICSC). For more information on the Carboniferous amniote record, see Reisz and Modesto (1996: figure 3), Müller and Reisz (2006), Maddin et al. (2019), Mann and Paterson (2019), Mann et al. (2019), and Pardo et al. (2020), the second and the third with phylogenetic analyses, as well as references in all six. Additionally, the oldest known diadectomorphs ("diadectamorphs" of Pardo et al., 2020) date from the Kasimovian ("Missourian" in Kissel, 2010) which follows the Moscovian; they may represent the sister-group of Amniota, or they may be what should have been called non-synapsid theropsids (Klembara et al., 2019; Marjanović and Laurin, 2019; Pardo et al., 2020; and references in all three).

The absence of amniotes (and diadectomorphs) in the Serpukhovian record preceding the Bashkirian should not be given much weight for paleoecological reasons, as reviewed by Pardo et al. (2020); note that "lepospondyls" like the Viséan *Kirktonecta* and *Westlothiana*, probably closely related to but outside Amniota, are almost unknown from this age as well (candidates were described by Carroll et al., 1991; Carroll and Chorn, 1995; Lombard and Bolt, 1999). Their absence from the somewhat richer Viséan record (discussed above) suffers in part from the same problem, in part from geographic restrictions. Thus, I refrain from recommending a maximum age other than that of the preceding Node 105, even though such an early age would imply very slow rates of morphological evolution in the earliest pan-mammals and sauropsids.

Node 107: Reptilia [PN] (Pan-Lepidosauria [PN] – Pan-Archelosauria [PN]); Node 108: Archelosauria [PN] (Pan-Testudines [PN] – Pan-Archosauria [PN])

The origin of the sauropsid crown group by a split into Pan-Lepidosauria and Pan-Archelosauria was assigned a minimum age of 252 Ma and a maximum age of 257 Ma, both in the Late

Permian. Ezcurra et al. (2014; correction: The PLOS ONE Staff, 2014) agreed that the oldest unambiguous reptile that can be clearly dated is the supposed pan-archosaur *Protorosaurus*, which is, however, 257.3 ± 1.6 Ma old as they also discussed. Therefore, they revised the minimum age to 255.7 Ma, the younger end of this confidence interval.

However, like all other recent phylogenetic analyses of molecular data, Irisarri et al. (2017) found the turtles to be closer to Archosauria [PN] than Lepidosauria [PN]. Thus, the question whether *Eunotosaurus* is a member of the turtle stem (Schoch and Sues, 2017, and references therein) becomes relevant, because the earliest occurrence of *Eunotosaurus* is roughly middle Capitanian in age (the Capitanian, the last stage of the Middle Permian, ended 259.1 ± 0.5 Ma ago and began 265.1 ± 0.4 Ma ago: ICSC), and further because *Protorosaurus* would presumably belong to Pan-Archosauria and thus calibrate Node 108, not 107.

For present purposes, I set the minimum age of Archelosauria (Node 108) as 263 Ma, the approximate midpoint of the Capitanian, and do not assign a minimum age to Reptilia (Node 107). However, in general, I have to, at our current level of understanding, recommend against using either of these nodes as a calibration. The reason are two major uncertainties about the topology of the phylogenetic tree.

First, if *Eunotosaurus* has moved from the "parareptiles" well outside Diapsida [PN]—or well inside Diapsida, though presumably still in its stem-group (Ford and Benson, 2019)—to the turtle stem within the crown group of Diapsida (i.e., Reptilia [PN]), do any other "parareptiles" follow it? The oldest known member of that assemblage, *Carbonodraco*, comes from the site of Linton in Ohio (Mann et al., 2019), which is about 307–308 Ma old (compare Reisz and Modesto, 1996; Carpenter et al., 2015), so that should be the minimum age of Archelosauria if all "parareptiles" are archelosaurs; the currently available phylogenetic analyses of "parareptiles" (Laurin and Piñeiro, 2018; MacDougall et al., 2019) have not adequately tested this question. While Schoch and Sues (2017) did test the mutual relationships of "parareptiles," *Eunotosaurus*, and diapsids and found *Eunotosaurus* nested in the latter, several nodes away from the former, these nodes were very poorly supported. The character and taxon samples of all existing matrices for analyses of amniote phylogeny need to be substantially improved (Ford and Benson, 2018, 2019; Laurin and Piñeiro, 2018; MacDougall et al., 2019; Mann et al., 2019); Ford and Benson (2019) made a large step in that direction, but deliberately excluded *Eunotosaurus* and the turtles from their analysis so as not to have to deal with all problems at the same time.

Second, the position of *Protorosaurus* as a pan-archosaur, accepted for decades, was thrown into doubt by Simões et al. (2018), who found it as such in their Bayesian analyses of morphological or combined data (Simões et al., 2018: ext. data figures 5, 6; also, after a few changes to the dataset, Garberoglio et al., 2019: figure S2; Sobral et al., 2020: figures S9, S10), but not in their parsimony analyses of morphological data without or with implied weights (ext. data figures 3, 4; likewise Garberoglio et al., 2019: figure S3; Sobral et al., 2020: figure S7, S8), where it came out as a stem-sauropsid; the question was unresolved in their Bayesian tip-dating or tip-and-node dating analyses of combined data (ext. data figures 7, 8). After a different set of changes to the dataset, Simões et al. (2020) found *Protorosaurus* as a pan-archosaur when they used MrBayes (supplementary figures 2–5) or when they used BEAST for dating with a correction (supplementary figure 7), but not when they used BEAST for dating without a correction (supplementary figure 6). Support was moderate throughout. However, these trees are hard to compare to that of Irisarri et al. (2017) because they all find the turtles outside the diapsid crown (with limited support); no extant archosaurs or turtles, and therefore no molecular data for them, are included in these datasets. Using a smaller dataset with much denser sampling of Triassic reptiles, Pritchard et al. (2018) found *Protorosaurus* closer to Archosauria than to Lepidosauria with very strong support (parsimony bootstrap value: 100%, Bayesian posterior probability: 99.06%), but whether that is on the archosaur or the archelosaur stem could not be determined because there were no turtles in that dataset.

The maximum age of either node is likewise difficult to narrow down. Uncontroversial diapsids have a notoriously patchy Paleozoic record (Ford and Benson, 2018, and references therein); the same holds for "parareptiles," which have only two known Carboniferous records so far (Modesto et al., 2015; Mann et al., 2019). I cannot express confidence in a maximum age other than that of Node 106, which I cannot distinguish from the maximum age of Node 105 as explained above. This leaves Node 107 without independent calibrations in the current taxon sample.

Node 113: Neognathae (Galloanserae [PN] – Neoaves)

The last common ancestor of *Anas*, *Gallus*, and *Meleagris* on one side and *Taeniopygia* on the other was assigned a minimum age of 66 Ma (the Cretaceous/Paleogene boundary) and a maximum age of 86.5 Ma (Coniacian/Santonian boundary, Late Cretaceous) following Benton and Donoghue (2007).

The oldest known neognath appears to be the presbyornithid stem-anserimorph (Elżanowski, 2014; Tambussi et al., 2019; within two steps of the most parsimonious trees of Field et al., 2020) *Teviornis* from somewhere low in the Late Cretaceous Nemegt Fm of Mongolia; it is known only from a carpometacarpus, two phalanges, and the distal end of a humerus that all seem to belong to the same right wing (Kurochkin et al., 2002). The most recent work on the specimen has bolstered its presbyornithid identity (De Pietri et al., 2016), even though the next younger presbyornithids are middle or late Paleocene (i.e., younger than 61.6 Ma: ICSC).

The age of the Nemegt Fm is difficult to pin down; radiometric dating of this or adjacent formations has not been possible, and the only fossils available for biostratigraphy are vertebrates that have to be compared to those of North America where marine correlations and radiometric dates are known. These comparisons favor a vaguely early Maastrichtian age, without ruling out a Campanian component. Magnetostratigraphic evidence was reported in a conference abstract by Hicks et al. (2001); I have not been able to find a follow-up publication. Hicks et al. (2001) stated that the sampled sections from the Nemegt and the conformably underlying Baruungoyot Fm "can

be quite reliably correlated to the Geomagnetic Reversal Time Scale [. . .] and clearly lie in the Campanian/Maastrichtian interval that extends from the uppermost part of subchron C33n, through chron 32 into the lower half of chron 31." Where the Baruungoyot/Nemegt boundary lies on this scale was not mentioned. The upper boundary of the Nemegt Fm is an unconformity with a Paleocene formation.

Hicks et al. (2001) also studied the Late Cretaceous Djadokhta Fm, finding that "a distinct reversal sequence is emerging that allows us to correlate the sections in a preliminary way to the late Campanian through Maastrichtian interval that ranges from C32 to C31." While I have not been able to find a publication by an overlapping set of authors on this finding, it agrees at least broadly with Dashzeveg et al. (2005: 18, 26, 27), whose own magnetostratigraphic work on the Djadokhta Fm indicated "that the sediments were deposited during the rapid sequence of polarity changes in the late part of the Campanian incorporating the end of Chron 33 and Chron 32 between about 75 and 71 Ma [. . .]. However, this tentative correlation to the Geomagnetic Polarity Timescale cannot yet be certainly established." Hasegawa et al. (2008) disagreed with the stratigraphy by Dashzeveg et al. (2005), but not with their dating.

Most often, the Djadokhta Fm has been thought to underlie the Baruungoyot Fm, but a contact between the two has not so far been identified (Dingus et al., 2008; cited without comment e.g., by Chinzorig et al., 2017); they could be partly coeval (references in Hasegawa et al., 2008). Still, it seems safe to say that most of the Nemegt Fm is younger than most of the Djadokhta Fm.

According to Milanese et al. (2018: Figure 12), the Campanian-Maastrichtian boundary (72.1 ± 0.2 Ma ago: ICSC) lies near the end of chron 32. The Djadokhta Fm thus corresponds to the end of the Campanian, the Baruungoyot Fm should have at most the same age, and the youngest magnetostratigraphic sample from the Nemegt Fm, in the earlier half of chron 31, should be about 70 Ma old. Given the stratigraphic position of *Teviornis* low within the formation and its nested phylogenetic position within Neognathae, I propose 71 Ma (within the same subchron as 70 Ma: Milanese et al., 2018: Figure 12) as the soft minimum age of the present calibration.

Field et al. (2020, p. 400) stated that the likely stem-pangallanseran "*Asteriornis* provides a firm calibration point for the minimum age of divergence of the major bird clades Galloanserae and Neoaves. We recommend that a minimum age of 66.7 million years is assigned to this pivotal neornithine node in future divergence time studies, reflecting the youngest possible age of the *Asteriornis* holotype including geochronological uncertainty." In their supplementary information (p. 13), however, they revealed being aware of *Teviornis*, citing De Pietri et al. (2016) for its position as a presbyornithid (and thus, by their own phylogenetic analyses, an anserimorph) without discussing it any further.

Should the fragmentary *Teviornis* fall out elsewhere, the minimum age might nonetheless not have to rest on *Asteriornis*, because Vegaviidae, a clade containing the late Maastrichtian (Clarke et al., 2005; Salazar et al., 2010) *Vegavis*, *Polarornis*, and *Neogaeornis* and probably the end-Campanian (McLachlan et al., 2017) *Maaqwi*, has been found on the anserimorph stem in some of the latest analyses (Agnolín et al., 2017; Tambussi et al., 2019). However, Mayr et al. (2018) discussed reasons for skepticism, and the analyses of McLachlan et al. (2017), Bailleul et al. (2019: supplementary trees 7–11, 16, 17), Field et al. (2020), and O'Connor et al. (2020) found the vegaviids they included close to but outside Aves (or at least Galloanserae in the case of Bailleul et al., 2019; O'Connor et al., 2020, who did not sample Neoaves or Palaeognathae in the analyses in question).

As the soft maximum age, I tentatively suggest 115 Ma, an estimate of the mid-Aptian age of the terrestrial Xiagou Fm of northwestern China, which has yielded a diversity of stem-birds but no particularly close relatives of the crown (Wang et al., 2013; Bailleul et al., 2019; O'Connor et al., 2020; and references therein).

Node 117: Testudines [PN] (Pan-Pleurodira [PN] – Pan-Cryptodira [PN])

The origin of the turtle crown group by split into the pleurodiran [PN] and cryptodiran [PN] total groups was assigned a minimum age of 210 Ma and no maximum age; this was taken from Noonan and Chippindale (2006), who had cited a work from 1990 as their source.

The calibration dates treated above, and correspondingly in the **Supplementary Material**, are almost all too young (some substantially so, others by just a few million years). This one, in contrast, is far too old. It rests on the outdated interpretation of the Norian (Late Triassic) *Proterochersis* as a stem-group pan-pleurodire. With one short series of exceptions (Gaffney et al., 2006, 2007; Gaffney and Jenkins, 2010), all 21st-century treatments of Mesozoic turtle phylogeny have found *Proterochersis* and all other turtles older than those mentioned below to lie well outside the crown group (Shao et al., 2018: figures S8, S9; Sterli et al., 2019, 2020; and references therein, in Gaffney and Jenkins, 2010; Romano et al., 2014a).

The three oldest known xinjiangchelyids [PN], of which one was referred to *Protoxinjiangchelys*, seem to be between 170 and 180 Ma old (Aalenian/Bajocian boundary, Middle Jurassic, to Toarcian, late Early Jurassic; Hu et al., 2020, and reference therein). In the last 3 years, the xinjiangchelyids have been found as stem-testudinates or as stem-pan-cryptodires (Shao et al., 2018; Evers et al., 2019; González Ruiz et al., 2019: Figure 6 and supplementary figure 4; Gentry et al., 2019; Anquetin and André, 2020; Sterli et al., 2020: supplementary figure "X" = 19), even in both positions when the same matrix was analyzed with different methods (Sterli et al., 2019: **supplementary file** SterlietalSupplementary_material_3.pdf).

The oldest known securely dated and securely identified crown-group turtle is thus the mid-late Oxfordian stem-pan-pleurodire *Caribemys* (de la Fuente and Iturralde-Vinent, 2001; Shao et al., 2018; mostly referred to *Notoemys* as *N. oxfordiensis* in more recent literature, e.g., Sterli et al., 2019). Given that the Oxfordian ended 157.3 ± 1.0 Ma ago (ICSC), I suggest a minimum age of 158 Ma.

The stem-pan-trionychian [PN] cryptodire *Sinaspideretes* (Tong et al., 2013), which would provide a minimum age for Cryptodira (node 118) rather than only Testudines, was long thought to have the same age or to be somewhat older. Of the three known specimens, at least one (the exact localities where

the type and the other specimen were found are unknown) comes from the Upper (Shang-) Shaximiao Fm (Tong et al., 2013), which conformably overlies a sequence of two supposedly Middle Jurassic formations and is overlain by two Upper Jurassic formations (Tong et al., 2011; Xing et al., 2013), so it should be about Oxfordian to Callovian in age. The biostratigraphic evidence for the age of the Upper Shaximiao Fm is conflicting; there was no consensus on whether it is Middle or Late Jurassic (Xing et al., 2013) before Wang et al. (2018) showed that the immediately underlying Lower (Xia-) Shaximiao Fm is at most 159 ± 2 Ma old, a confidence interval that lies entirely in the Late Jurassic (which began, with the Oxfordian, 163.5 ± 1.0 Ma ago: ICSC). Most likely, then, the same holds for all *Sinaspideretes* specimens, and none of them is older than *Caribemys*.

The unambiguously Early Jurassic and Triassic record of turtles throughout Pangea lies entirely on the stem and has a rather good stratigraphic fit (see Sterli et al., 2019, 2020). I therefore suggest a soft maximum age of 185 Ma (in the Pliensbachian: ICSC) that probably postdates all of these taxa but predates the oldest possible age of the oldest known xinjiangchelyids.

Node 129: Toxicofera (Pan-Serpentes [PN] – Anguimorpha + Pan-Iguania [PN])

This calibration was given a minimum age of 148 Ma (Tithonian, Late Jurassic) and no maximum age. Note that the minimum age was not operational because Node 131, Iguania [PN], was given an older minimum age of 165 Ma (see **Supplementary Material**); in other words, Node 129 was really not calibrated at all.

Indeed, I should first mention that the pan-squamate fossil record suffers from three problems that make it difficult to calibrate this node. First, it exhibits Carroll's Gap (Marjanović and Laurin, 2013a) very strongly. After the Middle Triassic stem-pan-squamate *Megachirella* and at least one Early Triassic pan-lepidosaur that may or may not be a pan-squamate (*Sophineta* in particular—compare the different phylogenetic analyses in Simões et al., 2018, 2020), the pan-squamate record as known today goes completely silent (see Node 131 for the one or two supposed exceptions) until the dam suddenly breaks in the Bathonian (Middle Jurassic) and representatives of the stem as well as, by current understanding, several parts of the crown appear in several sites in the northern continents and northernmost Gondwana. Second, these early representatives are all isolated and generally incomplete bones that preserve few diagnostic characters; the oldest complete skeletons come from one Tithonian (latest Jurassic) cluster of sites (Conrad, 2017), followed by a few Early Cretaceous ones as well as the oldest partially articulated material other than *Megachirella*. Third, the morphological datasets so far assembled for analysis of pan-squamate phylogeny are all so plagued by correlated characters and other problems that all of them support either Pan-Iguania as the sister-group to all other squamates, or the amphisbaenians (alone or even together with the dibamids) as the sister-group to Pan-Serpentes (e.g., Simões et al., 2020: supplementary figure 2), or both (e.g., Conrad, 2017: Figures 27, 28), while both are strongly contradicted by the molecular consensus (e.g., Irisarri et al., 2017; Garberoglio et al., 2019; Simões et al., 2020: supplementary figures 1, 3, 5, 8; Sobral et al., 2020: figure S10).

[As I try to redate the exact tree topology of Irisarri et al. (2017), it is not relevant to the present work that interesting doubts about parts of the molecular consensus have been raised from the molecular data, most recently and thoroughly by Mongiardino Koch and Gauthier (2018), who also reviewed that issue.]

The oldest known toxicoferans appear to be represented by four isolated vertebral centra from the Anoual Fm of Morocco, which is early Bathonian in age (Haddoumi et al., 2015). These bones were assigned to "cf. *Parviraptor*" by Haddoumi et al. (2015). Other material—vertebrae and jaw fragments from Europe and North America discussed in Panciroli et al. (2020)—was originally assigned to "cf." or "aff. *Parviraptor,*" including but not limited to the late Bathonian or earliest Callovian *Eophis*, the Kimmeridgian *Diablophis* and *Portugalophis*, and *Parviraptor* itself from around the Jurassic/Cretaceous (Tithonian/Berriasian) boundary. Traditionally regarded as representing the oldest anguimorphs, these fossils would calibrate Node 130, the split between Pan-Iguania [PN] and Anguimorpha; however, phylogenetic analyses following a redescription of much of the material have found it to constitute the oldest known pan-serpents, thus calibrating Node 129 (Caldwell et al., 2015; Martill et al., 2015; by implication Conrad, 2017; accepted without analysis by Garberoglio et al., 2019; Simões et al., 2020; Schineider Fachini et al., 2020). As the Bathonian began 168.3 ± 1.3 Ma ago and ended 166.1 ± 1.2 Ma ago, i.e., with uncertainty ranges that overlap in the middle (ICSC), the suggestion of 167 Ma by Caldwell et al. (2015) would then be a reasonable minimum age for this calibration.

Alifanov's (2019) casual referral of *Parviraptor* to an unusually large version of Mosasauria should not be construed to contradict this: the Cretaceous aquatic squamates, mosasaurs included, are probably all pan-serpents (see below), unless they lie on the common stem of Anguimorpha and Iguania (Simões et al., 2020: supplementary figure 8, with very low support).

As mentioned, all these remains are very fragmentary, and all are disarticulated; according to a reviewer, new, apparently unpublished material shows the "parviraptorids" are not snakes, and indeed Panciroli et al. (2020) were careful not to state in the text whether they agreed with the referral to the snake stem, designating "cf. *Parviraptor* sp." as "Squamata indet." in their faunal list (**Table 1**).

The next younger record of a possible toxicoferan is the just as fragmentary Callovian *Changetisaurus*, a supposed anguimorph, though Alifanov (2019) provided reasons to doubt that it is a toxicoferan. It is followed by the several species of *Dorsetisaurus*, another assemblage of skull fragments with osteoderms from the Kimmeridgian through Berriasian of Europe and North America, that was explicitly accepted as an anguimorph by Caldwell et al. (2015) and, on different grounds, Alifanov (2019), but has not, to the best of my knowledge, been included in any phylogenetic analysis. (Older and secondary literature has often claimed that the oldest *Dorsetisaurus* specimens are 148 Ma old, but the Kimmeridgian ended 152.1 ± 0.9 Ma ago: ICSC.)

Most of the rich record of Cretaceous aquatic squamates has traditionally been referred to Anguimorpha, but more likely belongs to Pan-Serpentes (e.g., Garberoglio et al., 2019; Palci et al., 2019; Sobral et al., 2020: figure S10; Simões et al., 2020: supplementary figures 3, 4, 6, 9; and references therein). It sets in in what seems to be the Hauterivian with *Kaganaias* (Evans et al., 2006; Campbell Mekarski et al., 2019); the Hauterivian ended ∼129.4 Ma ago (ICSC, uncertainty not quantified). If neither the "parviraptorids" nor *Changetisaurus* nor *Dorsetisaurus* are accepted as toxicoferans, the minimum age of Node 129 should thus be 130 Ma. To err on the side of caution, that is the age I have used here.

Due to Carroll's Gap (Marjanović and Laurin, 2013a), I agree with Irisarri et al. (2017) in not assigning a maximum age other than that for Node 125 (**Supplementary Material**).

Node 152: Placentalia [Atlantogenata – Boreo(eu)theria)]; Node 153: Boreo(eu)theria (Laurasiatheria – Euarchontoglires/Supraprimates)

The origin of Placentalia, the crown group of Eutheria, was given a minimum age of 95.3 Ma (Cenomanian, Late Cretaceous) and a maximum age of 113 Ma (Aptian/Albian boundary, Early Cretaceous) following Benton and Donoghue (2007). Its immediate descendant nodes were not constrained.

The minimum age rests on the assumption, commonly but not universally held in 2007, that the zhelestids are "ungulates," i.e., belong to Placentalia, or perhaps even that the zalambdalestids are related to Glires and therefore belong to Placentalia. For a long time now, as already pointed out by Parham et al. (2011), every reinvestigation of the anatomy of these Cretaceous animals, and every phylogenetic analysis that sampled Cretaceous eutherians densely (i.e., not including Zhou et al., 2019: supplementary inf. M), has found them on the eutherian stem, often not even particularly close to Placentalia (e.g., Novacek et al., 1997; Asher et al., 2005, 2019; Wible et al., 2009; Goswami et al., 2011; Halliday et al., 2015; Manz et al., 2015; Bi et al., 2018: figures 2 and SI-1; Wang et al., 2019: ext. data figure 5; and references in Parham et al., 2011; see also Fostowicz-Frelik and Kielan-Jaworowska, 2002).

A few terminal Cretaceous (late Maastrichtian) eutherians have been attributed to Placentalia in the past. This is at best dubious for all of them. *Protungulatum* (Wible et al., 2009; Halliday et al., 2015, 2019: figure 1 contrary to the text; Manz et al., 2015: figure 2a; Wang et al., 2019: ext. data figure 5; Mao et al., 2019: figure S9) and *Gypsonictops* (Halliday et al., 2015, 2019; Manz et al., 2015: figure 2; Bi et al., 2018; Wang et al., 2019: ext. data figure 5; Mao et al., 2019: figure S9) are now placed close to but consistently outside Placentalia. *Deccanolestes*—at least if the teeth and the tarsal bones belong together—is placed far away (Goswami et al., 2011 [see there also for *Sahnitherium*]; Manz et al., 2015: figures 2 and I-1; Penkrot and Zack, 2016; Halliday et al., 2019). The single worn tooth named *Kharmerungulatum*, which had been assigned to Placentalia mostly through comparison to *Protungulatum* in the first place (Prasad et al., 2007), has more recently been found outside Placentalia as well ("Although none of the strict consensus trees supported the placement of *Kharmerungulatum* within the placental crown group, the limited dental material for this taxon proved insufficient for resolving its phylogenetic relationships, and so it was removed a posteriori from the MPTs to produce the reduced strict consensus trees."—Goswami et al., 2011, p. 16334), specifically as an adapisoriculid like *Deccanolestes* when full molecular constraints were applied by Manz et al. (2015: figure 2b). The stylinodontid taeniodont *Schowalteria* (Fox, 2016, and references therein) belongs to a clade that survived into the Eocene; the conference abstract by Funston et al. (2020) reported that a very large phylogenetic analysis has found the group outside Placentalia.

The same reasons make it difficult to decide which of the earliest Paleocene eutherians should be accepted as securely enough identified placentals, but in any case, Williamson et al. (2019, p. 220) reported that the herbivorous periptychid *Ectoconus*, estimated to have reached about 100 kg, was "present within 0.4 Ma of the K-Pg boundary"; phylogenetic analyses have found it to be not only a placental, but a laurasiatherian—Halliday et al. (2015; regardless of constraints) found it and the other periptychids on the pholidotan stem; Halliday et al. (2019), using combined data and maximum likelihood, found a comparable result with much less resolution; Püschel et al. (2019), using a somewhat smaller matrix with, however, a focus on periptychids and new data on them, recovered them as stem-artiodactylomorphs. I therefore suggest 66 Ma, the Cretaceous/Paleogene boundary (66.021 ± 0.081 Ma: Clyde et al., 2016), as the minimum age for Node 153, the basal node of Boreoeutheria (a name apparently coined by accident by Murphy et al., 2001) or simply Boreotheria (explicitly coined by Waddell et al., 2001). For Node 152, I cannot recommend a separate minimum age.

Unambiguous placentals continue to be absent worldwide in the rich Maastrichtian record (see above as well as Halliday et al., 2016; Davies et al., 2017), and even ambiguous ones except *Gypsonictops* continue to be absent in the even richer Campanian record (although there are three isolated Turonian teeth indistinguishable from both species of *Gypsonictops*: Cohen and Cifelli, 2015; Cohen, 2017), despite the presence of stem-eutherians (all northern continents, Madagascar, and India), stem-metatherians (Asia and North America), and ecologically comparable spalacotheroids (Asia and North America), meridiolestidans (South America) and gondwanatheres (South America, Madagascar, India, and some point between the late Turonian and latest Campanian of Africa—O'Connor et al., 2019). Although the Late Cretaceous fossil record of Africa is too limited to exclude the presence of placentals, and Antarctica, Australia, and New Zealand have no known Late Cretaceous mammal record at all, biogeographic parsimony does not favor the presence of Campanian or Maastrichtian placentals on these paleocontinents (e.g., Huttenlocker et al., 2018): the closest known relatives of Placentalia come from North America, followed by Asian forms, while the Indian eutherians (discussed above) are quite distant from Placentalia and the incomplete tooth from Madagascar is similarly identified as zhelestid (Averianov et al., 2003). Neither the Cenozoic fossil record nor molecular phylogenetics suggest an African origin as most parsimonious either, let

alone a more southeastern one. Therefore, I suggest the Campanian/Maastrichtian boundary, rounded to 72 Ma, as the hard maximum age for Node 152. (I cannot make a separate recommendation for Node 153.) This is more generous than the result of Halliday et al. (2016), 95% of whose reconstructions of the age of Placentalia were 69.53 Ma old or younger. The discrepancy to the published molecular ages (references in Halliday et al., 2016) is most likely due to the effects of body size (Berv and Field, 2017; Phillips and Fruciano, 2018), or perhaps other factors like generation length, on rates of molecular evolution.

At this point, readers may be wondering why I have mentioned neither the extremely large phylogenetic analysis by O'Leary et al. (2013) nor the objections by Springer et al. (2019), who wrote in their abstract that "morphological cladistics has a poor track record of reconstructing higher-level relationships among the orders of placental mammals". It would be more accurate to say that phylogenetic analysis of morphological data has *no* track record of reconstructing the phylogeny of Placentalia, good *or* bad. To avoid long-branch attraction and long-branch repulsion, any such analysis of morphological data will have to sample the enormous and poorly understood diversity of Paleo- and Eocene eutherians very densely, which will have to entail sampling enough of the characters that unite and distinguish them without falling into the trap of accumulating redundant or otherwise correlated characters that inevitably distort the tree (Marjanović and Laurin, 2019; Sookias, 2019; Celik and Phillips, 2020; and references in all three). This is so much work, and so hard to get funded, that—at the most generous count—only three attempts at such a matrix have ever been made; I should also point out that matrices of such sizes were not computationally tractable until a few years ago, at least not in less than a few months of calculation time. The first attempt is the "phenomic" matrix by O'Leary et al. (2013); as Springer et al. (2019) pointed out repeatedly, it contains no less than 4541 characters—but several hundred of these are parsimony-uninformative (O'Leary et al., 2013), and many others are redundant, which means they represent a smaller number of independent characters of which many are weighted twice or more often. At 86 terminal taxa, almost all of which are extant, the taxon sample is hopelessly inadequate for eutherian phylogeny. It is no surprise that parts of the topology are highly implausible (e.g., the undisputed stem-whale *Rodhocetus* landing on the common ungulate [PN] stem, as pointed out by Springer et al., 2019) and that even such undisputed clades as Afrosoricida, Lipotyphla, and Artiodactyla are no longer recovered when the hundreds of soft-tissue characters, which cannot be scored for the extinct terminal taxa, are removed (Springer et al., 2019), which casts doubt on the ability of that matrix to place extinct taxa accurately. The second attempt began in the doctoral thesis of Zack (2009) and was further modified and merged with other datasets in Halliday's doctoral thesis that culminated in the publication of Halliday et al. (2015). The taxon sample contains an appreciable number of Cretaceous and Paleocene eutherians; the character sample is of course more modest and contains, as usual for mammals, a large proportion of tooth characters, some of which might be redundant (e.g., Kangas et al., 2004; Harjunmaa et al., 2014). The further improved version (Halliday et al., 2019) suffers from the drawback that all characters were reduced to two states to make the matrix tractable by maximum-likelihood software; this throws away a lot of information (probably for no gain: Sansom et al., 2018; King, 2019). The third is that of the PalM group; funded by an enormous grant, it involves a lot of people each revising a group of Paleo- or Eocene eutherians as their doctoral thesis and contributing the gained knowledge (e.g., Napoli et al., 2017) to a growing matrix (ultimately based on that of Wible et al., 2009) that will then be evaluated for character redundancy and other issues. The only phylogenetic publications that have yet resulted are conference abstracts, of which I have cited Püschel et al. (2019) and Funston et al. (2020) above.

Springer et al. (2019) went on to claim that "Sansom and Wills (2013) showed that fossils are more likely to move stemward than crownward when they are only known for biomineralized characters." Indeed, Sansom and Wills (2013) made that claim. They had taken 78 neontological matrices of extant animals with biomineralized tissues, deleted the data for soft-tissue characters from random taxa, and found that those taxa changed their phylogenetic position significantly more often than random, and further underwent "stemward slippage" as opposed to "crownward slippage" significantly more often than random. Deleting data from hard-tissue characters instead had no such effect. Sansom and Wills (2013) concluded that some mysterious factor causes hard-tissue characters to contain a systematically misleading signal much more often than soft-tissue characters do, and that therefore the phylogenetic positions of all taxa known only from hard tissues—in other words most animal fossils—are highly suspect of falsely appearing more rootward than they really are. Therefore, fossils assigned to various stem groups could really belong to the respective crown groups, and the minimum ages of divergence-date calibrations could be systematically too young (Sansom and Wills, 2013), just as Springer et al. (2019) believed. A much simpler explanation is available: hard-tissue characters are unreliable *specifically among extant species* because the hard-tissue anatomy of extant species is usually very poorly known. For example (Marjanović and Witzmann, 2015), the vertebrae of some of western and central Europe's most common newt species are simply unknown to science, even after 200 years or more of research, because neontologists have focused on soft-tissue anatomy, behavior, and, more recently, the genome while treating the skeleton as an afterthought. The vertebrae of salamandrids are at least known to contain a phylogenetic signal—whether the appendicular skeleton also does is anybody's guess at this point! As our knowledge of the skeletons of extant taxa would improve, so would, I predict, the ability of hard-tissue characters to accurately resolve the phylogenetic positions of extant taxa.

Node 188: Crown Group of Elasmobranchii (Selachimorpha – Batomorpha)

The origin of the elasmobranch crown group by split into Selachimorpha (sharks) and Batomorpha (rays and skates) was given a minimum age of 190 Ma (Sinemurian/Pliensbachian boundary, Early Jurassic) and no maximum age. (Note that the name Neoselachii is consistently treated in the paleontological literature as if defined by one or more apomorphies, not by tree

topology; it probably applies to a clade somewhat larger, and possibly much older, than its crown group.)

Any attempt to date this cladogenesis suffers from the fact that the elasmobranch fossil record consists mostly of "the tooth, the whole tooth and nothing but the tooth" (as has often been said about the Mesozoic mammalian fossil record); scales and the occasional fin spine do occur, but more substantial remains are very rare. The shape of tooth crowns is naturally prone to homoplasy, the number of phylogenetically informative characters it offers is easily overestimated due to correlations between them (e.g., Kangas et al., 2004; Harjunmaa et al., 2014; Celik and Phillips, 2020; see node 157 in the **Supplementary Material**), and histological studies, which are needed to determine the states of certain characters (e.g., Andreev and Cuny, 2012; Cuny et al., 2017), have not been carried out on all potentially interesting tooth taxa.

Consequently, there is not as much interest in phylogeny among specialists of early elasmobranchs than among specialists of early mammals or early dinosaurs. This goes so far as to affect the use of terminology: Andreev and Cuny (2012) mentioned "stem selachimorphs" in the title of their work, implying that they understood Selachimorpha as a clade name, but quietly revealed it to be the name of a paraphyletic assemblage on p. 263 by stating that bundled enameloid is "diagnostic for Neoselachii exclusive of batomorphs, i.e., Selachimorpha", and their consistent referral of Synechodontiformes (see below) to "Selachimorpha" is not necessarily a referral to the crown group—even though they called bato- and selachomorphs sister-groups in the next sentence.

A safe minimum age of 201 Ma, used here, is provided by the oldest unambiguous crown-group selachimorph, the total-group galeomorph *Agaleus*, dating from the Hettangian, apparently close to its beginning (Stumpf and Kriwet, 2019, especially figure 5, and references therein), which was the beginning of the Jurassic and happened 201.3 ± 0.2 Ma ago (ICSC); I round this down (stratigraphically up) to avoid breaching the mass extinction event at the Triassic/Jurassic boundary. The oldest batoid batomorph is only sightly younger, see Node 192 (**Supplementary Material**).

However, this may err very far on the side of caution. Indeed, for purposes beyond the present work, I must recommend against using the minimum age of this divergence to calibrate a timetree for at least as long as the histology of Paleozoic "shark" teeth has not been studied in much more detail in a phylogenetic context. As if by typographic error, the oldest widely accepted crown-group elasmobranch is not 190 but about 290 Ma old: the oldest fossils referred to the neoselachian *Synechodus* are four teeth of Sakmarian age (referred to *S. antiquus*, whose type tooth comes from the following Artinskian age: Ivanov, 2005; Stumpf and Kriwet, 2019), and the Sakmarian ended 290.1 ± 0.26 Ma ago (ICSC). Teeth referred to other species of *Synechodus* range into the Paleocene; *S. antiquus* is the only Permian species (Andreev and Cuny, 2012). The histology of *S. antiquus* remains unknown as of Koot et al. (2014); nonetheless, Cuny et al. (2017, p. 61) regarded *S. antiquus* as "[t]he first proven selachimorph". Rounding up, this would suggest suggest 291 Ma as the minimum age of this calibration.

(My previous suggestion—Marjanović, 2019—to use that age as a soft minimum was incoherent, as a reviewer pointed out. A soft minimum would imply that a tail of the probability distribution of the age of this node would extend to younger ages than 291 Ma, so that an age of 290 Ma would be treated as much more probable than an age of 201 Ma. The opposite is the case: both 291 and 202 are much more probable than 290, which is younger than one potential minimum age but far older than the other. If *Synechodus antiquus* is a crown-group elasmobranch, so that 291 Ma is "the correct" minimum age, 290 is impossible; if it is not a crown-group elasmobranch, so that 201 is "correct," 290 is so much older as to be much less probable than, say, 205 or 210.)

Potential crown-group elasmobranchs older than 291 Ma are known: Andreev and Cuny (2012) and Cuny et al. (2017, p. 69) suggested that the tooth taxa *Cooleyella* and *Ginteria* could be stem-batomorphs. The oldest known *Cooleyella* specimen dates from around the end of the Tournaisian (Richards et al., 2018), which occurred 346.7 ± 0.4 Ma ago (ICSC); *Ginteria* appeared in the following Viséan stage. Cuny et al. (2017, p. 21, p. 69) further pointed out that *Mcmurdodus*, a tooth taxon that first appeared around the Early/Middle Devonian (Emsian/Eifelian) boundary (Burrow et al., 2008), has occasionally been placed within Selachimorpha, even within Hexanchiformes in the selachimorph crown-group (Burrow et al., 2008, and references therein); they very tentatively suggested a stem-selachimorph position. Boisvert et al. (2019) wondered instead if it is a stem-chondrichthyan.

The absence of any however tentative suggestions of crown-elasmobranchs before *Mcmurdodus* in the rather rich total-group chondrichthyan microfossil record despite the traditional optimism of paleodontologists may, somewhat ironically, serve as a hard maximum age for this calibration; the ICSC places the Emsian/Eifelian boundary at 393.3 ± 1.2 Ma ago, so I suggest 395 Ma.

Analysis Methods

Johan Renaudie (Museum für Naturkunde, Berlin) kindly performed the divergence dating using the tree (topology and uncalibrated branch lengths), the model of evolution (CAT-GTR+Γ) and clock model (lognormal autocorrelated relaxed) inferred by Irisarri et al. (2017), and the data ("nuclear test data set": the variable sites of the 14,352 most complete amino acid positions of their "NoDP" dataset), but the calibrations presented above and in the **Supplementary Material** (all at once, not different subsets).

The intent was to also use the software Irisarri et al. (2017) had used (PhyloBayes, though the latest version, 4.1c: Lartillot, 2015). However, PhyloBayes is unable to treat some bounds as hard and others as soft in the same analysis; it can only treat all as soft, as Irisarri et al. (2017) had done, or all as hard. Consequently, we ran our analysis with all bounds treated as hard in order to account for the hard minima (discussed above in the section "*Materials and methods*: *Hard and soft minima and maxima*").

The launch code for our PhyloBayes analysis is: ./pb -d ali14352.phy -T final_tree.tre -cal dm4.txt -r outgroups -bd -cat -gtr -ln -dc dm4hardDC.1 &./pb -d ali14352.phy -T final_tree.tre -cal dm4.txt -r outgroups -bd -cat -gtr -ln -dc dm4hardDC.2.

Irisarri et al. (2017) ran 100 gene-jackknifed analyses for each of their two sets of calibrations. Lacking the necessary computational resources, we only ran two analyses of the full dataset, without jackknifing. The results (**Table 2** and **Figure 1**) are therefore less reliable, given the data, than those of Irisarri et al. (2017), but they fully suffice as a proof of concept to show that improved calibrations lead to changes to many inferred node ages.

Above, I describe phylogenetic uncertainty leading to two different minimum ages for Tetrapoda (Node 105), 335 Ma and "roughly" 350 Ma. Using the younger age results in a younger bound of 359 Ma on the 95% credibility interval of this node (mean age: 363 Ma, older bound: 365 Ma, i.e., the maximum age of the calibration: **Table 2**); therefore, I do not consider it necessary to set the minimum age of this node to 350 Ma and run a second analysis.

Having evaluated (in the preceding section) the inherent uncertainty of each calibration before the analyses unlike Irisarri et al. (2017), I did not cross-validate the calibrations. In the words of Pardo et al. (2020), "*a priori* assessment of the quality of *a priori* node calibrations must retain logical primacy in assessing the quality of a molecular clock". "Reductio ad absurdum" cases aside (e.g., van Tuinen and Hedges, 2004, pp. 46–47; Waggoner and Collins, 2004; Matsui et al., 2008; Phillips et al., 2009; Ruane et al., 2010), apparent inconsistencies between calibrations should be seen as indicating not that the calibrations are wrong, but that the rate of evolution varies substantially across the tree, as already expected from other considerations (e.g., Berv and Field, 2017).

RESULTS AND DISCUSSION

Bibliometry

Irisarri et al. (2017: supplementary table 8) cited 15 works as sources for their calibrations, six of them compilations made by paleontologists to help molecular biologists calibrate timetrees.

Not counting Irisarri et al. (2017) and the ICSC (which has been updated at least once a year since 2008), I cite 238 references to discuss minimum ages (mostly for the age or phylogenetic position of a potentially calibrating specimen), 27 to discuss maximum ages (mostly to argue if observed absence of a clade is reliable), and 15 for both purposes. Of the total of 280, 1 each dates to 1964, 1978, 1981, 1988, and 1991; 2 each to 1994, 1995 and 1996; 1 each to 1997 and 1998; 3 to 1999; 1 to 2000; 2 to 2001; 4 to 2002; 2 to 2003; 0 to 2004; 7 to 2005; 4 to 2006; 6 each to 2007 and 2008; 5 to 2009; 5 to 2010; 8 to 2011; 9 to 2012; 15 to 2013; 12 to 2014; 23 to 2015; 24 to 2016; 23 to 2017; 28 to 2018; 50 to 2019; 28 to 2020; 1 to 2021; and 1 was published as an accepted manuscript in 2020 and is expected to come out this year in final form. (Whenever applicable, these are the years of actual publication, i.e., public availability of the layouted and proofread work, not the year of intended publication which can be a year earlier, and not the year of print which is very often one or even two years later.) Only three of these are among the 14 used by Irisarri et al. (2017), and none of them are among the six compilations they cited.

Irisarri et al. submitted their manuscript on September 16, 2016. Assuming that half of the publications cited here that were published in 2016 came out too late to be used by Irisarri et al. (2017), the total proportion of the works cited here that would have been useful to them for calibrating their timetree but were not available amounts to 142 of 280, or 50.7%. Similarly, 252 of the works cited here, or 90%, were published since mid-2005. I conclude from this extreme "pull of the recent" that knowledge in this area has an extremely short half-life; calibration dates, therefore, cannot be taken from published compilations (including the present work) or other secondary sources, but must be checked every time anew against the current primary literature. This is time-consuming even in the digital age, much more so than I expected, and requires reading more works for context than actually end up cited (for some nodes three times as many); but there is no shortcut.

Changes in the Calibration Dates

Of the 30 minimum ages assigned by Irisarri et al. (2017), I find only one to be accurate by the current state of knowledge, that of Batrachia (Node 160: **Supplementary Material**) anchored by good old *Triadobatrachus* (see Ascarrunz et al., 2016, for the latest and most thorough redescription and stratigraphy, and Daza et al., 2020, for the latest and largest phylogenetic analysis).

The minimum age of Pleurodira (Node 124: **Supplementary Material**), which has long been known to be 100 Ma older than Irisarri et al. (2017) set it, turns out to be copied from the calibration of a much smaller clade in Noonan and Chippindale (2006), a secondary source whose minimum age for Pleurodira was actually better by a factor of four. The minimum age of Iguanidae (Node 132: **Supplementary Material**) turned out to be miscopied, most likely with a typographic error, from Noonan and Chippindale (2006), who had it as 25 Ma instead of the 125 Ma of Irisarri et al. (2017)—though 25 Ma is not tenable either, but too young by at least 28 Ma.

In four more cases (Osteichthyes: Node 102 [**Supplementary Material**]; Reptilia: Node 107; Placentalia: Node 152; Lalagobatrachia/Bombinanura: Node 170 [**Supplementary Material**]), I find myself unable to assign any minimum age specific to that node. In two of these cases (Reptilia and Placentalia), the specimen previously thought to constrain that node actually constrains a less inclusive clade (Archelosauria, Node 108; Boreo(eu)theria, Node 153) that was sampled but not constrained by Irisarri et al. (2017); I have used these minimum ages to constrain the latter two nodes.

As might be expected, 15 of the minimum ages are too young, by margins ranging from 1.4 to 100 Ma or, ignoring Pleurodira, 43.25 Ma (**Table 1**: last two columns). Unsurprisingly, this also holds for the two nodes that Irisarri et al. (2017) did not calibrate but I did: both of them were constrained by calibrated nodes whose minimum ages were too young for these two nodes. In eight cases, including Boreo(eu)theria (Node 153), the reason is the expected one, the more or less recent discovery of previously unknown fossils (mostly before 2016); the magnitude of the resulting changes ranges from 1.4 to 11 Ma. In four more cases, including the one used by Irisarri et al. (2017) to date Osteichthyes (Node 102) but by me to date the subsequent split of Dipnomorpha and Tetrapodomorpha (Node 104: **Supplementary Material**), the dating of the oldest known specimens has improved by 4–16.5 Ma. The specimen used to

constrain Tetrapoda (Node 105) is probably not a tetrapod, but the oldest known certain tetrapods are now nonetheless dated as roughly 5 Ma older than the minimum assigned by Irisarri et al. (2017); depending on the phylogenetic hypothesis, isolated bones or (!) footprints roughly 20 Ma older that were published in 2015 could represent the oldest tetrapods instead. The remaining six cases, including Reptilia (Node 107) and Archelosauria (Node 108) by implication, are caused by phylogenetic reassignments of previously known specimens (mostly before 2016) and have effects ranging from 4 Ma to 43.25 Ma.

The minimum ages of the remaining 13 nodes (including, accidentally, Iguanidae) are too old; the margins vary from 1 to 96 Ma. This includes the case of Toxicofera (Node 129), whose minimum age of 148 Ma assigned by Irisarri et al. (2017) was not operational as that node was in fact constrained by the minimum age of its constituent clade Iguania (Node 131: **Supplementary Material**), 165 Ma; both of these ages are too old—I find minimum ages of 130 Ma for Toxicofera and 72 Ma for Iguania. Interestingly, none of the changes to minimum ages are due to more precise dating. There is one case of the opposite: I have changed the minimum age of Pipidae (Node 178: **Supplementary Material**) from 86 to 84 Ma because the oldest known safely identified pipid, *Pachycentrata*, may be somewhat older than the Coniacian/Santonian boundary (86.3 ± 0.5 Ma ago: ICSC), but also somewhat younger, so the Santonian/Campanian boundary (83.6 ± 0.2 Ma ago: ICSC) is a safer approximation. All others are due to more or less recent findings that the oldest supposed members of the clades in question cannot, or at least cannot be confidently, assigned to these clades.

I agree with the reasoning for one of the maximum ages used by Irisarri et al. (2017), that for Archosauria (Node 109: **Supplementary Material**), though its numeric value had to be increased by 1 Ma due to improved dating of the Permian/Triassic boundary since the source Irisarri et al. (2017) used was published in 2005.

I find myself unable to assign a separate maximum age to 7 of the 18 remaining nodes that Irisarri et al. (2017) used maximum ages for; these nodes are only constrained by the maximum ages of more inclusive clades in my reanalysis. This includes the case of Chondrichthyes (Node 187: **Supplementary Material**), whose maximum age of 495 Ma assigned by Irisarri et al. (2017) was not operational as that node was in fact constrained by the maximum age of the root node, 462.5 Ma; I can likewise constrain it only by the maximum age of the root, 475 Ma. In one of these cases, the new implied maximum age is younger (by 28.5 Ma) than the previously explicit maximum; in the remainder, it is older by 27–110 Ma.

Of the remaining 11 maximum ages, six were too young by 12.5–125 Ma. In one case (the root: Gnathostomata, Node 100), the old maximum is younger than the new minimum, and in two more cases (Mammalia, Node 150, and Theria, Node 151: both **Supplementary Material**), phylogenetic (or, in the case of Theria, possibly stratigraphic) uncertainty is the reason; the remaining three merely show greater caution on my part in interpreting absence of evidence as evidence of absence.

The remaining five I consider too old by 3.2–93 Ma; these show greater confidence on my part in interpreting absence of evidence as evidence of absence in well-sampled parts of the fossil record. The same holds, naturally, for the six nodes that lacked maximum ages in Irisarri et al. (2017) but that I propose maximum ages for; one of these new ages, however (for Lepidosauria, Node 125: **Supplementary Material**), is older than the previously implied maximum age provided by the next more inclusive clade, and that by 33 Ma. The other five are 60.1 Ma to no less than 261.5 Ma younger than their previously implied equivalents.

Changes in the Divergence Dates

Reanalyzing the data of Irisarri et al. (2017) with their methods, but using the calibration ages proposed and discussed above and treating them all as hard bounds in PhyloBayes instead of treating all as soft (see section "*Materials and methods*": "*Hard and soft minima and maxima*" and "*Analysis methods*"), generally leads to implausibly old ages and large credibility intervals for the unconstrained nodes (**Figure 1** and **Table 2**): e.g., the last common ancestor of chickens and turkeys (Node 115) is placed around the Cretaceous/Paleogene boundary, with a 95% credibility interval that spans half of each period, and the credibility interval of the bird crown group (Aves [PN], Node 112) spans most of the Jurassic, with a younger bound less than 10 Ma younger than the age of the distant stem-avialan *Archaeopteryx* (just over 150 Ma), while the oldest known crown-birds are less than half as old, about 71 Ma (see section "Materials and Methods": Calibrations: Node 113).

There are exceptions, however. Most notably, the squamate radiation (nodes 126–129) is constrained only between the origin of Lepidosauria (**Supplementary Material**: Node 125: 244–290 Ma ago) and the origin of Toxicofera (Materials and methods: Calibrations: Node 129: minimum age 130 Ma), yet it is bunched up close to the latter date, unlike in Irisarri et al. (2017) where it was more spread out and generally older even though both calibrations were younger. For example, the unconstrained origin of Squamata [PN] (Node 126) was found to have a mean age of 199 Ma by Irisarri et al. (2017), but 153 Ma here (**Table 2**). The crucial difference may be that Lepidosauria did not have a maximum age, but this does not explain the very short internodes from Squamata to Iguania in my results. I should point out that the oldest likely squamate remains are close to 170 Ma old (reviewed in Panciroli et al., 2020).

In part, these implausible ages may be due to effects of body size (Berv and Field, 2017) or loosely related factors like generation length: most sampled squamates are small, while the two sampled palaeognath birds (Node 116, with an evidently spurious mean age of 163 Ma) are much larger than all sampled neognaths. This may be supported by the body size increase in snakes: their oldest sampled node (Macrostomata or Afrophidia: Node 136) is placed around the Early/Late Cretaceous boundary, followed by the origin of Endoglyptodonta (Node 138) in the Late Cretaceous, while any Late Cretaceous caenophidians (a clade containing Endoglyptodonta) remain unknown, all potential Cretaceous total-group macrostomates are beset with phylogenetic uncertainty, and considerably younger dates were found by Burbrink et al. (2020) despite the use of a mid-Cretaceous potential macrostomate as a minimum-age-only

calibration. Similarly, the fact that the entire credibility interval for Supraprimates/Euarchontoglires (Node 155) was younger than its calibrated minimum age when all bounds were treated as soft in Marjanović (2019) may be due to the fact that one of the two sampled supraprimates is *Homo*, the second-largest sampled mammal and the one with the second-longest generation span.

Whelan and Halanych (2016) found that the CAT-GTR model (at least as implemented in PhyloBayes) is prone to inferring inaccurate branch lengths, especially in large datasets; this may turn out to be another cause of the results described above. The omission of the constant characters from the dataset, intended to speed up calculations (Irisarri et al., 2017), may have exacerbated this problem by leading to inaccurate model parameters (Whelan and Halanych, 2016).

It is, however, noteworthy that all terminal branches inferred here are longer, in terms of time, than in Irisarri et al. (2017).

Naturally, the changes to the calibration dates have changed the inferred ages of many calibrated nodes and the sizes of their credibility intervals. For instance, Irisarri et al. (2017) inferred a mean age of 207 Ma for Batoidea, with a 90-Ma-long 95% credibility interval that stretched from 172 Ma ago to 262 Ma ago (Node 192; **Table 2**); that node was calibrated with a soft minimum age set to 176 Ma, but not only was no maximum age set, no other node between there and the root node (Gnathostomata, Node 100) had a maximum age either, so that effectively the maximum age for Batoidea was that of the root node, 462.5 Ma. Following the discovery of new fossils, I have increased the hard minimum age to 184 Ma; however, out of ecological considerations, I have also introduced a hard maximum age of 201 Ma, younger than the previously inferred mean age. Naturally, the new inferred mean age is also younger: 193 Ma, with a 95% credibility interval that spans the time between the calibration dates (**Table 2**).

Somewhat similarly, I have increased the minimum age of Mammalia (**Supplementary Material**: Node 150) from 162.5 to 179 Ma following improved dating of the oldest certain mammals, increased its maximum age from 191.4 to 233 Ma to account for phylogenetic uncertainty and the limits of the Norian (middle Late Triassic) fossil record, and treated both bounds as hard. While Irisarri et al. (2017) found a mean age of 165 Ma with a credibility interval from 161 to 172 Ma, straddling the minimum age but not reaching the maximum, I find an age range that reaches the new maximum but stays far away from the new minimum (mean: 229 Ma, 95% credibility interval from 217 to 233 Ma). While the next less inclusive calibrated node (151: Theria; **Supplementary Material**) has an increased maximum but a barely changed minimum age, both bounds of the next more inclusive calibrated node (106: Amniota) have increased by about 30 Ma, apparently pulling the inferred age of Mammalia with them.

Pitfalls in Interpreting the Descriptive Paleontological Literature

It is widely thought that paleontologists are particularly eager to publish their specimens as the oldest known record of some taxon. Indeed, it happens that five different species of different ages are published as the oldest record of the same taxon within 10 years. In such cases, finding a specimen that can establish a minimum age for that taxon can be as simple as finding the latest publication that makes such a claim, and that can be as simple as a Google Scholar search restricted to the last few years. However, there are harder cases; I will present two.

In the **Supplementary Material**, I argue for using the age of *Kopidosaurus*, about 53 million years, as the minimum age of Iguanidae (Node 132). *Kopidosaurus* was named and described from a largely complete skull by Scarpetta (2020a) in a publication where the words "oldest" and "older" do not occur at all, and "first" and "ancient" only occur in other contexts—even though Scarpetta (2019) had just published on calibration dates for molecular divergence date analyses. The reason is (S. Scarpetta, personal communication 2021) that he did not think *Kopidosaurus* was the oldest iguanid; one of the two matrices he used for phylogenetic analyses contained the 56-Ma-old *Suzanniwana*, and his analyses found it as an iguanid (Scarpetta, 2020a: supplementary information; Scarpetta, 2020b). Moreover, he was aware that the publication that named and described *Suzanniwana* (Smith, 2009a) also named and described *Anolbanolis* from the same site and age and argued that both of them—known from large numbers of isolated skull bones—were iguanids. Yet, *Anolbanolis* has never, to the best of my knowledge, been included in any phylogenetic analysis; and Conrad (2015), not mentioning *Anolbanolis* and not cited by Scarpetta (2020a,b), had found the phylogenetic position of *Suzanniwana* difficult to resolve in the analysis of a dataset that included a much larger sample of early pan-iguanians.

Smith (2009a, pp. 312–313), incidentally, did not advertise *Suzanniwana* and *Anolbanolis* as the oldest iguanids either, accepting instead at least some of the even older jaw fragments that had been described as iguanid as "surely iguanid", explicitly so for the "highly streamworn" over-62-Ma-old *Swainiguanoides* which had been described as "the oldest North American iguanid" (Sullivan, 1982). All of that and more was considered too uncertain by DeMar et al. (2017, p. 4, file S1: 26–28), who pointed out not only how fragmentary that material was (and that some of the Cretaceous specimens more likely belong to certain other squamate clades), but also that the presence of exclusive synapomorphies with Iguanidae (if confirmed) does not mean the specimens are actually inside that crown clade—they could be on its stem. As the "oldest definitive" iguanids, Dashzeveg et al. (2005: 4) accepted *Anolbanolis*, followed by the uncontroversial *Afairiguana*, which is younger than *Kopidosaurus*; curiously, they did not mention *Suzanniwana* at all.

The conclusion that the status of *Suzanniwana* and *Anolbanolis* (let alone *Swainiguanoides* and the like) is too uncertain and that *Kopidosaurus*, nowhere advertised for that purpose, should be used to set the minimum age for Node 132 was accessible to me as an outsider to the fossil record of iguanians (or indeed squamates in general), but it took me several days of searching and reading papers and their supplementary information, and I was lucky that the two papers I overlooked (pointed out by Scarpetta, personal communication 2021) do not change this conclusion.

It took me much less effort to find that, under some phylogenetic hypotheses, the oldest known tetrapod (Materials and methods: Calibrations: Node 105 – Tetrapoda) is *Casineria*, a specimen I have studied in person and published on (Marjanović and Laurin, 2019); yet, the idea had never occurred to me or apparently anyone else in the field, even though its possibility should have been evident since 2017 and even though the phylogenetic hypotheses in question are by no means outlandish—one of them is even majoritarian.

In short, the paleontological literature is not optimized for divergence dating; the questions of what is the oldest known member of a group or when exactly that group evolved often take a back seat to understanding the anatomy, biomechanics, ecology, extinction, phylogeny, or generally speaking evolution of that group in the minds of paleontologists—paleobiologists—and this is reflected in the literature. Mining it for bounds on divergence dates is still possible, as I hope to have shown, but also rather exhausting.

SUMMARY AND CONCLUSION

Irisarri et al. (2017) published the largest vertebrate timetree to date, calibrated with 30 minimum and 19 maximum ages for selected nodes (although one of each was not operational because the calibrations of other nodes set tighter constraints). With just 3 years of hindsight, only one of these dates stands up to scrutiny. Of the remaining 29 minimum ages, two had to be removed altogether, two had to be moved to previously uncalibrated nodes (with modifications to their numeric values), 15 were 4–100 Ma too young, and 13 were 1–96 Ma too old. Of the 19 maximum ages, seven had to be canceled altogether, while six were too young by 13–125 Ma, and five were too old by 3–93 Ma.

One of the minimum ages was taken from the wrong node in the cited secondary source, an earlier divergence-date analysis of molecular data (Noonan and Chippindale, 2006); another from the same source had a hundred million years added without explanation, most likely by typographic error. Only six of the 30 calibrated nodes were calibrated from primary literature. The calibration dates for seven nodes were taken from the compilation by Benton and Donoghue (2007), several from other compendia, and four from Noonan and Chippindale (2006) who had not succeeded in presenting the contemporary state of knowledge either.

Using software that was only able to treat all bounds as hard or all as soft (meaning that 2.5 or 5% of the credibility interval of each inferred node age must extend beyond the bound—younger than the minimum and older than the maximum age, where present), Irisarri et al. (2017) opted to treat all bounds as soft. For all minimum ages except one, this decision is not reproducible; it is even arguable for some of the maxima. This is not a purely theoretical problem; even the inferred mean ages of some calibrated nodes were younger than their minima in Marjanović (2019).

Redating of the tree of Irisarri et al. (2017) with the presumably improved calibrations results in many changes to the mean ages of nodes and to the sizes of their credibility intervals; not all of these changes are easily predictable.

Of the 280 references I have used to improve the calibrations, 50 were published in 2019, half of the total were published after mid-2016 [when Irisarri et al. (2017) seem to have completed the work on their manuscript], and 90% were published after mid-2005. Paleontology is a fast-moving field; secondary sources cannot keep up with the half-life of knowledge. A continually updated online compendium of calibration dates would be very useful, but the only attempt to create one (Ksepka et al., 2015) is no longer funded, has not been updated since early 2018, and had limited coverage. For the time being, each new attempt to calibrate node or tip ages will have to involve finding and studying the recent paleontological and chronostratigraphic literature on the taxa, strata, and sites in question; although the Internet has made this orders of magnitude easier, it remains labor-intensive, in part because the oldest record of a clade is often not published as such, but has to be inferred from comparing several sources on phylogeny, chronostratigraphy, and sometimes taphonomy or even phylogenetics, as I illustrate here.

I urge that such work be undertaken and sufficiently funded. Accurate and precise timetrees remain an essential component of our understanding of, for example, the model organisms that are used in biomedical research: how much they can tell us about ourselves depends on how much evolution has happened along both branches since our last common ancestor, and that is in part a function of time.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

DM designed the experiments, gathered the data, interpreted the results, prepared the figure and the tables, and wrote the manuscript.

FUNDING

I received no funding for this work; indeed I had to interrupt it for a long time for this reason.

ACKNOWLEDGMENTS

Glory to our pirate queen, without whose work this manuscript would at best have taken a lot longer to write and at worst would have been severely outdated before submission. Thanks to Albert Chen, Matteo Belvedere, and Jason Pardo for electronic reprints that would likely have been impossible to acquire in a timely manner otherwise; to Johan Renaudie for making me aware of another; to Olga Karicheva and the editorial office

for several deadline extensions; to Jason Silviria and Paige dePolo for discussion of early eutherians; to the editor, DB, for finding five reviewers; to all five reviewers and the editor for helpful comments; and to the editor and the editorial office for several more deadline extensions. PhyloBayes only runs on Unix systems; Johan Renaudie (Museum für Naturkunde, Berlin) has access to such and kindly performed the time-consuming analyses after expertly overcoming the gaps in the documentation of PhyloBayes. The first submission of this paper was released as a preprint at *bioRXiv* 2019.12.19.882829v1 (Marjanović, 2019). I thank Adam Yates for finding a misused term in that preprint (now corrected), Ben King for alerting me to his important paper (King, 2019), and Tiago Simões for a partial but detailed review of the preprint. The second submission was released as a preprint at *bioRXiv* 2019.12.19.882829v2. I thank Simon Scarpetta for his comments on potential pan-iguanians and their phylogeny and for alerting me to the publications by Smith (2009b) and Scarpetta (2020b). The third and the fourth submissions were also released as preprints at bioRXiv 2019.12.19.882829v3 and v4.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.521693/full#supplementary-material

REFERENCES

Agnolín, F. L., Brissón Egli, F., Chatterjee, S., Garcia Marsà, J. A., and Novas, F. E. (2017). Vegaviidae, a new clade of southern diving birds that survived the K/T boundary. *Sci. Nat.* 104:87. doi: 10.1007/s00114-017-1508-y

Ahlberg, P. E., and Clack, J. A. (2020). The smallest known Devonian tetrapod shows unexpectedly derived features. *R. Soc. Open Sci.* 7:192117. doi: 10.1098/rsos.192117

Alifanov, V. R. (2013). *Desertiguana gobiensis* gen. et sp. nov., a new lizard (Phrynosomatidae, Iguanomorpha) from the Upper Cretaceous of mongolia. *Paleont. J.* 47, 417–424. doi: 10.1134/S0031030113040023

Alifanov, V. R. (2019). Lizards of the families Dorsetisauridae and Xenosauridae (Anguimorpha) from the Aptian–Albian of Mongolia. *Paleont. J.* 53, 183–193. doi: 10.1134/S0031030119020023

Alifanov, V. R. (2020). A new lizard (Agamidae, Iguania) from the late Paleocene of south Mongolia. *Paleont. J.* 54, 410–413. doi: 10.1134/S0031030120040036

Anderson, J. S., Smithson, T. R., Mansky, C. F., Meyer, T., and Clack, J. (2015). A diverse tetrapod fauna at the base of 'Romer's Gap'. *PLOS ONE* 10:e0125446. doi: 10.1371/journal.pone.0125446

Andreev, P. S., and Cuny, G. (2012). New Triassic stem selachimorphs (Chondrichthyes, Elasmobranchii) and their bearing on the evolution of dental enameloid in Neoselachii. *J. Vert. Paleont.* 32, 255–266. doi: 10.1080/02724634.2012.644646

Andreev, P. [S]., Coates, M. I., Karatajūtė-Talimaa, V., Shelton, R. M., Cooper, P. R., Wang, N.-Z., et al. (2016a). The systematics of the Mongolepidida (Chondrichthyes) and the Ordovician origins of the clade. *PeerJ* 4:e1850. doi: 10.7717/peerj.1850

Andreev, P. S., Coates, M. I., Karatajūtė-Talimaa, V., Shelton, R. M., Cooper, P. R., and Sansom, I. J. (2016b). *Elegestolepis* and its kin, the earliest monodontode chondrichthyans. *J. Vert. Paleont.* 37:e1245664. doi: 10.1080/02724634.2017.1245664

Andreev, P. S., Coates, M. I., Shelton, R. M., Cooper, P. R., Smith, M. P., and Sansom, I. J. (2015). Upper ordovician chondrichthyan-like scales from North America. *Palaeontology* 58, 691–704. doi: 10.1111/pala.12167

Anquetin, J., and André, C. (2020). The last surviving thalassochelydia—a new turtle cranium from the Early Cretaceous. *PaleorXiv: 7pa5c version 3, peer-reviewed by PCI Paleo*. doi: 10.31233/osf.io/7pa5c

Apesteguía, S., Daza, J.-D., Simões, T. R., and Rage, J. C. (2016). The first iguanian lizard from the Mesozoic of Africa. *R. Soc. Open Sci.* 3:160462. doi: 10.1098/rsos.160462

Aranciaga Rolando, A. M., Agnolin, F. L., and Corsolini, J. (2019). A new pipoid frog (Anura, Pipimorpha) from the Paleogene of Patagonia. Paleobiogeographical implications. *C. R. Palevol* 18, 725–734. doi: 10.1016/j.crpv.2019.04.003

Arribas, I., Buscalioni, A. D., Royo Torres, R., Espílez, E., Mampel, L., and Alcalá, L. (2019). A new goniopholidid crocodyliform, *Hulkepholis rori* sp. nov. from the Camarillas Formation (early Barremian) in Galve, Spain) [sic]. *PeerJ* 7:e7911. doi: 10.7717/peerj.7911

Ascarrunz, E., Rage, J.-C., Legreneur, P., and Laurin, M. (2016). *Triadobatrachus massinoti*, the earliest known lissamphibian (Vertebrata: Tetrapoda) re-examined by μCT scan, and the evolution of trunk length in batrachians. *Contr. Zool.* 85, 201–234. doi: 10.1163/18759866-08502004

Asher, R. J., Meng, J., Wible, J. R., McKenna, M. C., Rougier, G. W., Dashzeveg, D., et al. (2005). Stem Lagomorpha and the antiquity of Glires. *Science* 307, 1091–1094. doi: 10.1126/science.1107808

Asher, R. J., Smith, M. R., Rankin, A., and Emry, R. J. (2019). Congruence, fossils and the evolutionary tree of rodents and lagomorphs. *R. Soc. Open Sci.* 6:190387. doi: 10.1098/rsos.190387

Averianov, A. O., Archibald, J. D., and Martin, T. (2003). Placental nature of the alleged marsupial from the Cretaceous of Madagascar. *Acta Palaeont. Pol.* 48, 149–151.

Ax, P. (1987). *The Phylogenetic System: The Systematization of Organisms on the Basis of Their Phylogenesis*. New York, NY: John Wiley & Sons.

Báez, A. M. (2013). Anurans from the Early Cretaceous Lagerstätte of Las Hoyas, Spain: New Evidence on the Mesozoic diversification of crown-clade Anura. *Cret. Res.* 41, 90–106. doi: 10.1016/j.cretres.2012.11.002

Báez, A. M., and Gómez, R. O. (2016). Revision of the skeletal morphology of *Eodiscoglossus santonjae*, an Early Cretaceous frog from northeastern Spain, with comments on its phylogenetic placement. *Foss. Imprint* 72, 67–77. doi: 10.14446/FI.2016.67

Báez, A. M., and Gómez, R. O. (2019). Redescription of the overlooked basal frog *Wealdenbatrachus* reveals increased diversity among Early Cretaceous anurans. *Cret. Res.* 99, 14–29. doi: 10.1016/j.cretres.2019.02.006

Bailleul, A. M., O'Connor, J. K., Zhang, S., Li, Z., Wang, Q., Lamanna, M. C., et al. (2019). An early Cretaceous enantiornithine (Aves) preserving an unlaid egg and probable medullary bone. *Nat. Commun.* 10:1275. doi: 10.1038/s41467-019-09259-x

Barido-Sottani, J., Aguirre-Fernández, G., Hopkins, M. J., Stadler, T., and Warnock, R. (2019). Ignoring stratigraphic age uncertainty leads to erroneous estimates of species divergence times under the fossilized birth–death process. *Proc. R. Soc. B* 286:0685. doi: 10.1098/rspb.2019.0685

Barido-Sottani, J., van Tiel, N. M. A., Hopkins, M. J., Wright, D. F., Stadler, T., and Warnock, R. C. M. (2020). Ignoring fossil age uncertainty leads to inaccurate topology and divergence time estimates in time[-]calibrated tree inference. *Front. Genet.* 8:183. doi: 10.3389/fevo.2020.00183

Benton, M. J., and Donoghue, P. C. J. (2007). [Correction with unabridged table 1 to:] Palaeontological evidence to date the tree of life. *Mol. Biol. Evol.* 24, 889–891. doi: 10.1093/molbev/msm017

Benton, M. J., Donoghue, P. C. J., Asher, R. J., Friedman, M., Near, T. J., and Vinther, J. (2015). Constraints on the timescale of animal evolutionary history. *Palaeont. Electron.* 18, 1–116. doi: 10.26879/424

Berv, J. S., and Field, D. J. (2017). Genomic signature of an avian Lilliput effect across the K-Pg extinction. *Syst. Biol.* 67, 1–13. doi: 10.1093/sysbio/syx064

Bi, S., Wible, J. R., Zheng, X., and Wang, X. (2019). "The Early Cretaceous eutherian Ambolestes and its implications for the eutherian-metatherian dichotomy," in *Proceedings of the 79th Annual Meeting: Society of Vertebrate Paleontology,* Vol. 64, ed. J. V. Paleont. Available online at: http://vertpaleo.org/Annual-Meeting/Annual-Meeting-Home/SVP-Program-book-v5_w-covers.aspx

Bi, S., Zheng, X., Wang, X., Cignetti, N. E., Yang, S., and Wible, J. R. (2018). An Early Cretaceous eutherian and the placental–marsupial dichotomy. *Nature* 558, 390–395. doi: 10.1038/s41586-018-0210-3

Blackburn, D. C., Paluh, D. J., Krone, I., Roberts, E. M., Stanley, E. L., and Stevens, N. J. (2019). The earliest fossil of the African clawed frog (genus *Xenopus*) from sub-Saharan Africa. *J. Herpetol.* 53, 125–130. doi: 10.1670/18-139

Boisvert, C. A., Johnston, P., Trinajstic, K., and Johanson, Z. (2019). "Chondrichthyan evolution, diversity, and senses," in *Heads, Jaws, and Muscles. Anatomical, Functional, and Developmental Diversity in Chordate Evolution*, eds J. Ziermann, R. Diaz Jr., and R. Diogo (Cham: Springer), 65–91. doi: 10.1007/978-3-319-93560-7_4

Bona, P., Ezcurra, M. D., Barrios, F., and Fernandez Blanco, M. V. (2018). A new Palaeocene crocodylian from southern Argentina sheds light on the early history of caimanines. *Proc. R. Soc. B* 285, 20180843. doi: 10.1098/rspb.2018.0843

Borsuk-Białynicka, M. (1996). The Late Cretaceous lizard *Pleurodontagama* and the origin of tooth permanency in Lepidosauria. *Acta Palaeont. Pol.* 41, 231–252.

Botella, H., Blom, H., Dorka, M., Ahlberg, P. E., and Janvier, P. (2007). Jaws and teeth of the earliest bony fishes. *Nature* 448, 583–586. doi: 10.1038/nature05989

Burbrink, F. T., Grazziotin, F. G., Pyron, R. A., Cundall, D., Donnellan, S., Irish, F., et al. (2020). Interrogating genomic-scale data for Squamata (lizards, snakes, and amphisbaenians) shows no support for key traditional morphological relationships. *Syst. Biol.* 69, 502–520. doi: 10.1093/sysbio/syz062

Burrow, C. J., and Young, G. C. (1999). An articulated teleostome fish from the Late Silurian (Ludlow) of Victoria, Australia. *Rec. West. Austral. Mus.* Suppl.57, 1–14.

Burrow, C. J., Hovestadt, D. C., Hovestadt-Euler, M., Turner, S., and Young, G. C. (2008). New information on the Devonian shark *Mcmurdodus*, based on material from western Queensland, Australia. *Acta Geol. Pol.* 58, 155–163.

Butler, R. J., Brusatte, S. L., Reich, M., Nesbitt, S. J., Schoch, R. R., and Hornung, J. J. (2011). The sail-backed reptile *Ctenosauriscus* from the latest Early Triassic of Germany and the timing and biogeography of the early archosaur radiation. *PlOS ONE* 6:e25693. doi: 10.1371/journal.pone.0025693

Cadena, E. (2015). A global phylogeny of Pelomedusoides turtles with new material of *Neochelys franzeni* Schleich, 1993 (Testudines, Podocnemididae) from the middle Eocene, Messel Pit, of Germany. *PeerJ* 3:e1221. doi: 10.7717/peerj.1221

Caldwell, M. W., Nydam, R. L., Palci, A., and Apesteguía, S. (2015). The oldest known snakes from the Middle Jurassic-Lower [sic] Cretaceous provide insights on snake evolution. *Nat. Commun.* 6:5996. doi: 10.1038/ncomms6996

Campbell Mekarski, M., Pierce, S. E., and Caldwell, M. W. (2019). Spatiotemporal distributions of non-ophidian ophidiomorphs, with implications for their origin, radiation, and extinction. *Front. Earth Sci.* 7:245. doi: 10.3389/feart.2019.00245

Cannatella, D. (2015). *Xenopus* in space and time: fossils, node calibrations, tip-dating, and paleobiogeography. *Cytogenet. Genome Res.* 145, 283–301. doi: 10.1159/000438910

Cantino, P. D., and de Queiroz, K. (2020). *International Code of Phylogenetic Nomenclature (PhyloCode). Version 6*. New York, NY: CRC Press. Available online at: http://phylonames.org/code

Carneiro, L. M. (2017). A new species of *Varalphadon* (Mammalia, Metatheria, Sparassodonta) from the upper Cenomanian of southern Utah, North America: Phylogenetic and biogeographic insights. *Cret. Res.* 84, 88–96. doi: 10.1016/j.cretres.2017.11.004

Carneiro, L. M., and Oliveira, É. V. (2017). The Eocene South American metatherian *Zeusdelphys complicatus* is not a protodidelphid but a hatcheriform: Paleobiogeographic implications. *Acta Palaeont. Pol.* 62, 497–507. doi: 10.4202/app.00351.2017

Carpenter, D. K., Falcon-Lang, H. J., Benton, M. J., and Grey, M. (2015). Early Pennsylvanian (Langsettian) fish assemblages from the Joggins Formation, Canada, and their implications for palaeoecology and palaeogeography. *Palaeontology* 58, 661–690. doi: 10.1111/pala.12164

Carroll, R. L. (1964). The earliest reptiles. *Zool. J. Linn. Soc.* 45, 61–83. doi: 10.1111/j.1096-3642.1964.tb00488.x

Carroll, R. L., and Chorn, J. (1995). Vertebral development in the oldest microsaur and the problem of "lepospondyl" relationships. *J. Vert. Paleont.* 15, 37–56. doi: 10.1080/02724634.1995.10011206

Carroll, R. L., Bybee, P., and Tidwell, W. D. (1991). The oldest microsaur (Amphibia). *J. Paleont.* 65, 314–322. doi: 10.1017/s0022336000020552

Celik, M. A., and Phillips, M. J. (2020). Conflict resolution for Mesozoic mammals: reconciling phylogenetic incongruence among anatomical regions. *Front. Genet.* 11:0651. doi: 10.3389/fgene.2020.00651

Cellinese, N., and Dell, C. (2020). *RegNum – The International Clade Names Repository [Website]*. https://www.phyloregnum.org/ (accessed April 26, 2021).

Chen, D., Blom, H., Sanchez, S., Tafforeau, P., and Ahlberg, P. E. (2016). The stem osteichthyan *Andreolepis* and the origin of tooth replacement. *Nature* 539, 237–241. doi: 10.1038/nature19812

Chester, S. G. B., Bloch, J. I., Boyer, D. M., and Clemens, W. A. (2015). Oldest known euarchontan tarsals and affinities of Paleocene *Purgatorius* to Primates. *Proc. Natl. Acad. Sci. U.S.A.* 112, 1487–1492. doi: 10.1073/pnas.1421707112

Chinzorig, T., Kobayashi, Y., Tsogtbaatar, K., Currie, P. J., Watabe, M., and Barsbold, R. (2017). First ornithomimid (Theropoda, Ornithomimosauria) from the Upper Cretaceous Djadokhta Formation of Tögrögiin Shiree, Mongolia. *Sci. Rep.* 7:5835. doi: 10.1038/s41598-017-05272-6

Choo, B., Zhu, M., Qu, Q., Yu, X., Jia, L., and Zhao, W. (2017). A new osteichthyan from the late Silurian of Yunnan, China. *PLOS ONE* 12:e0170929. doi: 10.1371/journal.pone.0170929

Clack, J. A. (2011). A new microsaur from the Early Carboniferous (Viséan) of East Kirkton, Scotland, showing soft-tissue evidence. *Spec. Pap. Palaeont.* 29, 45–55. doi: 10.1111/j.1475-4983.2011.01073.x

Clack, J. A., Bennett, C. E., Carpenter, D. K., Davies, S. J., Fraser, N. C., Kearsey, T. I., et al. (2016). Phylogenetic and environmental context of a Tournaisian tetrapod fauna. *Nat. Ecol. Evol.* 1:0002. doi: 10.1038/s41559-016-0002

Clack, J. A., Ruta, M., Milner, A. R., Marshall, J. E. A., Smithson, T. R., and Smithson, K. Z. (2019). *Acherontiscus caledoniae*: the earliest heterodont and durophagous tetrapod. *R. Soc. Open Sci.* 6:182087. doi: 10.1098/rsos.182087

Clarke, J. A., Tambussi, C. P., Noriega, J. I., Erickson, G. M., and Ketcham, R. A. (2005). Definitive fossil evidence for the extant avian radiation in the Cretaceous. *Nature* 433, 305–308. doi: 10.1038/nature03150

Clement, A. M., King, B., Giles, S., Choo, B., Ahlberg, P. E., Young, G. C., et al. (2018). Neurocranial anatomy of an enigmatic Early Devonian fish sheds light on early osteichthyan evolution. *eLife* 7:e34349. doi: 10.7554/eLife.34349

Clyde, W. C., Ramezani, J., Johnson, K. R., Bowring, S. A., and Jones, M. M. (2016). Direct high-precision U–Pb geochronology of the end-Cretaceous extinction and calibration of Paleocene astronomical timescales. *Earth Planet. Sci. Lett.* 452, 272–280. doi: 10.1016/j.epsl.2016.07.041

Cohen, J. E. (2017). *Radiation of Tribosphenic Mammals During the Early Late Cretaceous (Turonian) of North America*. Doctoral thesis. Norman: The University of Oklahoma.

Cohen, J. E., and Cifelli, R. L. (2015). "The first eutherian mammals from the early Late Cretaceous of North America," in *Proceedings of the 75th Annual Meeting:Society of Vertebrate Paleontology*, Vol. 108, ed. J. V. Paleont. Available online at: http://vertpaleo.org/PDFS/SVP-2015-Program-and-Abstract-Book-9-22-2015.aspx

Cohen, J. E., Davis, B. M., and Cifelli, R. L. (2020). Geologically oldest Pediomyoidea (Mammalia, Marsupialiformes) from the Late Cretaceous of North America, with implications for taxonomy and diet of earliest Late Cretaceous mammals. *J. Vert. Paleont.* 40:e1835935. doi: 10.1080/02724634.2020.1835935

Cohen, K. M., Harper, D. A. T., Gibbard, P. L., and Fan, J.-X. (2020). *International Chronostratigraphic Chart v2020/03 [Chart]*. International Commission on Stratigraphy. Available online at: http://www.stratigraphy.org/ICSchart/ChronostratChart2020-03.jpgor.pdf (accessed April 26, 2021).

Conrad, J. L. (2015). A new Eocene casquehead lizard (Reptilia, Corytophanidae) from North America. *PLOS ONE* 10:e0127900. doi: 10.1371/journal.pone.0127900

Conrad, J. L. (2017). A new lizard (Squamata) was the last meal of *Compsognathus* (Theropoda: Dinosauria) and is a holotype in a holotype. *Zool. J. Linn. Soc.* 183, 584–634. doi: 10.1093/zoolinnean/zlx055

Conrad, J. L., and Norell, M. A. (2007). A complete Late Cretaceous iguanian (Squamata, Reptilia) from the Gobi and identification of a new iguanian clade. *Am. Mus. Novit.* 3584, 1–41. doi: 10.1206/0003-0082(2007)3584[1:ACLCIS]2.0.CO;2

Cossette, A. P., and Brochu, C. A. (2018). A new specimen of the alligatoroid *Bottosaurus harlani* and the early history of character evolution in alligatorids. *J. Vert. Paleont.* 38:e1486321. doi: 10.1080/02726434.2018.1486321

Cossette, A. P., and Brochu, C. A. (2020). A systematic review of the giant alligatoroid *Deinosuchus* from the Campanian of North America and its implications for the relationships at the root of Crocodylia. *J. Vert. Paleont.* 40:e1767638. doi: 10.1080/02724634.2020.1767638

Cúneo, R., Ramezani, J., Scasso, R., Pol, D., Escapa, I., Zavattieri, A. M., et al. (2013). High-precision U–Pb geochronology and a new chronostratigraphy for the Cañadón Asfalto Basin, Chubut, central Patagonia: Implications for terrestrial faunal and floral evolution in Jurassic [sic]. *Gondw. Res.* 24, 1267–1275. doi: 10.1016/j.gr.2013.01.010

Cuny, G., Guinot, G., and Enault, S. (2017). *Evolution of Dental Tissues and Paleobiology in Selachians*. London: Elsevier. doi: 10.1016/C2015-0-06183-4

Danto, M., Witzmann, F., Kamenz, S. K., and Fröbisch, N. B. (2019). How informative is vertebral development for the origin of lissamphibians? *J. Zool.* 307, 292–305. doi: 10.1111/jzo.12648

Dashzeveg, D., Dingus, L., Loope, D. B., Swisher, C. C. III., Dulam, T., and Sweeney, M. R. (2005). New stratigraphic subdivision, depositional environment, and age estimate for the Upper Cretaceous Djadokhta Formation, southern Ulan Nur Basin, Mongolia. *Am. Mus. Novit.* 3498, 1–31. doi: 10.1206/0003-0082(2005)498[0001:NSSDEA]2.0.CO;2

Davies, T. M., Bell, M. A., Goswami, A., and Halliday, T. J. D. (2017). Completeness of the eutherian mammal fossil record and implications for reconstructing mammal evolution through the Cretaceous/Paleogene mass extinction. *Paleobiology* 43, 521–536. doi: 10.1017/pab.2017.20

Daza, J. D., Stanley, E. L., Bolet, A., Bauer, A. M., Arias, J. S., Čerňanský, A., et al. (2020). Enigmatic amphibians in mid-Cretaceous amber were chameleon-like ballistic feeders. *Science* 370, 687–691. doi: 10.1126/science.abb6005

Daza, J. D., Stanley, E. L., Wagner, P., Bauer, A. M., and Grimaldi, D. A. (2016). Mid-Cretaceous amber fossils illuminate the past diversity of tropical lizards. *Sci. Adv.* 2:e1501080. doi: 10.1126/sciadv.1501080

de la Fuente, M. S., and Iturralde-Vinent, M. (2001). A new pleurodiran turtle from the Jagua Formation (Oxfordian) of western Cuba. *J. Paleont.* 75, 860–869. doi: 10.1666/0022-3360(2001)075<0860:ANPTFT>2.0.CO;2

De Pietri, V. L., Scofield, R. P., Zelenkov, N., Boles, W. E., and Worthy, T. H. (2016). The unexpected survival of an ancient lineage of anseriform birds into the Neogene of Australia: the youngest record of Presbyornithidae. *R. Soc. Open Sci.* 3:150635. doi: 10.1098/rsos.150635

de Queiroz, K., Cantino, P. D., and Gauthier, J. A. (eds) (2020). *Phylonyms – A Companion to the PhyloCode*. New York, NY: CRC Press.

de Souza Carvalho, I., Agnolin, F., Aranciaga Rolando, M. A., Novas, F. E., Xavier-Neto, J., de Freitas, F. I., et al. (2019). A new genus of pipimorph frog (Anura) from the Early Cretaceous Crato Formation (Aptian) and the evolution of South American tongueless frogs. *J. S. Am. Earth Sci.* 92, 222–233. doi: 10.1016/j.jsames.2019.03.005

DeMar, D. G. Jr., Conrad, J. L., Head, J. J., Varricchio, D. J., and Wilson, G. P. (2017). A new Late Cretaceous iguanomorph from North America and the origin of New World Pleurodonta (Squamata, Iguania). *Proc. R. Soc. B* 284:20161902. doi: 10.1098/rspb.2016.1902

Dingus, L., Loope, D. B., Dashzeveg, D., Swisher, C. C. III., Minjin, C., Novacek, M. J., et al. (2008). The geology of Ukhaa Tolgod (Djadokhta Formation, Upper Cretaceous, Nemegt Basin, Mongolia). *Am. Mus. Novit.* 3616, 1–40. doi: 10.1206/442.1

Eldridge, M. D. B., Beck, R. M. D., Croft, D. A., Travouillon, K. J., and Fox, B. J. (2019). An emerging consensus in the evolution, phylogeny, and systematics of marsupials and their fossil relatives (Metatheria). *J. Mammal.* 100, 802–837. doi: 10.1093/jmammal/gyz018

Elżanowski, A. (2014). More evidence for plesiomorphy of the quadrate in the Eocene anseriform avian genus *Presbyornis*. *Acta Palaeont. Pol.* 59, 821–825. doi: 10.4202/app.00027.2013

Evans, S. E., and Borsuk-Białynicka, M. (2009a). A small lepidosauromorph reptile from the Early Triassic of Poland. *Palaeont. Pol.* 65, 179–202. Available online at: http://www.palaeontologia.pan.pl/PP65/PP65_179-202.pdf

Evans, S. E., and Borsuk-Białynicka, M. (2009b). The Early Triassic stem-frog *Czatkobatrachus* from Poland. *Palaeont. Pol.* 65, 79–105. Available online at: http://www.palaeontologia.pan.pl/PP65/PP65_079-106.pdf

Evans, S. E., Lally, C., Chure, D. C., Elder, A., and Maisano, J. A. (2005). A Late Jurassic salamander (Amphibia: Caudata) from the Morrison Formation of North America. *Zool. J. Linn. Soc.* 143, 599–616. doi: 10.1111/j.1096-3642.2005.00159.x

Evans, S. E., Manabe, M., Noro, M., Isaji, S., and Yamaguchi, M. (2006). A long-bodied lizard from the Lower Cretaceous of Japan. *Palaeontology* 49, 1143–1165. doi: 10.1111/j.1475-4983.2006.00598.x

Evans, S. E., Prasad, G. V. R., and Manhas, B. K. (2002). Fossil lizards from the Jurassic Kota Formation of India. *J. Vert. Paleont.* 22, 299–312. doi: 10.1671/0272-4634(2002)022[0299:FLFTJK]2.0.CO;2

Evers, S. W., Barrett, P. M., and Benson, R. B. J. (2019). Anatomy of *Rhinochelys pulchriceps* (Protostegidae) and marine adaptation during the early evolution of chelonioids. *PeerJ* 7:e6811. doi: 10.7717/peerj.6811

Ezcurra, M. D., Nesbitt, S. J., Bronzati, M., Dalla Vecchia, F. M., Agnolin, F. M., Benson, R. B. J., et al. (2020). Enigmatic dinosaur precursors bridge the gap to the origin of Pterosauria. *Nature* 588, 445–449. doi: 10.1038/s41586-020-3011-4

Ezcurra, M. D., Scheyer, T. M., and Butler, R. J. (2014). The origin and early evolution of Sauria: reassessing the Permian saurian fossil record and the timing of the crocodile-lizard divergence. *PLOS ONE* 9:e89165. doi: 10.1371/journal.pone.0089165

Field, D. J., Benito, J., Chen, A., Jagt, J. W. M., and Ksepka, D. T. (2020). Late Cretaceous neornithine from Europe illuminates the origins of crown birds. *Nature* 579, 397–401. doi: 10.1038/s41586-020-2096-0

Ford, D. P., and Benson, R. B. J. (2018). A redescription of *Orovenator mayorum* (Sauropsida, Diapsida) using high-resolution μCT, and the consequences for early amniote phylogeny. *Pap. Palaeont.* 5, 197–239. doi: 10.1002/spp2.1236

Ford, D. P., and Benson, R. B. J. (2019). The phylogeny of early amniotes and the affinities of Parareptilia and Varanopidae. *Nat. Ecol. Evol.* 4, 57–65. doi: 10.1038/s41559-019-1047-3

Fostowicz-Frelik, Ł., and Kielan-Jaworowska, Z. (2002). Lower incisor in zalambdalestid mammals (Eutheria) and its phylogenetic implications. *Acta Palaeont. Pol.* 47, 177–180.

Fox, R. C. (2016). The status of *Schowalteria clemensi*, the Late Cretaceous taeniodont (Mammalia). *J. Vert. Paleont.* 36:e1211666. doi: 10.1080/02724634.2016.1211666

Fox, R. C., and Scott, C. S. (2011). A new, early Puercan (earliest Paleocene) species of *Purgatorius* (Plesiadapiformes, Primates) from Saskatchewan, Canada. *J. Paleont.* 85, 537–548. doi: 10.1666/10-059.1

Fraguas, Á., Comas-Rengifo, M. J., Goy, A., and Gómez, J. J. (2018). Upper Sinemurian – Pliensbachian calcareous nannofossil biostratigraphy of the E Rodiles section (Asturias, N Spain): a reference section for the connection between the Boreal and Tethyan Realms. *Newsl. Strat.* 51, 227–244. doi: 10.1127/nos/2017/0401

Frey, L., Coates, M., Ginter, M., Hairapetian, V., Rücklin, M., Jerjen, I., et al. (2019). The early elasmobranch *Phoebodus*: phylogenetic relationships, ecomorphology and a new time-scale for shark evolution. *Proc. R. Soc. B* 286:20191336. doi: 10.1098/rspb.2019.1336

Friedman, M. (2015). The early evolution of ray-finned fishes. *Palaeontology* 58, 213–228. doi: 10.1111/pala.12150

Funston, G. F., Brusatte, S. L., Williamson, T. E., Wible, J. R., Shelley, S. L., Bertrand, O. C., et al. (2020). New information on the rise of mammals. *Vertebr. Anat. Morphol. Palaeont.* 8:38. doi: 10.18435/vamp29365

Gaffney, E. S., and Jenkins, F. A. Jr. (2010). The cranial morphology of *Kayentachelys*, an Early Jurassic cryptodire, and the early history of turtles. *Acta Zool.* 91, 335–368. doi: 10.1111/j.1463-6395.2009.00439.x

Gaffney, E. S., Rich, T. H., Vickers-Rich, P., Constantine, A., Vacca, R., and Kool, L. (2007). *Chubutemys*, a new eucryptodiran turtle from the Early Cretaceous of Argentina, and the relationships of the Meiolaniidae. *Am. Mus. Novit.* 3599, 1–35. doi: 10.1206/0003-0082(2007)3599[1:CANETF]2.0.CO;2

Gaffney, E. S., Tong, H., and Meylan, P. A. (2006). Evolution of the side-necked turtles: the families Bothremydidae, Euraxemydidae, and Araripemydidae. *Bull. Am. Mus. Nat. Hist.* 300, 1–698. doi: 10.1206/0003-0090(2006)300[1:EOTSTT]2.0.CO;2

Galli, K. G., Buchwaldt, R., Lucas, S. G., and Tanner, L. (2018). New chemical abrasion thermal ionization mass spectrometry dates from the Brushy Basin Member, Morrison Formation, western Colorado: implications for dinosaur evolution. *The J. Geol.* 126, 473–486. doi: 10.1086/699208

Gao, K., and Norell, M. A. (2000). Taxonomic composition and systematics of Late Cretaceous lizard assemblages from Ukhaa Tolgod and adjacent localities, Mongolian Gobi desert. *Bull. Am. Mus. Nat. Hist.* 249, 1–118. doi: 10.1206/0003-0090(2000)249<0001:TCASOL>2.0.CO;2

Gao, K.-Q., and Shubin, N. H. (2003). Earliest known crown-group salamanders. *Nature* 422, 424–428. doi: 10.1038/nature01491

Gao, K.-Q., and Shubin, N. H. (2012). Late Jurassic salamandroid from western Liaoning, China. *Proc. Natl. Acad. Sci. U.S.A.* 109, 5767–5772. doi: 10.1073/pnas.1009828109

Garberoglio, F. F., Apesteguía, S., Simões, T. R., Palci, A., Gómez, R. O., Nydam, R. L., et al. (2019). New skulls and skeletons of the Cretaceous legged snake *Najash*, and the evolution of the modern snake body plan. *Sci. Adv.* 5:eaax5833. doi: 10.1126/sciadv.aax5883

Gardner, J. D., and Rage, J.-C. (2016). The fossil record of lissamphibians from Africa, Madagascar, and the Arabian Plate. *Palaeobiodiv. Palaeoenv.* 96, 169–220. doi: 10.1007/s12549-015-0221-0

Gentry, A. D., Ebersole, J. A., and Kiernan, C. R. (2019). *Asmodochelys parhami*, a new fossil marine turtle from the Campanian Demopolis Chalk and the stratigraphic congruence of competing marine turtle phylogenies. *R. Soc. Open Sci.* 6:191950. doi: 10.1098/rsos.191950

Giles, S., Xu, G.-H., Near, T. J., and Friedman, M. (2017). Early members of 'living fossil' lineage imply later origin of modern ray-finned fishes. *Nature* 549, 265–268. doi: 10.1038/nature23654

Glienke, S. (2015). Two new species of the genus *Batropetes* (Tetrapoda, Lepospondyli) from the Central European Rotliegend (basal Permian) in Germany. *J. Vert. Paleont.* 35:e918041. doi: 10.1080/02724634.2014.918041

Gómez, R. O. (2016). A new pipid frog from the Upper Cretaceous of Patagonia and early evolution of crown-group Pipidae. *Cret. Res.* 62, 52–64. doi: 10.1016/j.cretres.2016.02.006

González Ruiz, P., de la Fuente, M. S., and Fernández, M. S. (2019). New cranial fossils of the Jurassic turtle *Neusticemys neuquina* and phylogenetic relationships of the only thalassochelydian known from the eastern Pacific. *J. Paleont.* 94, 145–164. doi: 10.1017/jpa.2019.74

Goodrich, E. S. (1916). On the classification of the Reptilia. *Proc. R. Soc. Lond. B* 89, 261–276. doi: 10.1098/rspb.1916.0012

Goswami, A., Prasad, G. V. R., Upchurch, P., Boyer, D. M., Seiffert, E. R., Verma, O., et al. (2011). A radiation of arboreal basal eutherian mammals beginning in the Late Cretaceous of India. *Proc. Natl. Acad. Sci. U.S.A.* 108, 16333–16338. doi: 10.1073/pnas.1108723108

Goswami, S., Gierlowski-Kordesch, E., and Ghosh, P. (2016). Sedimentology of the Early Jurassic limestone beds of the Kota Formation: record of carbonate wetlands in a continental rift basin of India. *J. Paleolimnol.* 59, 21–38. doi: 10.1007/s10933-016-9918-y

Graur, D., and Martin, W. (2004). Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision. *Trends Genet.* 20, 80–86. doi: 10.1016/j.tig.2003.12.003

Groh, S. S., Upchurch, P., Barrett, P. M., and Day, J. J. (2019). The phylogenetic relationships of neosuchian crocodiles and their implications for the convergent evolution of the longirostrine condition. *Zool. J. Linn. Soc.* 188, 473–506. doi: 10.1093/zoolinnean/zlz117

Guindon, S. (2020). Rates and rocks: strengths and weaknesses of molecular dating methods. *Front. Genet.* 11:526. doi: 10.3389/fgene.2020.00526

Haddoumi, H., Allain, R., Meslouh, S., Metais, G., Monbaron, M., Pons, D., et al. (2015). Guelb el Ahmar (Bathonian, Anoual Syncline, eastern Morocco): First continental flora and fauna including mammals from the Middle Jurassic of Africa. *Gondw. Res.* 29, 290–319. doi: 10.1016/j.gr.2014.12.004

Halliday, T. J. D., Brandalise, de Andrade, M., Benton, M. J., and Efimov, M. B. (2013). A re-evaluation of goniopholidid crocodylomorph material from Central Asia: biogeographic and phylogenetic implications. *Acta Palaeont. Pol.* 60, 291–312. doi: 10.4202/app.2013.0018

Halliday, T. J. D., dos Reis, M., Tamuri, A. U., Ferguson-Gow, H., Yang, Z., and Goswami, A. (2019). Rapid morphological evolution in placental mammals post-dates the origin of the crown group. *Proc. R. Soc. B* 286:20182418. doi: 10.1098/rspb.2018.2418

Halliday, T. J. D., Upchurch, P., and Goswami, A. (2015). Resolving the relationships of Paleocene placental mammals. *Biol. Rev.* 92, 521–550. doi: 10.1111/brv.12242

Halliday, T. J. D., Upchurch, P., and Goswami, A. (2016). Eutherians experienced elevated evolutionary rates in the immediate aftermath of the Cretaceous–Palaeogene mass extinction. *Proc. R. Soc. B* 283:20153026. doi: 10.1098/rspb.2015.3026

Haridy, Y., MacDougall, M. J., and Reisz, R. R. (2017). The lower jaw of the Early Permian parareptile *Delorhynchus*, first evidence of multiple denticulate coronoids in a reptile. *Zool. J. Linn. Soc.* 184, 791–803. doi: 10.1093/zoolinnean/zlx085

Harjunmaa, E., Seidel, K., Häkkinen, T., Renvoisé, E., Corfe, I. J., Kallonen, A., et al. (2014). Replaying evolutionary transitions from the dental fossil record. *Nature* 512, 44–48. doi: 10.1038/nature13613

Hasegawa, H., Tada, R., Ichinnorov, N., and Minjin, C. (2008). Lithostratigraphy and depositional environments of the Upper Cretaceous Djadokhta Formation, Ulan Nuur basin, southern Mongolia, and its paleoclimatic implication. *J. As. Earth Sci.* 35, 13–26. doi: 10.1016/j.jseaes.2008.11.010

Head, J. J. (2015). Fossil calibration dates for molecular phylogenetic analysis of snakes 1: Serpentes, Alethinophidia, Boidae, Pythonidae. *Palaeont. Electron.* 18:6FC. doi: 10.26879/487

Henrici, A. C. (1998). A new pipoid anuran from the Late Jurassic Morrison Formation at Dinosaur National Monument, Utah. *J. Vert. Paleont.* 18, 321–332. doi: 10.1080/02724634.1998.10011060

Hermanson, G., Iori, F. V., Evers, S. W., Langer, M. C., and Ferreira, G. S. (2020). A small podocnemidoid (Pleurodira, Pelomedusoides) from the Late Cretaceous of Brazil, and the innervation and carotid circulation of side-necked turtles. *Pap. Palaeont.* 6, 329–347. doi: 10.1002/spp2.1300

Hicks, J. F., Fastovsky, D., Nichols, D. J., and Watabe, M. (2001). "Magnetostratigraphic correlation of Late Cretaceous dinosaur-bearing localities in the Nemegt and Ulan [sic] Nuur basins, Gobi desert," in *Proceedings of the Geological Society of America Annual Meeting, Mongolia*. Available online at: https://gsa.confex.com/gsa/2001AM/finalprogram/abstract_28817.htm

Hime, P. M., Lemmon, A. R., Moriarty Lemmon, E. C., Prendini, E., Brown, J. M., Thomson, R. C., et al. (2020). Phylogenomics reveals ancient gene tree discordance in the amphibian tree of life. *Syst. Biol.* 70, 49–66. doi: 10.1093/sysbio/syaa034

Hu, X., Li, L., Dai, H., Wang, P., Buffetaut, É., Wei, G., et al. (2020). Turtle remains from the Middle Jurassic Xintiangou Formation of Yunyang, Sichuan Basin, China | Restes de tortues de la Formation Xintiangou (Jurassique moyen) de Yunyang, bassin du Sichuan, Chine. *Ann. Paléont.* 106:102440. doi: 10.1016/j.annpal.2020.102440

Hutchinson, M. N., Skinner, A., and Lee, M. S. Y. (2012). *Tikiguania* and the antiquity of squamate reptiles (lizards and snakes). *Biol. Lett.* 8, 665–669. doi: 10.1098/rsbl.2011.1216

Huttenlocker, A. K., Grossnickle, D. M., Kirkland, J. I., Schultz, J. A., and Luo, Z.-X. (2018). Late-surviving stem mammal links the lowermost Cretaceous of North America and Gondwana. *Nature* 558, 108–112. doi: 10.1038/s41586-018-0126-y

International Commission on Zoological Nomenclature (1999). *International Code of Zoological Nomenclature*, 4 Edn. London: International Trust for Zoological Nomenclature. Available onlie at: https://www.iczn.org/the-code/the-international-code-of-zoological-nomenclature/the-code-online/

Irisarri, I., Baurain, D., Brinkmann, H., Delsuc, F., Sire, J.-Y., Kupfer, A., et al. (2017). Phylotranscriptomic consolidation of the jawed vertebrate timetree. *Nat. Ecol. Evol.* 1, 1370–1378. doi: 10.1038/s41559-017-0240-5

Ivanov, A. (2005). Early Permian chondrichthyans of the Middle and South Urals. *Rev. Bras. Paleont.* 8, 127–138.

Jia, J., and Gao, K.-Q. (2016). A new basal salamandroid (Amphibia, Urodela) from the Late Jurassic of Qinglong, Hebei Province, China. *PLOS ONE* 11:e0153834. doi: 10.1371/journal.pone.0153834

Jia, J., and Gao, K.-Q. (2019). A new stem hynobiid salamander (Urodela, Cryptobranchoidea) from the Upper Jurassic (Oxfordian) of Liaoning Province, China. *J. Vert. Paleont.* 39:e1588285. doi: 10.1080/02724634.2019.1588285

Jiang, J.-Q., Cai, C.-Y., and Huang, D.-Y. (2015). Progonocimicids (Hemiptera, Coleorrhyncha) from the Middle Jurassic Haifanggou Formation, western Liaoning, northeast China support stratigraphic correlation with the Daohugou beds. *Alcheringa* 40, 53–61. doi: 10.1080/03115518.2015.1086053

Jones, M. E. H., Anderson, C. L., Hipsley, C. A., Müller, J., Evans, S. E., and Schoch, R. R. (2013). Integration of molecules and new fossils supports a Triassic

origin for Lepidosauria (lizards, snakes, and tuatara). *BMC Evol. Biol.* 13:208. doi: 10.1186/1471-2148-13-208

Joyce, J. G., Anquetin, J., Cadena, E.-A., Claude, J., Danilov, I. G., Evers, S. W., et al. (2021). A nomenclature for fossil and living turtles using phylogenetically defined clade names. *Swiss J. Palaeont.* 140:5. doi: 10.1186/s13358-020-00211-x

Kangas, A. T., Evans, A. R., Thesleff, I., and Jernvall, J. (2004). Nonindependence of mammalian dental characters. *Nature* 432, 211–214. doi: 10.1038/nature02927

Kearney, M., and Clark, J. M. (2003). Problems due to missing data in phylogenetic analyses including fossils: a critical review. *J. Vert. Paleont.* 23, 263–274.

King, B. (2019). Which morphological characters are influential in a Bayesian phylogenetic analysis? Examples from the earliest osteichthyans. *Biol. Lett.* 15, 20190288. doi: 10.1098/rsbl.2019.0288

Kissel, R. (2010). *Morphology, Phylogeny and Evolution of Diadectidae (Cotylosauria: Diadectomorpha)*. Doctoral thesis. Toronto, ON: University of Toronto.

Klembara, J., Hain, M., Ruta, M., Berman, D. S., Pierce, S. E., and Henrici, A. C. (2019). Inner ear morphology of diadectomorphs and seymouriamorphs (Tetrapoda) uncovered by high-resolution X-ray microcomputed tomography, and the origin of the amniote crown group. *Palaeontology* 63, 131–154. doi: 10.1111/pala.12448

Kligman, B. T., Marsh, A. D., Sues, H.-D., and Sidor, C. A. (2020). A new non-mammalian eucynodont from the Chinle Formation (Triassic: Norian), and implications for the early Mesozoic equatorial cynodont record. *Biol. Lett.* 16:20200631. doi: 10.1098/rsbl.2020.0631

Koot, M. B., Cuny, G., Orchard, M. J., Richoz, S., Hart, M. B., and Twitchett, R. J. (2014). New hybodontiform and neoselachian sharks from the Lower Triassic of Oman. *J. Syst. Palaeont.* 13, 891–917. doi: 10.1080/14772019.2014.963179

Krause, D. W. (1978). Paleocene primates from western Canada. *Can. J. Earth Sci.* 15, 1250–1271. doi: 10.1139/e78-133

Ksepka, D. T., Parham, J. F., Allman, J. F., Benton, M. J., Carrano, M. T., Cranston, K. A., et al. (2015). The Fossil Calibration Database—a new resource for divergence dating. *Syst. Biol.* 64, 853–859. doi: 10.1093/sysbio/syv025

Kurochkin, E. N., Dyke, G. J., and Karhu, A. A. (2002). A new presbyornithid bird (Aves, Anseriformes) from the Late Cretaceous of southern Mongolia. *Am. Mus. Novit.* 3386, 1–11. doi: 10.1206/0003-0082(2002)386<0001:ANPBAA>2.0.CO;2

Langer, M. C., de Oliveira Martins, N., Manzig, P. C., de Souza Ferreira, G., de Almeida Marsola, J. C., Fortes, E., et al. (2019). A new desert-dwelling dinosaur (Theropoda, Noasaurinae) from the Cretaceous of south Brazil. *Sci. Rep.* 9, 9379. doi: 10.1038/s41598-019-45306-9

Lartillot, N. (2015). *PhyloBayes [Software]. Version 4.1c*. Available online at http://www.atgc-montpellier.fr/phylobayes/ (accessed April 26, 2021).

Laurin, M., and Piñeiro, G. (2018). Response: commentary: a reassessment of the taxonomic position of mesosaurs, and a surprising phylogeny of early amniotes. *Front. Earth Sci.* 6:220. doi: 10.3389/feart.2018.00220

Laurin, M., Lapauze, O., and Marjanović, D. (2019). What do ossification sequences tell us about the origin of extant amphibians? *bioRXiv: 352609 version 4, peer-reviewed by PCI Paleo*. doi: 10.1101/352609v4

Lawver, D. R., Debee, A. M., Clarke, J. A., and Rougier, G. W. (2011). A new enantiornithine bird from the Upper Cretaceous La Colonia Formation of Patagonia, Argentina. *Ann. Carnegie Mus.* 80, 35–42. doi: 10.2992/007.080.0104

Lee, M. S. Y., and Yates, A. M. (2018). Tip-dating and homoplasy: reconciling the shallow molecular divergences of modern gharials with their long fossil record. *Proc. R. Soc. B* 285:20181071. doi: 10.1098/rspb.2018.1071

Lefebvre, B., Gutiérrez-Marco, J. C., Lehnert, O., Martin, E. L. O., Nowak, H., Akodad, M., et al. (2017). Age calibration of the Lower Ordovician Fezouata Lagerstätte, Morocco. *Lethaia* 51, 296–311. doi: 10.1111/let.12240

Li, P.-P., Gao, K.-Q., Hou, L.-H., and Xu, X. (2007). A gliding lizard from the Early Cretaceous of China. *Proc. Natl. Acad. Sci. U.S.A* 104, 5507–5509. doi: 10.1073/pnas.0609552104

Liang, C., Liu, Y., Hu, Z., Li, X., Li, W., Zheng, C., et al. (2019). Provenance study from petrography and geochronology of Middle Jurassic Haifanggou Formation in Xingcheng Basin, western Liaoning Province. *Geol. J.* 55, 2420–2446. doi: 10.1002/gj.3509

Liu, Y.-H., Gai, Z.-K., and Zhu, M. (2017). New findings of galeaspids (Agnatha) from the Lower Devonian of Qujing, Yunnan, China. *Vert. PalAsiat.* 56, 1–15.

Lombard, R. E., and Bolt, J. R. (1999). A microsaur from the Mississippian of Illinois and a standard format for morphological characters. *J. Paleont.* 73, 908–923. doi: 10.1017/s0022336000040749

López-Arbarello, A., and Sferco, E. (2018). Neopterygian phylogeny: the merger assay. *R. Soc. Open Sci.* 5:172337. doi: 10.1098/rsos.172337

López-Torres, S., and Fostowicz-Frelik, Ł. (2018). A new Eocene anagalid (Mammalia: Euarchontoglires) from Mongolia and its implications for the group's phylogeny and dispersal. *Sci. Rep.* 8:13955. doi: 10.1038/s41598-018-32086-x

Lu, J., Giles, S., Friedman, M., and Zhu, M. (2017). A new stem sarcopterygian illuminates patterns of character evolution in early bony fishes. *Nat. Commun.* 8:1932. doi: 10.1038/s41467-017-01801-z

Lu, J., Zhu, M., Long, J. A., Zhao, W., Senden, T. J., Jia, L., et al. (2012). The earliest known stem-tetrapod[omorph] from the Lower Devonian of China. *Nat. Commun.* 3:1160. doi: 10.1038/ncomms2170

Luo, Z.-X., Yuan, C.-X., Meng, Q.-J., and Ji, Q. (2011). A Jurassic eutherian mammal and divergence of marsupials and placentals. *Nature* 476, 442–445. doi: 10.1038/nature10291

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. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 475, 1–11. doi: 10.1016/j.palaeo.2017.02.019

MacDougall, M. J., Winge, A., Ponstein, J., Jansen, M., Reisz, R. R., and Fröbisch, J. (2019). New information on the early [sic] Permian lanthanosuchoid *Feeserpeton oklahomensis* [sic] based on computed tomography. *PeerJ* 7:e7753. doi: 10.7717/peerj.7753

Maddin, H. C., Mann, A., and Hebert, B. (2019). Varanopid from the Carboniferous of Nova Scotia reveals evidence of parental care in amniotes. *Nat. Ecol. Evol.* 4, 50–56. doi: 10.1038/s41559-019-1030-z

Maidment, S. C. R., and Muxworthy, A. (2019). A chronostratigraphic framework for the Upper Jurassic Morrison Formation, western U.S.A. *J. Sedim. Res.* 89, 1017–1038. doi: 10.2110/jsr.2019.54

Mann, A., and Paterson, R. S. (2019). Cranial osteology and systematics of the enigmatic early 'sail-backed' synapsid *Echinerpeton intermedium* Reisz, 1972, and a review of the earliest 'pelycosaurs'. *J. Syst. Palaeont.* 18, 529–539. doi: 10.1080/14772019.2019.1648323

Mann, A., Gee, B. M., Pardo, J. D., Marjanović, D., Adams, G. R., Calthorpe, A. S., et al. (2020). Reassessment of historic 'microsaurs' from Joggins, Nova Scotia, reveals hidden diversity in the earliest amniote ecosystem. *Pap. Palaeontol.* 6, 605–625. doi: 10.1002/spp2.1316

Mann, A., McDaniel, E. J., McColville, E. R., and Maddin, H. C. (2019). *Carbonodraco lundi* gen et sp. nov., the oldest parareptile, from Linton, Ohio, and new insights into the early radiation of reptiles. *R. Soc. Open Sci.* 6:191191. doi: 10.1098/rsos.191191

Manz, C. L., Chester, S. G. B., Bloch, J. I., Silcox, M. T., and Sargis, E. J. (2015). New partial skeletons of Palaeocene Nyctitheriidae and evaluation of proposed euarchontan affinities. *Biol. Lett.* 11:20140911. doi: 10.1098/rsbl.2014.0911

Mao, F., Hu, Y., Li, C., Wang, Y., Hill Chase, M., Smith, A. K., et al. (2019). Integrated hearing and chewing modules decoupled in a Cretaceous stem therian mammal. *Science* 367, 305–308. doi: 10.1126/science.aay9220

Marjanović, D. (2019). Recalibrating the transcriptomic timetree of jawed vertebrates [first preprint of this paper]. *bioRXiv*:2019.12.19.882829v1 [Preprint]. doi: 10.1101/2019.12.19.882829v1

Marjanović, D., and Laurin, M. (2007). Fossils, molecules, divergence times, and the origin of lissamphibians. *Syst. Biol.* 56, 369–388. doi: 10.1080/10635150701397635

Marjanović, D., and Laurin, M. (2013a). The origin(s) of extant amphibians: a review with emphasis on the "lepospondyl hypothesis". *Geodiversitas* 35, 207–272. doi: 10.5252/g2013n1a8

Marjanović, D., and Laurin, M. (2013b). An updated paleontological timetree of lissamphibians, with comments on the anatomy of Jurassic crown-group salamanders (Urodela). *Hist. Biol.* 26, 535–550. doi: 10.1080/08912963.2013.797972

Marjanović, D., and Laurin, M. (2019). Phylogeny of Paleozoic limbed vertebrates reassessed through revision and expansion of the largest published relevant data matrix. *PeerJ* 6:e5565. doi: 10.7717/peerj.5565

Marjanović, D., and Witzmann, F. (2015). An extremely peramorphic newt (Urodela: Salamandridae: Pleurodelini) from the latest Oligocene of Germany,

and a new phylogenetic analysis of extant and extinct salamandrids. *PLOS ONE* 10:e0137068. doi: 10.1371/journal.pone.0137068

Maron, M., Muttoni, G., Rigo, M., Gianolla, P., and Kent, D. V. (2018). New magnetobiostratigraphic results from the Ladinian of the Dolomites and implications for the Triassic geomagnetic polarity timescale. *Palaeogeogr. Palaeoclimat. Palaeoecol.* 517, 52–73. doi: 10.1016/j.palaeo.2018.11.024

Marshall, C. R. (2019). Using the fossil record to evaluate timetree timescales. *Front. Genet.* 10:1049. doi: 10.3389/fgene.2019.01049

Martill, D. M., Tischlinger, H., and Longrich, N. R. (2015). A four-legged snake from the Early Cretaceous of Gondwana. *Science* 349, 416–419. doi: 10.1126/science.aaa9208

Mateus, O., Puértolas-Pascual, E., and Callapez, M. (2018). A new eusuchian crocodylomorph from the Cenomanian (Late Cretaceous) of Portugal reveals novel implications on the origin of Crocodylia. *Zool. J. Linn. Soc.* 186, 501–528. doi: 10.1093/zoolinnean/zly064

Matschiner, M. (2019). Selective sampling of species and fossils influences age estimates under the fossilized birth–death model. *Front. Genet.* 10:1064. doi: 10.3389/fgene.2019.01064

Matsui, M., Tominaga, A., Liu, W.-Z., and Tanaka-Ueno, T. (2008). Reduced genetic variation in the Japanese giant salamander, *Andrias japonicus* (Amphibia: Caudata). *Mol. Phyl. Evol.* 49, 318–326. doi: 10.1016/j.ympev.2008.07.020

Matsumoto, R., and Evans, S. E. (2018). The first record of albanerpetontid amphibians (Amphibia: Albanerpetontidae) from East Asia. *PLOS ONE* 13:e0189767. doi: 10.1371/journal.pone.0189767

Mayr, G., De Pietri, V. L., Scofield, R. P., and Worthy, T. H. (2018). On the taxonomic composition and phylogenetic affinities of the recently proposed clade Vegaviidae Agnolín et al., 2017–neornithine birds from the Upper Cretaceous of the Southern Hemisphere. *Cret. Res.* 86, 178–185. doi: 10.1016/j.cretres.2018.02.013

McLachlan, S. M. S., Kaiser, G. W., and Longrich, N. R. (2017). *Maaqwi cascadensis*: A large, marine diving bird (Avialae: Ornithurae) from the Upper Cretaceous of British Columbia, Canada. *PLOS ONE* 12:e0189473. doi: 10.1371/journal.pone.0189473

Meng, J. (2014). Mesozoic mammals of China: implications for phylogeny and early evolution of mammals. *Natl. Sci. Rev.* 1, 521–542. doi: 10.1093/nsr/nwu070

Milanese, F. N., Olivero, E. B., Raffi, M. E., Franceschinis, P. R., Gallo, L. C., Skinner, S. M., et al. (2018). Mid Campanian-Lower Maastrichtian magnetostratigraphy of the James Ross Basin, Antarctica: Chronostratigraphical implications. *Basin Res.* 31, 562–583. doi: 10.1111/bre.12334

Milner, A. C. (2019). A morphological revision of *Keraterpeton*, the earliest horned nectridean from the Pennsylvanian of England and Ireland. *Earth Env. Sci. Trans. R. Soc. Edinb.* 109, 237–253. doi: 10.1017/S1755691018000579

Milner, A. R., and Sequeira, S. E. K. (1994). The temnospondyl amphibians from the Viséan of East Kirkton, West Lothian, Scotland. *Trans. R. Soc. Edinb. Earth Sci.* 84, 331–361. doi: 10.1017/s0263593300006155

Missiaen, P., Escarguel, G., Hartenberger, J.-L., and Smith, T. (2012). A large new collection of *Palaeostylops* from the Paleocene of the Flaming Cliffs area (Ulan-Nur Basin, Gobi Desert, Mongolia), and an evaluation of the phylogenetic affinities of Arctostylopidae (Mammalia, Gliriformes). *Geobios* 45, 311–322. doi: 10.1016/j.geobios.2011.10.004

Modesto, S. P., Scott, D. M., MacDougall, M. J., Sues, H.-D., Evans, D. C., and Reisz, R. R. (2015). The oldest parareptile and the early diversification of reptiles. *Proc. R. Soc. B* 282:20141912. doi: 10.1098/rspb.2014.1912

Mongiardino Koch, N., and Gauthier, J. A. (2018). Noise and biases in genomic data may underlie radically different hypotheses for the position of Iguania within Squamata. *PLOS ONE* 13:e0202729. doi: 10.1371/journal.pone.0202729

Mongiardino Koch, N., and Parry, L. A. (2020). Death is on our side: paleontological data drastically modify phylogenetic hypotheses. *Syst. Biol.* 69, 1052–1067. doi: 10.1093/sysbio/syaa023

Morris, J. L., Puttick, M. N., Clark, J. W., Edwards, D., Kenrick, P., Pressel, S., et al. (2018). The timescale of early land plant evolution. *Proc. Natl. Acad. Sci. U.S.A.* 115, E2274–E2283. doi: 10.1073/pnas.1719588115

Müller, J., and Reisz, R. R. (2006). The phylogeny of early eureptiles: comparing parsimony and Bayesian approaches in the investigation of a basal fossil clade. *Syst. Biol.* 55, 503–511. doi: 10.1080/10635150600755396

Müller, J., and Reisz, R. R. (2007). Four well-constrained calibration points from the vertebrate fossil record for molecular clock estimates. *BioEssays* 27, 1069–1075. doi: 10.1002/bies.20286

Murphy, W. J., Eizirik, E., O'Brien, S. J., Madsen, O., Scally, M., Douady, C. J., et al. (2001). Resolution of the early placental mammal radiation using Bayesian phylogenetics. *Science* 294, 2348–2351. doi: 10.1126/science.1067179

Napoli, J. G., Williamson, T. E., Shelley, S. L., and Brusatte, S. L. (2017). A digital endocranial cast of the early Paleocene (Puercan) 'archaic' mammal *Onychodectes tisonensis* (Eutheria: Taeniodonta). *J. Mammal. Evol.* 25, 179–195. doi: 10.1007/s10914-017-9381-1

Naylor, B. G. (1981). Cryptobranchid salamanders from the Paleocene and Miocene of Saskatchewan. *Copeia* 1981:76. doi: 10.2307/1444042

Ni, X., Li, Q., Li, L., and Beard, K. C. (2016). Oligocene primates from China reveal divergence between African and Asian primate evolution. *Science* 352, 673–677. doi: 10.1126/science.aaf2107

Noonan, B. P., and Chippindale, P. T. (2006). Vicariant origin of Malagasy reptiles supports Late Cretaceous Antarctic land bridge. *The Am. Nat.* 168, 730–741. doi: 10.2307/4122245

Novacek, M., Rougier, G., Wible, J., McKenna, M. C., Dashzeveg, D., and Horovitz, I. (1997). Epipubic bones in eutherian mammals from the Late Cretaceous of Mongolia. *Nature* 389, 483–486. doi: 10.1038/39020

O'Connor, P. M., Krause, D. W., Stevens, N. J., Groenke, J. R., MacPhee, R. D. E., Kalthoff, D. C., et al. (2019). A new mammal from the Turonian–Campanian (Upper Cretaceous) Galula Formation, southwestern Tanzania. *Acta Palaeont. Pol.* 64, 65–84. doi: 10.4202/app.00568.2018

O'Connor, P. M., Turner, A. H., Groenke, J. R., Felice, R. N., Rogers, R. R., Krause, D. W., et al. (2020). Late Cretaceous bird from Madagascar reveals unique development of beaks. *Nature* 588, 272–276. doi: 10.1038/s41586-020-2945-x

O'Leary, M. A., Bloch, J. I., Flynn, J. J., Gaudin, T. J., Giallombardo, A., Giannini, N. P., et al. (2013). The placental mammal ancestor and the post-K-Pg radiation of placentals. *Science* 339, 662–667. doi: 10.1126/science.1229237

Olsen, P. E. (1988). "Paleontology and paleoecology of the Newark Supergroup (early Mesozoic, eastern North America)," in *Triassic-Jurassic Rifting: Continental Breakup and the Origin of the Atlantic Ocean and Passive Margins*, Part A. *Developments in Geotectonics* 22, ed. W. Manspeizer (Amsterdam: Elsevier), 185–230.

Olsen, P. E. (2010). *Fossil Great Lakes of the Newark Supergroup – 30 Years Later [Field Trip Guide]. NYSGA Field Trip 4*. doi: 10.1.1.407.418\&rep=rep1\&type=pdf

Palci, A., Hutchinson, M. N., Caldwell, M. W., Smith, K. T., and Lee, M. S. Y. (2019). The homologies and evolutionary reduction of the pelvis and hindlimbs in snakes, with the first report of ossified pelvic vestiges in an anomalepidid (*Liotyphlops beui*). *Zool. J. Linn. Soc.* 188, 630–652. doi: 10.1093/zoolinnean/zlz098

Panciroli, E., Benson, R. B. J., Walsh, S., Butler, R. J., Andrade Castro, T., Jones, M. E. H., et al. (2020). Diverse vertebrate assemblage of the Kilmaluag Formation (Bathonian, Middle Jurassic) of Skye, Scotland. *Earth Env. Sci. Trans. R. Soc. Edinb.* 111, 1–22. doi: 10.1017/S1755691020000055

Pardo, J. D., Holmes, R., and Anderson, J. S. (2018). An enigmatic braincase from Five Points, Ohio (Westphalian D) further supports a stem tetrapod position for aïstopods. *Earth Env. Sci. Trans. R. Soc. Edinb.* 109, 255–264. doi: 10.1017/S1755691018000567

Pardo, J. D., Lennie, K., and Anderson, J. S. (2020). Can we reliably calibrate deep nodes in the Retrapod tree? Case studies in deep tetrapod divergences. *Front. Genet.* 11:506749. doi: 10.3389/fgene.2020.506749

Pardo, J. D., Szostakiwskyj, M., Ahlberg, P. E., and Anderson, J. S. (2017). Hidden morphological diversity among early tetrapods. *Nature* 546, 642–645. doi: 10.1038/nature22966

Parham, J. F., Donoghue, P. C. J., Bell, C. J., Calway, T. D., Head, J. J., Holroyd, P. A., et al. (2011). Best practices for justifying fossil calibrations. *Syst. Biol.* 61, 346–359. doi: 10.1093/sysbio/syr107

Paton, R. L., Smithson, T. R., and Clack, J. A. (1999). An amniote-like skeleton from the Early Carboniferous of Scotland. *Nature* 398, 508–513. doi: 10.1038/19071

Pearson, M. R. (2016). *Phylogeny and systematic history of early salamanders*, Doctoral thesis. London: University College London.

Penkrot, T. A., and Zack, S. P. (2016). Tarsals of Sespedectinae (?Lipotyphla) from the middle Eocene of southern California, and the affinities of Eocene

'erinaceomorphs'. *J. Vert. Paleont.* 36:e1212059. doi: 10.1080/02724634.2016.1212059

Pérez-García, A. (2019). The African Aptian *Francemys gadoufaouaensis* gen. et sp. nov.: New data on the early diversification of Pelomedusoides (Testudines, Pleurodira) in northern Gondwana. *Cret. Res.* 102, 112–126. doi: 10.1016/j.cretres.2019.06.003

Phillips, M. J., and Fruciano, C. (2018). The soft explosive model of placental mammal evolution. *BMC Evol. Biol.* 18:104. doi: 10.1186/s12862-018-1218-x

Phillips, M. J., Bennett, T. H., and Lee, M. S. Y. (2009). Molecules, morphology, and ecology indicate a recent, amphibious ancestry for echidnas. *Proc. Natl. Acad. Sci. U.S.A.* 106, 17089–17094. doi: 10.1073/pnas.0904649106

Powell, C. L. E., Waskin, S., and Battistuzzi, F. U. (2020). Quantifying the error of secondary vs. distant primary calibrations in a simulated environment. *Front. Genet.* 11:252. doi: 10.3389/fgene.2020.00252

Prasad, G. V. R., and Manhas, B. K. (2007). A new docodont mammal from the Jurassic Kota Formation of India. *Palaeont. Electron.* 10:7A. Available online at: https://palaeo-electronica.org/2007_2/00117/index.html

Prasad, G. V. R., Parmar, V., and Kumar, D. (2014). "Recent vertebrate fossil discoveries from the Jurassic Kota Formation of India," in *Proceedings of the 74th Annual Meeting:Abstract, Society of Vertebrate Paleontology*, Vol. 208, ed. J. V. Paleont. Available online at: http://vertpaleo.org/CMSPages/GetFile.aspx?nodeguid=b51eb899-273b-4edb-ae8a-ae067050a3a8

Prasad, G. V. R., Verma, O., Sahni, A., Parmar, V., and Khosla, A. (2007). A Cretaceous hoofed mammal from India. *Science* 318:937. doi: 10.1126/science.1149267

Prevosti, F. J., and Chemisquy, M. A. (2009). The impact of missing data on real morphological phylogenies: influence of the number and distribution of missing entries. *Cladistics* 26, 326–339. doi: 10.1111/j.1096-0031.2009.00289.x

Pritchard, A. C., Gauthier, J. A., Hanson, M., Bever, G. S., and Bhullar, Bh.-A. S. (2018). A tiny triassic saurian from connecticut and the early evolution of the diapsid feeding apparatus. *Nat. Commun.* 9:1213. doi: 10.1038/s41467-018-03508-1

Püschel, H. P., Shelley, S. L., Williamson, T. E., Wible, J. R., and Brusatte, S. (2019). "Testing the phylogeny of Periptychidae and "archaic" Paleocene mammals under different optimality criteria," in *Proceedings of the 9th Annual Meeting: Society of Vertebrate Paleontology*, Vol. 174, ed. J. V. Paleont. Available online at: http://vertpaleo.org/Annual-Meeting/Annual-Meeting-Home/SVP-Program-book-v5_w-covers.aspx

Rage, J.-C., and Dutheil, D. B. (2008). Amphibians and squamates from the Cretaceous (Cenomanian) of Morocco – A preliminary study, with description of a new genus of pipid frog. *Palaeontogr. A* 285, 1–22. doi: 10.1127/pala/285/2008/1

Reeder, T. W., Townsend, T. M., Mulcahy, D. G., Noonan, B. P., Wood, P. L. Jr., Sites, J. W. Jr., et al. (2015). Integrated analyses resolve conflicts over squamate reptile phylogeny and reveal unexpected placements for fossil taxa. *PLOS ONE* 10:e0118199. doi: 10.1371/journal.pone.0118199

Reinhard, S., Voitel, S., and Kupfer, A. (2013). External fertilisation and paternal care in the paedomorphic salamander *Siren intermedia* Barnes, 1826 (Urodela: Sirenidae). *Zool. Anz.* 253, 1–15. doi: 10.1016/j.jcz.2013.06.002

Reisz, R. R., and Modesto, S. P. (1996). *Archerpeton anthracos* from the Joggins Formation of Nova Scotia: a microsaur, not a reptile. *Can. J. Earth Sci.* 33, 703–709. doi: 10.1139/e96-053

Renesto, S., and Bernardi, M. (2013). Redescription and phylogenetic relationships of *Megachirella wachtleri* Renesto et Posenato, 2003 (Reptilia, Diapsida). *Paläont. Z.* 88, 197–210. doi: 10.1007/s12542-013-0194-0

Reynoso, V.-H. (2005). Possible evidence of a venom apparatus in a Middle Jurassic sphenodontian from the Huizachal red beds of Tamaulipas, México. *J. Vert. Paleont.* 25, 646–654. doi: 10.1671/0272-4634(2005)025[0646:PEOAVA]2.0.CO;2

Richards, K. R., Sherwin, J. E., Smithson, T. R., Bennion, R. F., Davies, S. J., Marshall, J. E. A., et al. (2018). Diverse and durophagous: Early Carboniferous chondrichthyans from the Scottish Borders. *Earth Env. Sci. Trans. R. Soc. Edinb.* 108, 67–87. doi: 10.1017/S1755691018000166

Romano, C., Koot, M. B., Kogan, I., Brayard, A., Minikh, A. V., Brinkmann, W., et al. (2014a). Permian–Triassic Osteichthyes (bony fishes): diversity dynamics and body size evolution. *Biol. Rev.* 91, 106–147. doi: 10.1111/brv.12161

Romano, P. S. R., Gallo, V., Ramos, R. R. C., and Antonioli, L. (2014b). *Atolchelys lepida*, a new side-necked turtle from the Early Cretaceous of Brazil and the age of crown Pleurodira. *Biol. Lett.* 10:20140290. doi: 10.1098/rsbl.2014.0290

Romo de Vivar, P. R., Martinelli, A. G., Fonseca, P. H. M., and Bento Soares, M. (2020). To be or not to be: the hidden side of *Cargninia enigmatica* and other puzzling remains of Lepidosauromorpha from the Upper Triassic of Brazil. *J. Vert. Paleont.* 40:438. doi: 10.1080/02724634.2020.1828438

Ronchi, A., Santi, G., Marchetti, L., Bernardi, M., and Gianolla, P. (2018). First report on swimming trace fossils of fish from the Upper Permian and Lower Triassic of the Dolomites (Italy). *Ann. Soc. Geol. Pol.* 88, 111–125. doi: 10.14241/asgp.2018.013

Rong, Y.-F. (2018). Restudy of *Regalerpeton weichangensis* [sic] (Amphibia: Urodela) from the Lower Cretaceous of Hebei. *China. Vert. PalAsiat.* 56, 121–136. doi: 10.19615/j.cnki.1000-3118.170627

Rong, Y.-F., Vasilyan, D., Dong, L.-P., and Wang, Y. (2020). Revision of *Chunerpeton tianyiense* (Lissamphibia, Caudata): is it a cryptobranchid salamander? *Palaeoworld* (in press). doi: 10.1016/j.palwor.2020.12.001

Rougier, G. W., Chornogubsky, L., Casadio, S., Páez Arango, N., and Giallombardo, A. (2008). Mammals from the Allen Formation, Late Cretaceous, Argentina. *Cret. Res.* 30, 223–238. doi: 10.1016/j.cretres.2008.07.006

Ruane, S., Pyron, R. A., and Burbrink, F. T. (2010). Phylogenetic relationships of the Cretaceous frog *Beelzebufo* from Madagascar and the placement of fossil constraints based on temporal and phylogenetic evidence. *J. Evol. Biol.* 24, 274–285. doi: 10.1111/j.1420-9101.2010.02164.x

Ruta, M., Clack, J. A., and Smithson, T. R. (2020). A review of the stem amniote *Eldeceeon rolfei* from the Viséan of East Kirkton, Scotland. *Earth Env. Sci. Trans. R. Soc. Edinb.* 111, 173–192. doi: 10.1017/S1755691020000079

Salazar, C., Stinnesbeck, W., and Quinzio-Sinn, L. A. (2010). Ammonites from the Maastrichtian (Upper Cretaceous) Quiriquina Formation in central Chile. *N. Jb. Geol. Paläont. Abh.* 257, 181–236. doi: 10.1127/0077-7749/2010/0072

Sansom, I. J., and Andreev, P. S. (2018). "The Ordovician enigma. Fish, first appearances and phylogenetic controversies," in *Evolution and Development of Fishes*, eds Z. Johanson, C. Underwood, and M. Richter (Cambridge: Cambridge University Press), 59–70. doi: 10.1017/9781316832172.004

Sansom, I. J., Davies, N. S., Coates, M. I., Nicoll, R. S., and Ritchie, A. (2012). Chondrichthyan-like scales from the Middle Ordovician of Australia. *Palaeontology* 55, 243–247. doi: 10.1111/j.1475-4983.2012.01127.x

Sansom, R. S., and Wills, M. A. (2013). Fossilization causes organisms to appear erroneously primitive by distorting evolutionary trees. *Sci. Rep.* 3:2545. doi: 10.1038/srep02545

Sansom, R. S., Choate, P. G., Keating, J. N., and Randle, E. (2018). Parsimony, not Bayesian analysis, recovers more stratigraphically congruent phylogenetic trees. *Biol. Lett.* 14:20180263. doi: 10.1098/rsbl.2018.0263

Scarpetta, S. G. (2019). The first known fossil *Uma*: ecological evolution and the origins of North American fringe-toed lizards. *BMC Evol. Biol.* 19:178. doi: 10.1186/s12862-019-1501-5

Scarpetta, S. G. (2020a). Effects of phylogenetic uncertainty on fossil identification illustrated by a new and enigmatic Eocene iguanian. *Sci. Rep.* 10:15734. doi: 10.1038/s41598-020-72509-2

Scarpetta, S. G. (2020b). Combined-evidence analyses of ultraconserved elements and morphological data: an empirical example in iguanian lizards. *Biol. Lett.* 16:20200356. doi: 10.1098/rsbl.2020.0356

Schineider Fachini, T., Onary, S., Palci, A., Lee, M. S. Y., Bronzati, M., and Schmaltz Hsiou, A. (2020). Cretaceous blind snake from Brazil fills major gap in snake evolution. *iScience* 23:101834. doi: 10.1016/j.isci.2020.101834

Schoch, R. R., and Milner, A. R. (2014). *Temnospondyli I. Part 3A2 of Handbook of Paleoherpetology*, ed. H.-D. Sues (Munich: Dr. Friedrich Pfeil).

Schoch, R. R., and Sues, H.-D. (2017). Osteology of the Middle Triassic stem-turtle *Pappochelys rosinae* and the early evolution of the turtle skeleton. *J. Syst. Palaeont.* 16, 927–965. doi: 10.1080/14772019.2017.1354936

Schoch, R. R., Werneburg, R., and Voigt, S. (2020). A Triassic stem-salamander from Kyrgyzstan and the origin of salamanders. *Proc. Natl. Acad. Sci. U.S.A.* 117, 11584–11588. doi: 10.1073/pnas.2001424117

Scott, C. S., Fox, R. C., and Redman, C. M. (2016). A new species of the basal plesiadapiform *Purgatorius* (Mammalia, Primates) from the early Paleocene Ravenscrag Formation, Cypress Hills, southwest Saskatchewan, Canada: further taxonomic and dietary diversity in the earliest primates. *Can. J. Earth Sci.* 53, 343–354. doi: 10.1139/cjes-2015-0238

Shao, S., Li, L., Yang, Y., and Zhou, C. (2018). Hyperphalangy in a new sinemydid turtle from the Early Cretaceous Jehol Biota. *PeerJ* 6:e5371. doi: 10.7717/peerj.5371

Silcox, M. T., Bloch, J. I., Boyer, D. M., Chester, S. G. B., and López-Torres, S. (2017). The evolutionary radiation of plesiadapiforms. *Evol. Anthropol.* 26, 74–94. doi: 10.1002/evan.21526

Simões, T. R., Caldwell, M. W., Tałanda, M., Bernardi, M., Palci, A., Vernygora, O., et al. (2018). The origin of squamates revealed by a Middle Triassic lizard from the Italian Alps. *Nature* 557, 706–709. doi: 10.1038/s41586-018-0093-3

Simões, T. R., Caldwell, M. W., Weinschütz, L. C., Wilner, E., and Kellner, A. W. A. (2017). Mesozoic lizards from Brazil and their role in early squamate evolution in South America. *J. Herpetol.* 51, 307–315. doi: 10.1670/16-007

Simões, T. R., Vernygora, O., Caldwell, M. W., and Pierce, S. E. (2020). Megaevolutionary dynamics and the timing of evolutionary innovation in reptiles. *Nat. Commun.* 11:3322. doi: 10.1038/s41467-020-17190-9

Simões, T. R., Wilner, E., Caldwell, M. W., Weinschütz, L. C., and Kellner, A. W. A. (2015). A stem acrodontan lizard in the Cretaceous of Brazil revises early lizard evolution in Gondwana. *Nat. Commun.* 6:8149. doi: 10.1038/ncomms9149

Skutschas, P. P. (2013). Mesozoic salamanders and albanerpetontids of Middle Asia, Kazakhstan, and Siberia. *Palaeobiodiv. Palaeoenv.* 93, 441–457. doi: 10.1007/s12549-013-0126-8

Skutschas, P. P. (2014). *Kiyatriton leshchinskiyi* Averianov et Voronkevich, 2001, a crown-group salamander from the Lower Cretaceous of Western Siberia. *Russia. Cret. Res.* 51, 88–94. doi: 10.1016/j.cretres.2014.05.014

Skutschas, P. P. (2015). A new crown-group salamander from the Middle Jurassic of Western Siberia, Russia. *Palaeobio. Palaeoenv.* 96, 41–48. doi: 10.1007/s12549-015-0216-x

Smith, K. T. (2009a). A new lizard assemblage from the earliest Eocene (zone Wa0) of the Bighorn Basin, Wyoming, USA: biogeography during the warmest interval of the Cenozoic. *J. Syst. Palaeont.* 7, 299–358. doi: 10.1017/S1477201909002752

Smith, K. T. (2009b). Eocene lizards of the clade *Geiseltaliellus* from Messel and Geiseltal, Germany, and the early radiation of Iguanidae (Reptilia: Squamata). *Bull. Peabody Mus. Nat. Hist.* 50, 219–306. doi: 10.3374/014.050.0201

Smithson, T. R., Carroll, R. L., Panchen, A. L., and Andrews, S. M. (1994). *Westlothiana lizziae* from the Viséan of East Kirkton, West Lothian, Scotland, and the amniote stem. *Trans. R. Soc. Edinb. Earth Sci.* 84, 383–412. doi: 10.1017/s0263593300006192

Smithson, T. R., Wood, S. P., Marshall, J. E. A., and Clack, J. A. (2012). Earliest Carboniferous tetrapod and arthropod faunas from Scotland populate Romer's Gap. *Proc. Natl. Acad. Sci. U.S.A.* 109, 4532–4537. doi: 10.1073/pnas.1117332109

Sobral, G., Simões, T. R., and Schoch, R. R. (2020). A tiny new Middle Triassic stem-lepidosauromorph from Germany: implications for the early evolution of lepidosauromorphs and the Vellberg fauna. *Sci. Rep.* 10:2273. doi: 10.1038/s41598-020-58883-x

Solé, F., Smith, T., de Bast, E., Codrea, V., and Gheerbrant, E. (2016). New carnivoraforms from the latest Paleocene of Europe and their bearing on the origin and radiation of Carnivoraformes (Carnivoramorpha, Mammalia). *J. Vert. Paleont.* 36:e1082480. doi: 10.1080/02724634.2016.1082480

Sookias, R. B. (2019). Exploring the effects of character construction and choice, outgroups and analytical method on phylogenetic inference from discrete characters in extant crocodilians [sic]. *Zool. J. Linn. Soc.* 189, 670–699. doi: 10.1093/zoolinnean/zlz015

Spaulding, M., and Flynn, J. J. (2012). Phylogeny of the Carnivoramorpha: the impact of postcranial characters. *J. Syst. Palaeont.* 10, 653–677. doi: 10.1080/14772019.2011.630681

Springer, M. S., Foley, N. M., Brady, P. L., Gatesy, J., and Murphy, W. J. (2019). Evolutionary models for the diversification of placental mammals across the KPg boundary. *Front. Genet.* 10:1241. doi: 10.3389/fgene.2019.01241

Sterli, J., de la Fuente, M. S., and Rougier, G. W. (2019). New remains of *Condorchelys antiqua* (Testudinata) from the Early-Middle Jurassic of Patagonia: anatomy, phylogeny, and paedomorphosis in the early evolution of turtles. *J. Vert. Paleont.* 38:e1480112. doi: 10.1080/02724634.2018.1480112

Sterli, J., Martínez, R. N., Cerda, I. A., and Apaldetti, C. (2020). Appearances can be deceptive: bizarre shell microanatomy and histology in a new Triassic turtle (Testudinata) from Argentina at the dawn of turtles. *Pap. Palaeont.* 7, 1097–1132. doi: 10.1002/spp2.1334

Stocker, M. R., Nesbitt, S. J., Kligman, B. T., Paluh, D. J., Marsh, A. D., Blackburn, D. C., et al. (2019). The earliest equatorial record of frogs from the Late Triassic of Arizona. *Biol. Lett.* 15:20180922. doi: 10.1098/rsbl.2018.0922

Stumpf, S., and Kriwet, J. (2019). A new Pliensbachian elasmobranch (Vertebrata, Chondrichthyes) assemblage from Europe, and its contribution to the understanding of late Early Jurassic elasmobranch diversity and distributional patterns. *Paläont. Z.* 93, 637–658. doi: 10.1007/s12542-019-00451-4

Sullivan, R. M. (1982). Fossil lizards from Swain Quarry "Fort Union Formation," middle Paleocene (Torrejonian), Carbon County, Wyoming. *J. Paleont.* 56, 996–1010.

Sun, Z.-y, Tintori, A., Lombardo, C., and Jiang, D.-y (2016). New miniature neopterygians from the Middle Triassic of Yunnan Province, South China. *N. Jb. Geol. Paläont. Abh.* 282, 135–156. doi: 10.1127/njgpa/2016/0610

Sweetman, S. C., Smith, G., and Martill, D. M. (2017). Highly derived eutherian mammals from the earliest Cretaceous of southern Britain. *Acta Palaeont. Pol.* 62, 657–665. doi: 10.4202/app.00408.2017

Tambussi, C. P., Degrange, F. J., de Mendoza, R. S., Sferco, E., and Santillana, S. (2019). A stem anseriform from the early Palaeocene of Antarctica provides new key evidence in the early evolution of waterfowl. *Zool. J. Linn. Soc.* 186, 673–700. doi: 10.1093/zoolinnean/zly085

Tanner, L. H., and Lucas, S. G. (2015). The Triassic-Jurassic strata of the Newark Basin, USA: a complete and accurate astronomically-tuned timescale? *Stratigraphy* 12, 47–65.

The PLOS ONE Staff (2014). Correction: The origin and early evolution of Sauria: reassessing the Permian saurian fossil record and the timing of the crocodile-lizard divergence. *PLOS ONE* 9:e97828. doi: 10.1371/journal.pone.0097828

Tomiya, S. (2011). A new basal caniform (Mammalia: Carnivora) from the middle Eocene of North America and remarks on the phylogeny of early carnivorans. *PlOS ONE* 6:e24146. doi: 10.1371/journal.pone.0024146

Tomiya, S., and Tseng, Z. J. (2016). Whence the beardogs? Reappraisal of the Middle to Late [sic] Eocene '*Miacis*' from Texas, USA, and the origin of Amphicyonidae (Mammalia, Carnivora). *R. Soc. Open Sci.* 3:160518. doi: 10.1098/rsos.160518

Tong, H., Danilov, I., Ye, Y., Ouyang, H., and Peng, G. (2011). Middle Jurassic turtles from the Sichuan Basin, China: a review. *Geol. Mag.* 149, 675–695. doi: 10.1017/S0016756811000859

Tong, H., Li, L., and Ouyang, H. (2013). A revision of *Sinaspideretes wimani* Young & Chow, 1953 (Testudines: Cryptodira: Trionychoidae) from the Jurassic of the Sichuan Basin, China. *Geol. Mag.* 151, 600–610. doi: 10.1017/S0016756813000575

Trujillo, K. C., Carrano, M. T., and Chamberlain, K. R. (2015). "A U-Pb zircon age for Reed's Quarry 9, Upper jurassic morrison formation, Albany County, WY," in *Proceedings of the 75th Annual Meeting: Abstract, Society of Vertebrate Paleontology*, Vol. 226, ed. J. V. Paleont. Available online at: http://vertpaleo.org/PDFS/SVP-2015-Program-and-Abstract-Book-9-22-2015.aspx

Tykoski, R. S., Rowe, T. B., Ketcham, R. A., and Colbert, M. W. (2002). *Calsoyasuchus valliceps*, a new crocodyliform from the Early Jurassic Kayenta Formation of Arizona. *J. Vert. Paleont.* 22, 593–611. doi: 10.1671/0272-4634(2002)022[0593:CVANCF]2.0.CO;2

van Tuinen, M., and Hedges, S. B. (2004). The effects of external and internal fossil calibrations on the avian evolutionary timescale. *J. Paleont.* 78, 45–50. doi: 10.1666/0022-33602004078<0045:TEOEAI>2.0.CO;2

Vasilyan, D., and Böhme, M. (2012). Pronounced peramorphosis in lissamphibians—*Aviturus exsecratus* (Urodela, Cryptobranchidae) from the paleocene–eocene thermal maximum of Mongolia. *PLOS ONE* 7:e40665. doi: 10.1371/journal.pone.0040665

Vasilyan, D., Böhme, M., Chkhikvadze, V. M., Semenov, Y. A., and Joyce, W. G. (2013). A new giant salamander (Urodela, Pancryptobrancha) from the Miocene of Eastern Europe (Grytsiv, Ukraine). *J. Vert. Paleont.* 33, 301–318. doi: 10.1080/02724634.2013.722151

Vijayakumar, S. P., Pyron, R. A., Dinesh, K. P., Torsekar, V. R., Srikanthan, A. N., Swamy, P., et al. (2019). A new ancient lineage of frog (Anura: Nyctibatrachidae: Astrobatrachinae subfam. nov.) endemic to the Western Ghats of Peninsular India. *PeerJ* 7:e6457. doi: 10.7717/peerj.6457

von Heyden, L. (1870). *Entomologische Reise nach dem südlichen Spanien, der Sierra Guadarrama und Sierra Morena, Portugal und den Cantabrischen Gebirgen [...]*. Berlin: Entomologischer Verein. Available online at: https://books.google.com/books?id=FXBTAAAAcAAJ

Waddell, P. J., Kishino, H., and Ota, R. (2001). A phylogenetic foundation for comparative mammalian genomics. *Genome Inform.* 12, 141–154. doi: 10.11234/gi1990.12.141

Waggoner, B., and Collins, A. G. (2004). *Reductio ad absurdum*: testing the evolutionary relationships of Ediacaran and Paleozoic problematic fossils using molecular divergence dates. *J. Paleont.* 78, 51–61. doi: 10.1666/0022-33602004078<0051:RAATTE>2.0.CO;2

Wagner, P., Stanley, E. L., Daza, J. D., and Bauer, A. M. (2021). A new agamid lizard in mid-Cretaceous amber from northern Myanmar. *Cret. Res.* 124:104813. doi: 10.1016/j.cretres.2021.104813

Wang, H., Meng, J., and Wang, Y. (2019). Cretaceous fossil reveals a new pattern in mammalian middle ear evolution. *Nature* 576, 102–105. doi: 10.1038/s41586-019-1792-0

Wang, J., Ye, Y., Pei, R., Tian, Y., Feng, C., Zheng, D., et al. (2018). Age of Jurassic basal sauropods in Sichuan, China: a reappraisal of basal sauropod evolution. *GSA Bull.* 130, 1493–1500. doi: 10.1130/B31910.1

Wang, Y., and Evans, S. E. (2006). A new short-bodied salamander from the Upper Jurassic/Lower Cretaceous of China. *Acta Palaeont. Pol.* 51, 127–130.

Wang, Y., Dong, L., and Evans, S. E. (2015). Polydactyly and other limb abnormalities in the Jurassic salamander *Chunerpeton* from China. *Palaeobiol. Palaeoenv.* 96, 49–59. doi: 10.1007/s12549-015-0219-7

Wang, Y.-M., O'Connor, J. K., Li, D.-Q., and You, H.-L. (2013). Previously unrecognized ornithuromorph bird diversity in the Early Cretaceous Changma Basin, Gansu Province, northwestern China. *PLOS ONE* 8:e77693. doi: 10.1371/journal.pone.0077693

Wang, Y.-Q., Li, C.-K., Li, Q., and Li, D.-S. (2016). A synopsis of Paleocene stratigraphy and vertebrate paleontology in the Qianshan Basin, Anhui, China. *Vert. PalAsiat.* 54, 89–120.

Wesley, G. D., and Flynn, J. J. (2003). A revision of *Tapocyon* (Carnivoramorpha), including analysis of the first cranial specimens and identification of a new species. *J. Paleont.* 77, 769–783. doi: 10.1017/S0022336000044486

Whelan, N. V., and Halanych, K. M. (2016). Who let the CAT out of the bag? Accurately dealing with substitutional heterogeneity in phylogenomic analyses. *Syst. Biol.* 66, 232–255. doi: 10.1093/sysbio/syw084

Wible, J. R., Rougier, G. W., Novacek, M. J., and Asher, R. J. (2009). The eutherian mammal *Maelestes gobiensis* from the Late Cretaceous of Mongolia and the phylogeny of Cretaceous Eutheria. *Bull. Am. Mus. Nat. Hist.* 327, 1–123. doi: 10.1206/623.1

Wiens, J. J. (2003a). Incomplete taxa, incomplete characters, and phylogenetic accuracy: is there a missing data problem? *J. Vert. Paleont.* 23, 297–310. doi: 10.1671/0272-4634(2003)023[0297:ITICAP]2.0.CO;2

Wiens, J. J. (2003b). Missing data, incomplete taxa, and phylogenetic accuracy. *Syst. Biol.* 52, 528–538. doi: 10.1080/10635150390218330

Wiens, J. J. (2005a). Can incomplete taxa rescue phylogenetic analyses from long-branch attraction? *Syst. Biol.* 54, 731–742. doi: 10.1080/10635150500234583

Wiens, J. J. (2005b). Missing data and the design of phylogenetic analyses. *J. Biomed. Informatics* 39, 34–42. doi: 10.1016/j.jbi.2005.04.001

Wiens, J. J., Bonett, R. M., and Chippindale, P. T. (2005). Ontogeny discombobulates phylogeny: paedomorphosis and higher-level salamander relationships. *Syst. Biol.* 54, 91–110. doi: 10.1080/10635150590906037

Williamson, T. E., Flynn, A., Peppe, D. J., Heizler, M. T., Leslie, C. E., Secord, R., et al. (2019). "A revised, high-resolution age model for the Paleocene of the San Juan Basin, New Mexico, U.S.A[,] and implications for faunal and floral dynamics during the dawn of the age of mammals," in *Proceedings of the 79th Annual Meeting: Abstract, Society of Vertebrate Paleontology*, Vol. 219, ed. J. V. Paleont. Available online at: http://vertpaleo.org/Annual-Meeting/Annual-Meeting-Home/SVP-Program-book-v5_w-covers.aspx

Wolfe, J. M., Daley, A. C., Legg, D. A., and Edgecombe, G. D. (2016). Fossil calibrations for the arthropod Tree of Life. *Earth-Sci. Rev.* 160, 43–110. doi: 10.1016/j.earscirev.2016.06.008

Woodburne, M. O., Goin, F. J., Raigemborn, M. S., Heizler, M., Gelfo, J. N., and Oliveira, E. V. (2014). Revised timing of the South American early Paleogene land mammal ages. *J. S. Am. Earth Sci.* 54, 109–119. doi: 10.1016/j.jsames.2014.05.003

Woodhead, J., Reisz, R., Fox, D., Drysdale, R., Hellstrom, J., Maas, R., et al. (2010). Speleothem climate records from deep time? Exploring the potential with an example from the Permian. *Geology* 38, 455–458. doi: 10.1130/G30354.1

Xing, L.-D., Lockley, M. G., Wei, C., Gierliński, G. D., Li, J.-J., Persons, W. S. IV., et al. (2013). Two theropod track assemblages from the Jurassic of Chongqing, China, and the Jurassic stratigraphy of Sichuan Basin. *Vert. PalAsiat.* 51, 107–130.

Xu, G.-H. (2019). Osteology and phylogeny of *Robustichthys luopingensis*, the largest holostean fish in the Middle Triassic. *PeerJ* 7:e7184. doi: 10.7717/peerj.7184

Yuan, C.-X., Ji, Q., Meng, Q.-J., Tabrum, A. R., and Luo, Z.-X. (2013). Earliest evolution of multituberculate mammals revealed by a new Jurassic fossil. *Science* 341, 779–783. doi: 10.1126/science.1237970

Zack, S. P. (2009). *The Phylogeny of Eutherian Mammals: A New Analysis Emphasizing Dental and Postcranial Morphology of Paleogene taxa*, Doctoral thesis. Baltimore: Johns Hopkins University.

Zhang, Y., Wang, Y., Zhan, R., Fan, J., Zhou, Z., and Fang, X. (2014). *Ordovician and Silurian Stratigraphy and Palaeontology of Yunnan, Southwest China. A guide to the field excursion across the South China, Indochina and Sibumasu [Blocks]*. Beijing: Science Press.

Zhou, C.-F., Bhullar, Bh.-A. S., Neander, A. I, Martin, T., and Luo, Z.-X. (2019). New Jurassic mammaliaform sheds light on early evolution of mammal-like hyoid bones. *Science* 365, 276–279. doi: 10.1126/science.aau9345

Zhu, M., and Fan, J. (1995). *Youngolepis* from the Xishancun Formation (Early Lochkovian) of Qujing, China. *Geobios* 28(Suppl. 2), 293–299. doi: 10.1016/S0016-6995(95)80130-8

Žigaitė-Moro, Ž, Karatajūtė-Talimaa, V., Joachimski, M. M., and Jeffries, T. (2018). Silurian vertebrates from northern Mongolia: diversity, ecology and environment [abstract, 5th International Palaeontological Congress].Available online at: https://ipc5.sciencesconf.org/data/IPC5_Abstract_Book.pdf

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Marjanović. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership