

## Supplementary Information

### 1. Supplemental Methods

#### (a) Data Used

Our analyses utilize a 374 character matrix of 89 theropod dinosaurs from Xu et al. (2011), as provided by Lee and Worthy (2012), and stratigraphic occurrence data for the same taxa, obtained from the Paleobiology Database via Fossilworks (fossilworks.org; downloaded 24th November 2013). Data were obtained through the “Classic” interface in order to check for and exclude uncertain occurrences (cf., aff., ?, “” etc.). This temporal data consists of interval ages for both first and last appearance times, rather than precise point-estimates for these dates. We used the first appearances as the ‘times of observation’ for tips (Bapst, 2014), accounting for stratigraphic interval uncertainty, except for *Zanabazar junior*, which was constrained to 66 Ma, and thus provide an anchor for placing the trees resulting from tip-dating analyses on a common time-scale with trees from APT approaches.

#### (b) Bayesian Tip-Dating Analyses

*Models employed.* We applied three tip-dating approaches to this dataset, which varied mainly in the tree prior used (e.g. SA BDSS versus noSA BDSS) and the software application utilized. These analyses held a number of elements in common: all used the Mk<sub>v</sub> model of character change (Lewis, 2001) with gamma-distributed rate heterogeneity, and with equal transition frequencies between states (Wright et al., in press). The choice of this particular model was not arbitrary. Mk<sub>v</sub> is the standard model used for morphological characters in Bayesian analyses. In order to allow different characters to evolve at different rates (as might be expected from a priori expectations), we allow those per-character rates to vary under a gamma distribution, similar to previous tip-dating analyses (e.g. Slater, 2013). We did attempt analyses where transition frequencies between states were not equal with MrBayes, as that may better fit our prior intuition of how morphological characters evolve, but this attempt ran into issues (see bottom of this section). Tip-dates were always treated as random variables with uniform prior distributions, with bounds based on the stratigraphic uncertainty on the respective first appearances times. Analyses were run until they converged.

Two of our three tip-dating analyses were conducted with BEAST2 (i.e. BEAST version 2.3.1; Bouckaert et al., 2014). To build the complex XML files required by BEAST2, we used the R package BEASTmasterR (Matzke 2014). In estimating our trees and model parameters, we used the BEAST2 packages morph-models and SA (Gavryushkina et al. 2014). These BEAST2 add-ons are needed to use morphological data and conduct sampled ancestor BDSS analyses. As both SA and noSA BDSS fossil tip-dating analyses have been performed in the literature, we chose to apply BDSS tree priors in BEAST2 both with and without sampled ancestors (SA and noSA, or sampled-ancestor-birth-death ‘SABD’ and ‘BDSS’ as they are sometimes referred to in BEAST2 literature). We also performed divergence time estimation in MrBayes (version 3.2.6; Ronquist et al., 2012). For this analysis, we also used the ‘fossilized-birth-death’ (FBD) model (Heath et al., 2014; Zhang et al., 2016) with sampled ancestors, which is intended to be identical in its assumptions and the resulting output of the SABD model of BEAST2. Thus, in the main text, we refer to both as SA-BDSS, and the alternative as noSA-BDSS.

*Background on Birth-Death-Sequential-Sampling models.* Our goal is to understand whether general practices in paleobiology, such as using both SA and noSA BDSS models in tip-dating, produce interchangeable results or not. Thus, which of these BDSS models is most appropriate is tangential to our ultimate purpose. In general, though, we expect that sampled ancestor BDSS is a better description of the fossil record than a model without sampled ancestors. The BDSS of BEAST was originally implemented for use in tip-dating viral phylogenies in epidemiology, hence why Gavryushkina et al. (2014) go so far as to term it the ‘transmission birth-death process’. Because of the historical origins of this model in disease transmission modeling, this model presumes that sampling a lineage is always equivalent to the death of that lineage. This makes sense if tips are virus samples taken from patients who have been removed from an environment where they might otherwise transmit their contagion. However, if applied to paleontological data, this results in the unsettling assumption that either fossilization of one or more individuals somehow kills entire populations (!) or that extinction is somehow necessary for burial and preservation. SA-BDSS-type models are thus probably superior to currently-available noSA-BDSS models for use with fossil data.

Given the historical debates about the assumption or inference of “ancestors” in phylogenetic analyses, it is worth discussing the assumptions of SA-BDSS models in this regard. The use of the sampled-ancestor models does not necessitate assuming that sampled ancestral taxa exist in a given dataset, but merely acknowledging that if sampling rates are high enough, they could exist: using sampled-ancestor models means that we are relaxing our assumptions, rather than making stronger assumptions about the complete absence of sampled ancestors. Furthermore, it is difficult to entirely exclude the possibility of sampled ancestors in any fossil record, as simple modeling with birth-death-sampling processes reveals that individual taxa have at least a 10% of being a sampled indirect ancestor for some latter sampled descendant even at extremely low sampling rates (Foote, 1996; see Figure 2).

Initially, MCMC sampling was run for 20 million generations for all analyses. We observed that MrBayes had poor convergence after this number of generations. Therefore, we performed an additional 30 million generations (for a total of 50 million generations), and stopped the analysis when convergence was reached. The discrepancy between the number of generations required by BEAST2 and MrBayes is likely related to BEAST2 performing multiple moves (perturbations of the model parameters to generate the next step of the analysis) per MCMC generation (Vaughan et al. 2014).

*MCMC runs.* We assessed convergence and mixing via post-analysis examination of effective sample sizes and parameter distributions with Tracer (version 1.6). From this, we determined that the MrBayes analysis suffered from poor starting values and thus we uniformly applied a 30% burn-in to the BEAST2 and MrBayes analyses. For all Bayesian tip-dating analyses, post-burnin maximum-clade-credibility trees (MCCTs) with posterior probabilities were built using TreeAnnotator (version 2.3.1). Post-burnin half-compatibility (similar to majority-rule consensus) summary tree were estimated with MrBayes (for the MrBayes analysis) and, for the BEAST2 analyses, half-compatibility trees were built using a custom script (see supplemental data files) in R (version 3.2.3; R Core Team, 2016), using functions from packages ape (version 3.4; Paradis et al., 2004) and phangorn (version 2.0.2; Schliep, 2009).

*Additional MrBayes Attempt.* We attempted an alternative MrBayes analysis in which we used the symmetric Dirichlet hyperprior to allow characters to depart from the default Mk model assumption of

equal rates of change in both forward transitions and reversals (Wright et al., in press). However, we found that this hyperprior led to extremely poor mixing (as measured by low effective sample sizes in Tracer) in this parameter, and in several others. We also tried using a fixed value for this hyperprior (i.e., using the hyperprior as a prior). This improved mixing, but we still observed lower ESS (estimated sample size) for these estimations than for those without this prior. For this reason, the results presented in the paper correspond to the default assumption of equal forward and backward transition rates.

### **(c) Applying a posteriori Time-scaling (APT) Approaches**

In addition to tip-dating, we also applied several APT methods to a random sample of 100 most parsimonious trees obtained from an equal-weighted maximum parsimony analysis using TNT (v1.1; Goloboff et al. 2008), with a series of 20 replicates of new technology searches, followed by “bbreak=tbr” to find all MPTs, recovering 540 MPTs in total. These trees were time-scaled with a minimum dating approach, where clades are as old as their earliest appearing descendant (Norell, 1992; Smith, 1994), and the *cal3* approach, where node ages are stochastically sampled relative to their likelihood under a birth-death-sampling model, given prior rates of branching, extinction and sampling in the fossil record (Bapst, 2013). To maximize comparability between *cal3* and our tip-dating analyses, the input rates for *cal3* are the median parameter estimates from the BEAST2 SA posterior (Tables S1, S2). Tip dates were stochastically resampled under a uniform distribution, with bounds set by the interval each taxon first occurs in, except for *Z. junior* (as noted above). In addition, for the sake of contrasting results of comparative methods, the ‘minimum branch length’ (MBL) approach was utilized, as it is more commonly used in empirical studies (including Benson et al., 2013). Under MBL, the inter-node edge lengths of a minimum-dated tree are adjusted so that every branch must be equal to or greater than some set value (here, 1 Ma). All APT methods were applied with R (version 3.2.3; R Core Team, 2016), via package paleotree (v2.6, Bapst, 2012).

### **(d) Comparing Phylogenetic Comparative Methods**

For the sake of comparing node age distributions and the results of downstream comparative analyses, we randomly sampled 100 trees from the post-burn-in posteriors of our tip-dating analyses. We replicated analyses from Benson et al. (2014) across our samples of time-scaled phylogenies, in which they used maximum likelihood methods to analyze body-evolution using the R package geiger (version 2.0.6; Pennell et al., 2014). They compared the fit of several models, three of which they evaluated: one with a constant rate of change (Brownian Motion, BM), another where rates decrease over time (Early Burst, EB) and a third model with a single optimum acting as an attractor (Ornstein-Uhlenbeck, OU). Of the 89 theropods in the Xu et al. dataset, 47 had data which we could obtain from Benson et al., and we assigned species-level values by following their supplied R script, so that our data treatment protocols match as closely as possible. We quantified relative fit as the Akaike weights (Burnham and Anderson, 2002) of these models for each individual time-scaled tree, based on the corrected Akaike information criterion, as Benson et al. did.

## **2. Supplemental Results**

### **(a) Assessment of Tip-Dating Topologies**

Our tip-dating analyses had considerable differences in topology, especially between the analyses run under different programs (Figures S4-S8). There is weak support for troodontids as sister to Avialae in the MrBayes analyses (posterior 0.5142), while the converse relationship of troodontids as sister to dromaeosaurids is better supported in the BEAST2 noSA analysis (posterior 0.53). BEAST2 SA does not assign either hypothesis a posterior greater than 0.5, with the half-compatibility summary of the posterior (similar to a majority-rule consensus) showing a polytomy of the troodontids, dromaeosaurids, Avialae and a small clade composed of *Anchiornis* and *Xiaotingia*. The Alvarezsauridae are poorly constrained in the tip-dating analyses: the MrBayes analysis (Fig S8) does not find any stronger support for them beyond an unresolved polytomy of multiple lineages composed of non-tyrannosauroid coelurosaurs. The BEAST2 no-SA and SA analyses (Fig S6-S7) resolves Alvarezsauridae as either sister to Paraves or to a group composed of therizinosaurs and oviraptorosaurs (neither solution has a posterior probability greater than 0.5; however, their inclusion in that group is supported; posterior probabilities of 0.93 and 0.94 respectively).

A number of authors have criticized Lee and Worthy's (2012) presented maximum-likelihood (ML) topology (Fig S5) as having a number of seemingly unlikely or atypical relationships (e.g. Turner et al., 2012; Spencer and Wilburg, 2013; Xu and Pol, 2013). It is important for us to evaluate our results with respect to those criticisms. Most strikingly, Lee and Worthy's analysis did not find a monophyletic Tyrannosauroidea, converting the tyrannosauroids to a paraphyletic group ancestral to the remaining Coelurosauria, when most analyses before and since have recovered a monophyletic Tyrannosauroidea (Turner et al., 2012; Xu and Pol, 2013). However, all of our tip-dating analyses (Figs S6-S8) disagreed with Lee and Worthy's results, finding instead strong support for a monophyletic Tyrannosauroidea (posterior probabilities of 0.97 for BEAST2 noSA, 0.99 for BEAST2 SA and 0.9994 for MrBayes SA).

Turner et al. (2012) criticized the derived placement of *Epidexipteryx* and *Epidendrosaurus* (the Scansoriopterygidae) within the Avialae, in the RAXML topology presented by Lee and Worthy (2012). This placement is very stratigraphically incongruent for these early occurring taxa. Turner et al.'s own cladistic analysis, based on a revision of the Xu et al. character set, placed these taxa in different groups. One would, of course, *de facto* expect that tip-dating analyses would also correct for placements that awkwardly increase stratigraphic incongruence. Indeed, this bears out in this study, with the BEAST2 SA analysis placing them as a monophyletic group (posterior probability = 1) sister to the oviraptorosaurs (posterior probability of 0.53). The noSA analysis also places them as a monophyletic group (posterior probability = 1; Fig S6), but the half-compatibility topology leaves them in an unresolved polytomy with the oviraptorosaurs and the therizinosaurs (posterior probability = 0.58). The MrBayes SA analysis (Fig S8) also finds them to be monophyletic (posterior probability = 1), but places them as sister to a clade composed of both the oviraptorosaurs and therizinosaurs (posterior probability of 0.53).

*Archaeopteryx* and *Wellnhoferia* form a paraphyletic stem to the Avialae in Lee and Worthy's analysis (Figure S5), which Xu and Pol (2013) view with suspicion given that these two taxa have been considered probable taxonomic synonyms, and thus should form a monophyletic group in their view, as in the original Xu et al. analysis (Fig S4). (Note, however, that a monophyletic grouping would preclude the possibility that *Archaeopteryx* and *Wellnhoferia* are the same taxon *and* ancestral to more derived Avialae.) The BEAST2 analyses also prefer these as a paraphyletic grouping, with *Wellnhoferia* more closely related to the rest of Avialae than to *Archaeopteryx* (posterior probabilities of 0.75 for noSA, 0.77 for SA). In

contrast, MrBayes SA finds fairly high support for a monophyletic pairing (posterior probability of 0.9991).

### **(b) Statistical Analysis of Cross-Clade Variation in Divergence Date Estimates**

We can analyze the statistical variation in divergence dates by choosing an arbitrary minimum-dating tree as an exemplar, and then comparing to a another set of trees, calculating the shift in node ages for shared clades that contain an identical set of tip taxa relative to that clade's age on the exemplar. The per-node median of this shift in age estimates revealed that tip-dating analyses predict generally older node ages than the APT approaches, with BEAST2 and MrBayes predicting ages about 4-6 Ma older than the minimum-age exemplar (Fig S2(a)). Tip-dating sometimes suggested median node ages 30-40 Ma older than the exemplar. MrBayes SA appears to have a narrow distribution of shifts in node age (but with a flat tail) and appears to often prefer node ages younger than the exemplar, implying that it is preferring younger ages within the stratigraphic uncertainty. However, the variance in the per-node shift in node ages is high in the BEAST2 and *cal3* samples, but much lower in both minimum dating (where any variation simply represents stratigraphic uncertainty) and MrBayes (Fig S2(b)).

### **(c) Comparison of Sampled Ancestor Frequencies**

Our dataset of trees time-scaled with *cal3* using first appearance dates as the times of observation did not contain any taxa placed as sampled ancestors, unlike BEAST2 and MrBayes SA. This is likely because of the manner in which *cal3* selects node ages (and thus places taxa as ancestral), where the range of possible node ages is discretized, and selected relative to their probability weights, which may make it overly hard for a single date (the first appearance) to be selected, unlike the sampled-ancestor moves used by BEAST2 and MrBayes (Gavryushkina et al., 2014; Zhang et al., 2016). Thus, to evaluate how similar *cal3* ancestral inference was to the two tip-dating approaches that allow for sampled-ancestors, we conducted an additional comparison with trees time-scaled with *cal3*, but with the times of observation set to the last appearance date (LAD) instead.

This *cal3*-LAD sample estimates ancestors much more frequently than either of our tip-dating methods, with a median of 17 ancestors inferred on each tree examined, resulting in taxa being inferred to be ancestors over a much larger proportion of examined phylogenies (Fig S3). This is expected, as using the last appearance times allows *cal3* to consider the possibility of branching events anywhere between the first and last appearance of a taxon, any of which would be considered as treating a taxon as a sampled-ancestor (a capability of *cal3* not yet implemented in MrBayes or BEAST2). There is a fairly strong relationship between the taxa inferred as ancestors under *cal3* and either tip-dating method. Similar rank-order correlations were measured between the frequency of ancestor inference for taxa under *cal3* and BEAST2 SA (Spearman rho = 0.46, p-value = 4.19e-06) and MrBayes SA (Spearman rho = 0.46, p-value = 5.87e-06).

## **3. Supplemental Files in Dryad Repository**

The supplementary data package for this study consists of several sets of files.

The first two sets are input xml files for our two BEAST2 tip-dating analyses, and Nexus files containing the posterior tree sample, the maximum clade credibility (MCC; i.e. the phylogeny with the highest sampled posterior probability) tree (with a 30% burn-in) and the half-compatibility ('majority-rule') summary tree (with a 30% burn-in).

B2noSA\_inputFile.xml  
B2noSA\_majrule\_burn03.nex  
B2noSA\_mcc\_burn03.nex  
B2noSA\_treelog.nex

B2SA\_inputFile.xml  
B2SA\_majrule\_burn03.nex  
B2SA\_mcc\_burn03.nex  
B2SA\_treelog.nex

Summarizing BEAST2 tree samples as half-compatibility ('majority-rule') summary trees with posterior support values is not straightforward with typical software, so we wrote an R script to do this.

makeBEAST2\_majRule\_02-05-16.R

The next set contains nexus files associated with the MrBayes tip-dating analysis, including the input script, a nexus file containing 100 trees randomly sampled from the posterior (post 30% burn-in), and the MCC and half-compatibility ('majority-rule') summary trees (both with 30% burn-in) in nexus format.

MrBayesSA\_100treeSample\_burn03.nex  
MrBayesSA\_input\_script.nex  
MrBayesSA\_majrule\_burn03.nex  
MrBayesSA\_mcc\_burn03.nex

For comparing to tip-dating analyses, we conducted maximum-parsimony analyses using TNT and obtained 540 most-parsimonious topologies. These are provided in a Newick string format file, which can be read with function `read.tree` in R package `ape`:

TNT\_most\_parsimonious\_trees.tre

In doing our analyses, we also needed two topologies published previously by Xu et al. 2011 (a majority-rule tree from a maximum-parsimony analysis, provided in their Figure S6) and Lee and Worthy 2012 (a maximum-likelihood tree found using a RAXML analysis and figured in the main text of their paper) using this character matrix. These topologies don't seem to be available freely online, so we reproduced them from their figures and provide them here in Nexus format:

theropod\_tree\_XuEtAll2011\_FigS6\_02-01-16.nex  
theropod\_tree\_LeeWorthy2011\_RAXML\_02-01-16.nex

We also needed a dataset of body mass estimates from theropods for comparative analyses, taken from Benson et al. 2014. The original format provided by Benson et al. was not immediately applicable to our dataset, so we reformatted the data as a plain-text table:

mass\_data\_for\_PCMs\_BensonEtal14\_10-27-14.txt

Age data for taxa in this analysis were formatted for use in R analyses as a 'timeList' object, which is an R 'list' object composed of two separate matrices: one indicating the start and end times of geologic intervals, and a second indicating the first and last intervals of taxa (by referencing the first matrix). This is saved as a ASCII text representation that can be read back into R via the function 'dget()'.

timeList\_sorted\_asIs\_theropods\_DWB\_11-05-14.txt

All post-inference analysis and visualization were done in R, via an Rmarkdown script in RStudio. The script (a .Rmd), the resulting markdown PDF with output and figures, and the saved workspace file are included in our supplemental data materials:

birdtreecomparison\_03-07-16.Rmd  
birdtreecomparison\_03-07-16.pdf  
birdtree\_workspace\_03-07-2016.Rdata

Finally, all published figures were created entirely from within R, using the saved workspace from the Rmarkdown script and the following R script:

figures\_theropod\_06-17-16.R

All analyses were performed with BEAST v2.3.1, MrBayes v3.2.6, TNT v11, and R v3.2.3 with R packages ape v3.4, geiger v2.06, paleotree v2.6, and phangorn v2.02.

### **Supplementary References**

- Bapst, D. W. 2012. paleotree: an R package for paleontological and phylogenetic analyses of evolution. *Methods in Ecology and Evolution* 3(5):803-807.
- Bapst, D. W. 2014. Assessing the effect of time-scaling methods on phylogeny-based analyses in the fossil record. *Paleobiology* 40(3):331-351.
- Bouckaert, R., J. Heled, D. Kühnert, T. Vaughan, C.-H. Wu, D. Xie, M. A. Suchard, A. Rambaut, and A. J. Drummond. 2014. BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Comput Biol* 10(4):e1003537.
- Burnham, K. P., and D. R. Anderson. 2002. *Model Selection and Multitmodel Inference: A Practical Information-Theoretic Approach*. Springer, New York.
- Foote, M. 1996. On the Probability of Ancestors in the Fossil Record. *Paleobiology* 22(2):141-151.
- Gavryushkina, A., D. Welch, T. Stadler, and A. J. Drummond. 2014. Bayesian Inference of Sampled Ancestor Trees for Epidemiology and Fossil Calibration. *PLoS Comput Biol* 10(12):e1003919.
- Goloboff, P., J. Farris, and K. Nixon. 2008. TNT, a free program for phylogenetic analysis. *Cladistics* 24(5):774 - 786.

- Heath, T. A., J. P. Huelsenbeck, and T. Stadler. 2014. The fossilized birth–death process for coherent calibration of divergence-time estimates. *Proceedings of the National Academy of Sciences* 111(29):E2957-E2966.
- Lee, M. S. Y., and T. H. Worthy. 2012. Likelihood reinstates Archaeopteryx as a primitive bird. *Biology Letters* 8(2):299-303.
- Lewis, P. O. 2001. A Likelihood Approach to Estimating Phylogeny from Discrete Morphological Character Data. *Systematic Biology* 50(6):913-925.
- Matzke, N. J. (2014). "BEASTmaster: Automated conversion of NEXUS data to BEAST2 XML format, for fossil tip-dating and other uses." Online at PhyloWiki, <http://phylo.wikidot.com/beastmaster>. Accessed 12/01/2015.
- Norell, M. A. 1992. Taxic origin and temporal diversity: the effect of phylogeny. Pp. 89-118. *In* M. J. Novacek, and Q. D. Wheeler, eds. *Extinction and phylogeny*. Columbia University Press, New York.
- Paradis, E., J. Claude, and K. Strimmer. 2004. APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* 20(2):289-290.
- Pennell, M. W., J. M. Eastman, G. J. Slater, J. W. Brown, J. C. Uyeda, R. G. FitzJohn, M. E. Alfaro, and L. J. Harmon. 2014. geiger v2.0: an expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. *Bioinformatics* 30(15):2216-2218.
- R Core Team. 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Ronquist, F., M. Teslenko, P. van der Mark, D. L. Ayres, A. Darling, S. Höhna, B. Larget, L. Liu, M. A. Suchard, and J. P. Huelsenbeck. 2012. MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. *Systematic Biology* 61(3):539-542.
- Schliep, K. P. 2011. phangorn: phylogenetic analysis in R. *Bioinformatics* 27(4):592-593.
- Smith, A. B. 1994. *Systematics and the fossil record: documenting evolutionary patterns*. Blackwell Scientific, Oxford.
- Spencer, M. R., and E. W. Wilberg. 2013. Efficacy or convenience? Model-based approaches to phylogeny estimation using morphological data. *Cladistics* 29(6):663-671.
- Turner, A. H., P. J. Makovicky, and M. A. Norell. 2012. A Review of Dromaeosaurid Systematics and Paravian Phylogeny. *Bulletin of the American Museum of Natural History*:1-206.
- Vaughan, T. G., D. Kühnert, A. Popinga, D. Welch, and A. J. Drummond. 2014. Efficient Bayesian inference under the structured coalescent. *Bioinformatics* 30(16):2272-2279.
- Wright, A. M., G. T. Lloyd, and D. M. Hillis. In press. Modeling character change heterogeneity in phylogenetic analyses of morphology through the use of priors. *Systematic Biology*.
- Xu, X., and D. Pol. 2013. Archaeopteryx, paravian phylogenetic analyses, and the use of probability-based methods for palaeontological datasets. *Journal of Systematic Palaeontology* 12(3):323-334.
- Xu, X., H. You, K. Du, and F. Han. 2011. An Archaeopteryx-like theropod from China and the origin of Avialae. *Nature* 475(7357):465-470.
- Zhang, C., T. Stadler, S. Klopfstein, T. A. Heath, and F. Ronquist. 2016. Total-Evidence Dating under the Fossilized Birth–Death Process. *Systematic Biology* 65(2):228-249.



## Supplemental Tables

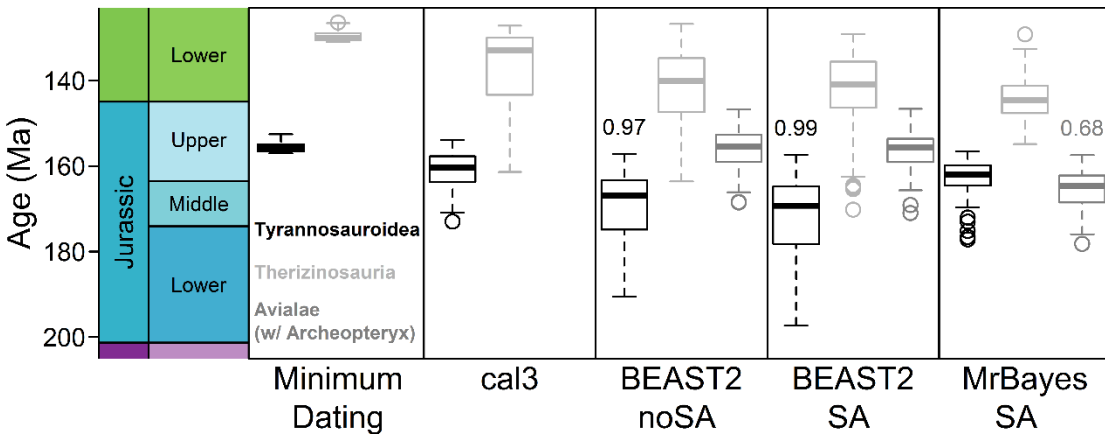
	<b>BEAST2 noSA</b>	<b>BEAST2 SA</b>
<b>lambda</b>	0.049	0.046
<b>mu</b>	0.005	0.005
<b>psi</b>	0.042	0.040

**Table S1. Comparison of post-burn-in posteriors medians for rate parameters of the birth-death-sampling processes for the BEAST2 analyses.** Here, lambda is the speciation rate, mu is the extinction rate, and psi is the sampling rate, all in units of per lineage Ma.

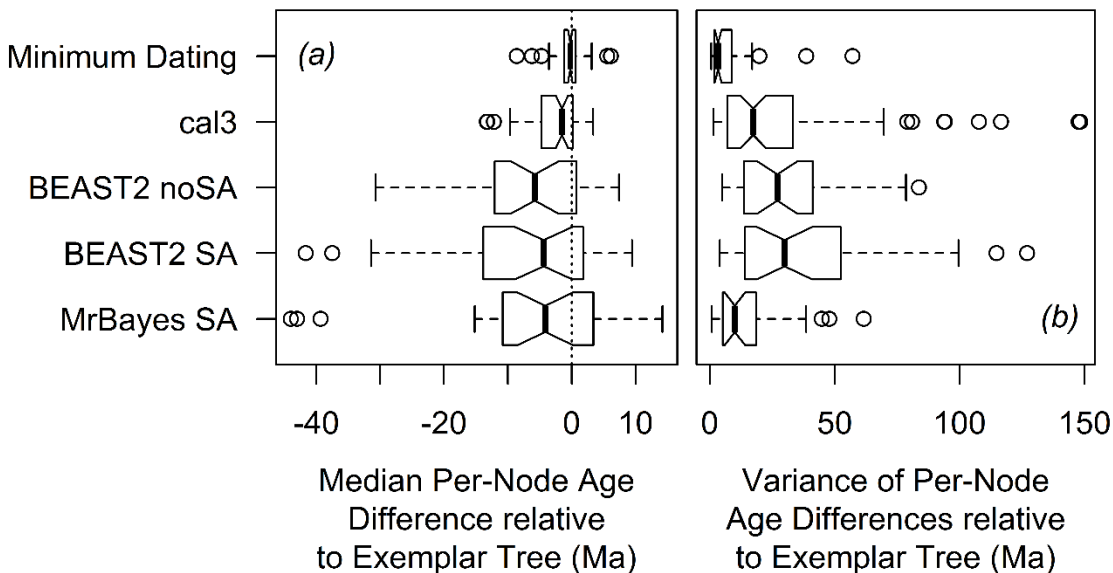
	<b>BEAST2 noSA</b>	<b>BEAST2 SA</b>	<b>MrBayes SA</b>
<b>Net Speciation</b>	0.0435	0.0414	0.0086
<b>Relative Extinction</b>	0.1061	0.1080	0.9865
<b>Relative Fossilization</b>	0.8900	0.8887	0.0058

**Table S2. Comparison of post-burn-in posteriors medians for parameters related to birth-death-sampling processes, across the three tip-dating analyses.** Here, we converted parameter estimates from BEAST2 to the notation from MrBayes, where Net Speciation is equal to  $\lambda - \mu$ , Relative Extinction is  $\mu / \lambda$ , and Relative Fossilization is  $\psi / (\mu + \psi)$ .

## Supplemental Figures



**Figure S1. Additional divergence date comparisons across time-scaling approaches, for several non-nesting clades.** The clades examined here are two non-avian clades: the Tyrannosauroidae (black) and the Therizinosauria (light gray), and an alternative ‘Avialae’ (medium gray) based not on a branch-based definition but rather defined as the MRCA of *Archaeopteryx* and modern birds, when that clade does not also include *Troodon* and *Deinonychus*. The numbers shown represent the posterior probabilities of those particular clades; when no number is shown, the posterior was 1, or effectively 1 when rounded to two significant digits. The approaches used are described in the caption for Fig 1.



**Figure S2. Comparing node ages for identical clades with a minimum-age exemplar tree.** The approaches used are described in the caption for Fig 1. (a) The median shift in node ages plotted as boxplots, with a dashed line indicating 0 Ma (no shift in the node ages). Negative shifts are those that where the node age in the exemplar is younger than the compared sample. (b) The variance of the per-node shifts in date estimates. Notches on the boxes indicate the expected confidence of the respective median values.

## Captions for Full Page Figures, On Following Pages:

**Figure S3. Proportion of sampled phylogenies in which a taxon was considered to be a sampled ancestor under three different approaches that allow for sampled ancestors: (a) *cal3*, (b) BEAST2 SA and (c) MrBayes SA.** Taxa are sorted by the sum proportions across all three approaches, and then secondarily sorted by taxon name. The *cal3* analyses shown here used the entire range for assessing potential ancestry (i.e. the last appearance time), rather than the first appearance only (unlike the tip-dating analyses), as the *cal3* analysis which only used the first appearance time inferred no ancestor-descendant relationships at all.

**Figure S4. The majority-rule topology from Xu et al. (2011)'s maximum-parsimony analysis, as shown in their figure S6, replotted with R package ape (Paradis et al., 2004).** Depicted branches are not to scale, instead apparent branch lengths are only to provide an ultrametric appearance for visual clarity. *Archaeopteryx* is labeled in red, and a blue dot indicates the node for the branch-based clade Avialae (i.e., all taxa more close related to modern birds than *Troodon* or *Deinonychus*).

**Figure S5. The maximum-likelihood topology from Lee and Worthy's (2012) RAXML analysis, as shown in their main text, replotted.** As in Figure S4, *Archaeopteryx* is labeled in red, and a blue dot indicates the node for the branch-based clade Avialae. Depicted branch lengths are not to scale.

**Figure S6. The half-compatibility topology from our BEAST2 noSA-BDSS tip-dating analysis, with nodes labeled with their posterior probabilities.** As in Figure S4, *Archaeopteryx* is labeled in red, and a blue dot indicates the node for the branch-based clade Avialae. Depicted branch lengths are not to scale.

**Figure S7. The half-compatibility topology from our BEAST2 SA-BDSS tip-dating analysis, with nodes labeled with their posterior probabilities.** As in Figure S4, *Archaeopteryx* is labeled in red, and a blue dot indicates the node for the branch-based clade Avialae. Depicted branch lengths are not to scale.

**Figure S8. The half-compatibility topology from our MrBayes SA-BDSS tip-dating analysis, with nodes labeled with their posterior probabilities.** As in Figure S4, *Archaeopteryx* is labeled in red, and a blue dot indicates the node for the branch-based clade Avialae. Depicted branch lengths are not to scale.

