Topography, divergence dates, and macroevolutionary inferences vary between different tip-dating approaches applied to fossil theropods (Dinosauria)

D. W. Bapst1,2, A. M. Wright3,4, N. J. Matzke5 and G. T. Lloyd6

1Department of Geology University of California Davis, One Shields Avenue Davis, CA 94568, USA
2Geological Engineering, South Dakota School of Mines and Technology, Rapid City, SD 57701, USA
3Department of Ecology, Evolution, and Organismal Biology, Iowa State University, Ames, IA 50011, USA
4Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, KS, 66605, USA
5Division of Evolution, Ecology, and Genetics, Research School of Biology, Australian National University, Canberra, Australian Capital Territory 2601, Australia
6Department of Biological Sciences, Faculty of Science, Macquarie University, Sydney, New South Wales 2109, Australia

Dated phylogenies of fossil taxa allow palaeobiologists to estimate the timing of major divergences and placement of extinct lineages, and to test macroevolutionary hypotheses. Recently developed Bayesian ‘tip-dating’ methods simultaneously infer and date the branching relationships among fossil taxa, and infer putative ancestral relationships. Using a previously published dataset for extinct theropod dinosaurs, we contrast the dated relationships inferred by several tip-dating approaches and evaluate potential downstream effects on phylogenetic comparative methods. We also compare tip-dating analyses to maximum-parsimony trees time-scaled via alternative a posteriori approaches including via the probabilistic cal3 method. Among tip-dating analyses, we find opposing but strongly supported relationships, despite similarity in inferred ancestors. Overall, tip-dating methods infer divergence dates often millions (or tens of millions) of years older than the earliest stratigraphic appearance of that clade. Model-comparison analyses of the pattern of body-size evolution found that the support for evolutionary mode can vary across and between tree samples from cal3 and tip-dating approaches. These differences suggest that model and software choice in dating analyses can have a substantial impact on the dated phylogenies obtained and broader evolutionary inferences.

1. Introduction

How fossil organisms are related to each other and to living lineages is a matter of interest both to the general public and the scientific community. This matter surpasses systematic placement, because our estimates of branching relationships and their timing have direct implications on macroevolutionary inferences. Few examples are better than Archaeopteryx, which has long caught public attention as a potential early bird, a position questioned by a recent maximum-parsimony phylogenetic analysis [1] but seemingly reaffirmed by a later maximum-likelihood analysis [2].

Parsimony versus model-based phylogenetics is only one great debate in palaeontological systematics: for decades, there has been disagreement about whether to consider stratigraphic occurrences when inferring relationships [3]. Recently, the oft-criticized parsimony-based ‘stratocladistics’ [4] has been reborn as Bayesian ‘tip-dating’ phylogenetics [5], where non-ultrametric time-scaled phylogenies of extinct fossil tip taxa are inferred as a function of both clock-like models of character
change and a tree prior, describing the distributions of divergence dates [6,7]. Most recently, these tree priors belong to the birth–
death-serial-sampling (BDSS) family of models, which involve both
diversification and sampling processes in the fossil record
[8]. Tip-dating with BDSS is implemented in Bayesian phyloge-
netics applications, such as BEAST2 and MrBayes, including
allowing for fossil taxa to be considered as potential sampled
ancestors [9,10]. Sampled-ancestor BDSS (‘SA-BDSS’, also
known as sampled-ancestor-birth–death or fossilized-
BDSS (noSA-BDSS’ or transmission birth–death process), where
sampling is synchronous with extinction [12]. Fossilization is unli-
kely to coincide with extinction, and thus noSA-BDSS may be
more fitting to pathogen phylogenetics in epidemiology.

Additionally, palaeobiologists often use a posteriori time-scaling
(APT) to secondarily date existing cladograms of extinct taxa.
While some APT methods are arbitrary rescaling algorithms,
the cal3 approach probabilistically dates divergences relative to
an SA-BDSS variant [13].

The diversity of approaches, models and software that can
be used to obtain a fossil-only time-scaled phylogeny calls for
an empirical comparison of tip-dating and probabilistic APT
methods. We choose to perform such an examination using the
matrix from Xu et al. [1], paired with stratigraphic occur-
rences. Although this matrix was outdated by later revisions
[14], its usage in studies employing different phylogenetic
methods makes it an attractive basis for a case study compar-
ing the results of dating approaches, which differ in the model
assumed and their implementation. Analysing the original Xu
et al. matrix also allows us to test whether Bayesian tip-dating
avoids atypical relationships [15,16] inferred by Lee & Worthy
[2]. Additionally, the emergence of avian dinosaurs has been a
focus for macromolecular fossil analyses never infer any ancestors, but similar correlations were
found with ancestor frequencies from cal3 using last appearance
ences never infer any ancestors, but similar correlations were
found with ancestor frequencies from cal3 using last appearance
times (see electronic supplementary material, Results).

While Archaeopteryx is popularly referred to as an ‘ancestral
bird’, it is a sampled ancestor in only 5% of the MrBayes pos-
terior (0% for BEAST2 SA), and then only to its close relative
Wishikloforia, not the more nested Avialae.

Comparisons of divergence dates for four nested avian clades (using a branch-based definition) show differences in
clad age estimates across approaches (figure 1). All APT
methods propose similar median ages for all four clades, much
younger than tip-dating estimates. This is due to maximum-
parsimony analyses placing the early-appearing Epidexipteryx
and Epidendrosaurus (i.e. the Scansoriopterygidae) as members of a branch-based Avialae (also observed in [1,2]), which con-
strains the age of the Avialae to the Middle Jurassic or older.
Tip-dating analyses vary in their placement of the Scansorioptery-
gidae but do not place them with the Avialae (see electronic
supplementary material, Results). Divergence date estimates
from cal3 for alternative non-Avian clades (Tyrannosauroidea,
Therizinosauroidea) resemble distributions obtained from tip-
dating (electronic supplementary material, figure S3), illustrating
how APT approaches are ultimately constrained by input topol-
gies. Even among tip-dating methods, there are differences,
with BEAST2 noSA estimating earlier root ages than SA ana-
lyses, and BEAST2 SA having wider age distributions than
MrBayes SA. Comparing age estimates for clades containing
identical taxa reveals that tip-dating approaches estimate
median divergence dates approximately 4.6–6 million
(myr) older than the earliest stratigraphic occurrence, although
root-ward nodes have median ages as much as 30–40 myr
older (see electronic supplementary materials, Results).

The original body-size analysis [17] used several APT
approaches, including the 1 myr minimum-branch-length
(MBL) approach. Under all time-scaling variants, they
found strong support for single-optima OU for Theropoda
and Maniraptora. Our reanalysis with alternative dated phyllogenies agrees, with high support for OU across all
approaches, particularly MBL (figure 2). However, our analy-
sis reveals that model support varies considerably across trees
from the same dating approach, with some phyllogenies
providing greater support for BM, a pattern that is most evident in cal3 and BEAST2 tree samples.

4. Discussion

While the Bayesian tip-dating analyses return broadly similar phylogenies, the contrast in topology, divergence dates and model support patterns between approaches suggests that workers need to carefully evaluate the models and priors applied, and the plausibility of complex models when data-sets are limited [20]. Tip-dating methods appear to favour divergence dates that are several Ma older than the minimum age, sometimes tens of millions of years (figure 1 and electronic supplementary material, figures S1–S2). One explanation may be that by treating taxa in tip-dating analyses as single tips (i.e. a single point occurrence), even though more than 20% are known from multiple occurrences across millions of years, the inferred level of sampling may be so low that the average morphological clock rate dominates, swamping increases in the rate of character change and erroneously leading to older dates. The differences between MrBayes and BEAST2 SA-BDSS analyses are difficult to explain given their congruence in a previous comparison (electronic supplementary material, table S3 in [10]). As that study had both extant and extinct taxa, our discrepancy might be due to MrBayes having poor MCMC mixing when all tips are extinct.

Our comparative analyses support previous findings of constrained body-size evolution [18], but there is variation among dating methods in the relative support for OU across trees. Variation in model support among sampled posterior trees reinforces the importance of not taking a single point estimate of phylogeny for downstream analyses [21], and highlights the need to evaluate dated phylogenies from multiple approaches. Future studies should investigate body-size evolution through additional analyses than model choice [22], particularly given the known bias of some dating methods toward supporting OU [23]. The similarity of cal3 and the BEAST2 comparative analyses suggests that cal3 may be a suitable alternative when tip-dating is inapplicable.

Figure 1. Age estimates for four nested theropod clades, across five different approaches for obtaining a dated phylogeny using the Xu et al. [1] dataset. Plotted boxes represent the first quantile, the median and the third quantile, respectively. From left to right, the approaches used are minimum node age dating and cal3, both applied to a sample of 100 randomly selected maximum-parsimony topologies, and three tip-dating approaches, the noSA and SA with BEAST2, and SA with MrBayes. The four clades examined (ordered left-to-right for each approach) are the root node (essentially, the Avetheropoda) and three clades with ‘branch-based’ definitions: the Coelurosauria (all taxa more closely related to modern birds than Sinraptor and Allosaurus), the Maniraptora (... than Ornithomimus) and the Avialae (... than Troodon or Deinonychus). (Online version in colour.)

Figure 2. The fit of models of body-size evolution across different sets of trees, each from a different approach for obtaining dated phylogenies. The right-most set is trees with edge lengths constrained to a minimum length of 1 myr; for all others, see caption for figure 1. Models are fit to a single dated tree, with Akaike weights for each model, for that tree (which sum to 1), figured as a bar along with other trees from that sample, as a block of 100 stacked barplots with borders removed. The barplots for each sample are reordered relative to their support for Brownian motion (BM), versus Ornstein–Uhlenbeck (OU) and early burst (EB). (Online version in colour.)
Palaeobiologists will likely become major users of tip-dating and probabilistic APT approaches to generate dated phylogenies, replacing the arbitrary APT approaches. However, these techniques are still maturing. Careful consideration and applying multiple dating approaches may be necessary to isolate artefacts and identify what consensus does exist across models and implementations.

Data accessibility. All data, input files and programming scripts for recreating all analyses and figures can be found (separate from the electronic supplementary material) at the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.n2g80.

References

5. O’Reilly JE, dos Reis M, Donoghue PCJ. 2015 Dating late artefacts and identify what consensus does exist across models and implementations.

Authors’ contributions. G.L. compiled the data and performed parsimony analyses; A.W. and N.M. performed tip-dating analyses; D.B. performed statistical comparisons and created figures; D.B., A.W., N.M. and G.L. wrote the manuscript and all authors agree to be held accountable for the content therein and approve the final version of the manuscript.

Competing interests. We declare we have no competing interests.

Funding. D.B. was supported by NSF EAR-1147537; A.W. was supported by NSF DEB-1256993; N.M. was supported by NSF EFJ-0832858 and ARC grant no. DE150101773; and G.L. was supported by ARC grant no. DE140101879.

Acknowledgments. This is Paleobiology Database publication #266.