Title: Ground truthing tip-dating methods using fossil Canidae reveals major differences in performance

Running head: Ground truthing tip-dating with canids

Authors: Nicholas J. Matzke[1,2,*]; April Wright[3]

[1]Discovery Early Career Researcher Award (DECRA) Fellow, Moritz Lab, Division of Ecology, Evolution, and Genetics, Research School of Biology, Canberra, ACT 2601 AUSTRALIA

[2]Work performed at: National Institute for Mathematical and Biological Synthesis (NIMBioS, www.nimbios.org), 1122 Volunteer Blvd., Suite 106, University of Tennessee, Knoxville, TN 37996-3410; and O'Meara Lab, 425a Hesler, Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN 37996

[3]Heath Lab, 228 Bessey Hall, Iowa State University, 2200 Osborn Dr., Ames, Iowa, 50011-4009

*Corresponding author. Email: nick.matzke@anu.edu.au

Abstract

Tip-dating methods are becoming popular alternatives to traditional "node-dating." However, they have not been extensively tested. We “ground-truth” the most popular methods against a dated tree of fossil Canidae derived from monographs by Wang and Tedford. Using a revised canid morphology dataset we compare MrBayes 3.2.5 to Beast 2.1.3 combined with *BEASTmasteR* (phylo.wikidot.com/beastmaster), an R package that automates the conversion of dates, priors, and NEXUS character matrices into the complex Beast2 XML format. We find that unconstrained MrBayes analysis under the uniform node age prior fails to retrieve reasonable results, exhibiting extremely high uncertainty in dates. On the other hand, Beast2 inference matches the ground-truth well, under both birth-death serially sampled (BDSS, disallowing direct ancestors) and sampled ancestor (SABD) tree models, as does MrBayes using BDSS. MrBayes using SABD seems to have difficulty converging in some analyses. These results, on a high quality fossil dataset, indicate that while tip-dating is very promising, methodological issues in tip-dating can have drastic effects, and require close attention, especially on more typical datasets where the distinction between "method problems" and "data problems" will be more difficult to detect.

Keywords: Tip-dating, total evidence dating, Canidae, MrBayes, BEASTmasteR, uniform node age prior

Main text (2100 words)

Testing phylogenetic inference methods against an externally known truth is highly desirable, but is rarely possible except when an experimenter manufactures a known evolutionary history either with simulations [1] or splitting populations of microbial/virus cultures [2]. Even when conducted, it is debatable to what extent manufactured histories are comparable to the complexity of real evolutionary histories, where heterogeneity of rates, environment, and data acquisition are likely to be significant [3].

Our goal is to assess Bayesian total evidence ("tip-dating") methods. These are methods where the ages of fossil OTUs (operational taxonomic units / terminal taxa) are used as the primary source of dating information, rather than priors on node dates. The latter, while valuable, is subject to a number of well-known criticisms [4-7] such as subjectivity and incomplete use of information. In addition, node-dating weakens inferences to the extent that it essentially constrains *a priori* some of the clade dates that we would prefer to infer.

A number of tip-dating methods and models recently have become available [5, 6, 8-11] and are being applied to empirical datasets [5-7, 12-27]. However, tip-dating results seem to vary widely between methods and datasets. Some papers conclude that their tip-dating results are implausible (e.g. [14, 27, 28]), and some infer dates that are wildly uncertain [17, 22, 29]. Evaluation of the methods against each other, or against expectations based on the fossil record, is hampered by the complexity of Bayesian analyses: differences in results might be produced by differences in clock models, tree models, site models, priors (user-set or default) on any of the parameters used in these models, issues in implementation (bugs in the code, decisions about defaults, MCMC operators, etc.), user error in setting up the analysis or post-analysis processing, or issues with the data itself. In addition, several of the issues above could be in play and interacting in any particular study.

As tip-dating methods are still in development, it is useful to "ground-truth" the major methods and models on an "ideal" empirical dataset, one where the fossil record is of sufficiently high quality that the true evolutionary tree and dates are broadly known even without complex computational methods. An ideal dataset would also meet the assumptions made by the models (Table 1). When multiple methods are run on the same ground-truth dataset, not only can differences in inference be attributed to differences in the method, but an assessment can be made about which methods are making inferences closer to the "known" truth. Methods that fail on the ideal dataset are unlikely to provide useful inferences on typical, non-ideal datasets.

Fossil datasets ideal for ground-truthing are few and far between, due to the vagaries of preservation and description, but one for which a strong argument can be made is the fossil Canidae (dog family; [30]). Canids avoid the challenges faced by most datasets (Table 1). The bulk of canid evolution occurred in North America from the Eocene to present, and their fossil record is approximately continuous, with fossil diversity greater than extant diversity (approximately 35 living species, at least 123 well-described fossil taxa). In addition, the group has been thoroughly revised in three major monographs on the three subfamilies of Canidae: the extinct Hesperocyoninae (~27 species, 40-15 Ma; [31], the extinct Borophaginae (~66 species, 34-2 Ma; [32], and the extinct and extant Caninae (>40 fossil species, 34 Ma-present; [33]). All living dogs thus represent a small surviving branch, originating 10-12 Ma, of a much more massive tree of fossil Canidae. Thus, apart from utility for ground-truthing, the Canidae serve as a useful group for comparing trait evolution inferences made with living-only versus living+fossil datasets [24, 34, 35].

Methods

Ground-truth tree and characters. The ground-truth tree was digitized from the monographs of Wang and Tedford, with judgment calls resolved in favor of preserving the authors' depiction of divergence times (SM). Morphological characters and dates came from the published matrix of Slater (2015) [24, 36].

Tip-dating analyses. MrBayes analyses were conducted by modification of Slater's commands file. A large number of variant MrBayes analyses (35 total) were constructed to investigate several issues that were noticed between MrBayes versions, documentation, and Slater's commands file (Supplemental Material, SM). Due to the many node date calibrations, these did not substantially change Slater's inferred tree, but if left unchanged they would have a large impact on unconstrained analyses. As the issues are subtle and due in part due to MrBayes' implementation or documentation, they may cause problems for other MrBayes users (SM).

We focused on six analyses (four MrBayes 3.2.5 analyses and two Beast2.1.3) to compare to the ground-truth tree, and to Slater's published analysis (mb1_orig) analyses. These were: (1) *mb1*: Slater's original uniform node age prior analysis including node date calibrations, with some corrections; (2) *mb8*: uniform node age prior, no node dates, flat priors on clock parameters, uniform(45,100) prior on the root age; (3) *mb9*: mb8 but with SABD tree prior and flat priors on speciation, extinction, and sampling rate; (4) *mb10*: mb9 but BDSS, i.e. disallowing sampled ancestors; (5) *r1*: Beast2 BDSS analysis with flat priors were used for each major parameter (mean and SD of the lognormal relaxed clock; and birth, death, and serial sampling rates); (6) *r2*: Beast2 SABD analysis with the same priors. Beast2 analyses were constructed with BEASTmasteR [37]; full details on the analyses and post-processing steps are presented in SM.

Results

The dated trees from the six focal analyses are compared in Figure 1 (plots of all trees are available in SM), and key priors and results are shown in Supplemental Table S1. The general picture is clear: the unconstrained MrBayes uniform node age prior analysis (mb8) produces implausibly old ages and huge uncertainties, with the age of Canidae overlapping the K-Pg boundary. This behavior was also noted by Slater (2015). The ground-truth dates of crown *Canis* (which includes genera *Cuon*, *Lycaon*, and *Xenocyon*) and crown Caninae are 3.2 and 11.7 Ma, but mb8 makes mean estimates of 27.5 and 38.9 Ma, and even the very wide HPDs (width 22-25 my) do not overlap the truth. More surprisingly, even Slater's highly constrained analysis (mb1), although much closer, does not produce HPDs (5.1-9.6 Ma; 17.8-25.5 Ma) that overlap the ground truth for these nodes. In contrast, both Beast2 analysis (r1 and r2) and MrBayes BDSS (mb10) produce mean estimates close to the truth, with narrower HPD widths (2-3 my).

The MrBayes SABD analysis (mb9) wrongly estimated these node ages as identical with the age of Canidae; this is due to mb9 misplacing *Lycaon pictus* (African wild dog) and *Cuon javanicus* (Dhole) in the extinct Borophagines. If this is ignored, the estimates are much closer, although the crown *Canis* estimate still fails to overlap (Supplemental Table S1, notes 5 and 6). A suggestion to repeat mb9 with 4 runs instead of 2 (Mike Lee, personal communication) did produce an mb9 result that placed these taxa in the conventional position (SM), but we present the unconventional result to emphasize that it appears much greater care is required to achieve convergence with MrBayes SABD than with other methods.

Comparing topological distances (Supplemental Table S1) provides a more systematic assessment of topology differences between analyses. Randomly chosen trees in the post-burnin posterior distribution of Slater's undated MrBayes analysis (mb2; SM) have a mean symmetric distance of 31.0% (95% C.I.=±8.3%). The dated MCC trees of all 8 focal analyses fall within this range, except for the MrBayes-uniform and -SABD analyses.

A comprehensive view of the correlation between date estimates and ground truth is shown in linear regression plots (SM) comparing the ages of nodes that are shared between the ground truth tree and estimated tree (i.e., have the same descendant OTUs after removing OTUs not shared by both trees; 127 OTUs were shared). While all analyses have statistically significant correlation to ground truth ages, mb8 has a lower R^2 (0.71) than the others (>0.9), and also has a systematic bias towards older ages (intercept=6.67±2.94; for other analyses intercept is ~0-1). All of the BD analyses have a bias towards underestimating dates near the base of the tree, where the ground truth tree is dominated by Hesperocyoninae, with a lower density of fossil OTUs and thus longer branches; the BD methods tend to infer shorter branches here, leading to younger ages. This bias leads to an underestimation of ages by about 2 my at age 20 Ma, to 4-5 my at 30 Ma. Prediction of tip dates generally succeeded in overlapping the true value, although uncertainty is usually high (5+ my) and extreme in the case of mb8 (17+ my). SABD analyses consistently produce younger date estimates than BDSS analyses, but the effect is small (~1 my).

Discussion

The result of greatest interest is the poor performance of the MrBayes uniform node age prior even in a "perfect-case" dataset. Whether or not this is surprising depends on researcher background. We suggest that reasoning from first principles suggests that effective tip-dating under the uniform node age prior will be difficult-to-impossible without strongly informative priors on node dates and/or clock rate and variability. Apart from such constraints, nothing in the tip dates or the uniform node age prior restricts the age of nodes below the dated tips; thus the node ages are, in effect, scaled up and down as the root age is sampled according to the root age prior (a required setting for the MrBayes uniform node age prior). Without informative priors, the clock rate and variability parameters will adjust along with the tree height; highly uncertain node dates will result.

Despite what first principles suggest, we suspect may surprise some researchers. The MrBayes uniform node age prior was the leading model in the early tip-dating literature (11/16 papers as of mid-2015, 9 of them as the exclusive Bayesian tip-dating method), and until recently (October 2014, v. 3.2.3) the uniform node age prior was the only option available in MrBayes. Early tip-dating efforts in Beast/Beast2 required tedious manual editing of XML and/or elaborate scripting efforts (such as BEASTmasteR), whereas MrBayes was relatively easy to use. Therefore, many early attempts at tip-dating used the uniform node age prior.

In contrast to the disappointing results with the uniform node age prior, analyses using the BDSS or SABD tree prior (mb10, r1, r2) fared well against ground truth. Given only the characters and tip-dates, and with uninformative priors on parameters and the root age, these analyses were able to estimate node ages with high accuracy. Surprisingly, these analyses outperformed the uniform node age prior even when this analysis was given substantial additional information in the form of many node calibrations (mb1). It seems that even well-constrained uniform node age prior analyses have a tendency to space node ages unrealistically evenly between calibrations and tip dates, regardless of morphological branch lengths. This can be seen in our mb1 analysis, for example in the old age for crown *Canis*; we have also observed the phenomenon with other datasets, and it seems to be a feature of published uniform node age prior analyses as well. Admittedly we have not devised a way to quantify the observation of “unrealistically even” node date estimates, but the phenomenon does appear to be real, at least when uniform node age prior results can be compared side-by-side to ground-truth and BDSS/SABD trees (e.g., mb1 vs. mb10, r1, r2), or to undated trees (e.g. mb2). The disagreement between the

MrBayes BDSS and SABD (m10 and m9) analyses about the position of *Lycaon+Cuon* is puzzling and is discussed further in SM.

Conclusions

Tip-dating with the uniform node age prior was explicitly introduced [6] as an alternative to node-dating, attractive precisely because tip-dating avoided various undesirable compromises that researchers are forced to make to when constructing node-age priors. Ronquist et al. [6] also critiqued Stadler's [38] BDSS prior as being "complete but unrealistic," particularly due to assumptions about constant birth/death/sampling rates and sampling in the Recent. They offered the uniform prior as an alternative, free of these difficulties. If, however, strongly informative node-age priors are required to produce reasonable results under the uniform node age prior, the main appeal of this prior is lost. The exploration of birth-death-sampling models for MrBayes [11] suggests that the future of tip-dating is likely to lay in adding realism to the BDSS-like models, rather than in attempting to devise wholly agnostic dating priors.

A great deal of work remains in the area of tip-dating in terms of methods testing and implementing more realistic methods. We have shown that "ground-truth" datasets, though rare and imperfect, are extremely useful in evaluating methods and models, bringing to light issues that would be less noticeable with lower-quality datasets and/or more complex setups (e.g., informative priors on parameters and node dates).

Data accessibility. All scripts, data files, and results files are available via a zipfile on Dryad (doi:XXX) [Note: Files also available at: https://github.com/nmatzke/Matzke_Wright_2016]

Competing interests. We have no competing interests.

Authors' Contributions. NJM wrote *BEASTmasteR*, conducted the Beast2 computational analyses and drafted the manuscript. AW contributed to MrBayes dating efforts and edited and corrected the manuscript.

Acknowledgements. We thank David Bapst, Graeme Lloyd, Jeremy Beaulieu, Kathryn Massana, Brian O'Meara, and Mike Lee for helpful comments and discussion, as well as the participants and attendees of the 2014 Society of Vertebrate Paleontology workshop and symposium on tip-dating. We also thank the BEAST developers and the *beast-users* Google Group, particularly Remco Bouckaert.

Funding. NJM was supported by NIMBioS fellowship under NSF Award #EFJ0832858, and ARC DECRA fellowship DE150101773. Work on this topic began under the NSF Bivalves in Time and Space grant (DEB-0919451). AW was supported by NSF DEB-1256993.

Captions for Figures and Tables

Figure 1. Plots of (a) the ground-truth tree, (b) Slater's original analysis, and (c-h) the six focal analyses. Illustrative nodes (dates listed in Supplemental Table S1) are (1) the common ancestor of crown (living) *Canis*, (2) the common ancestor of living Caninae (2), and (3) the common ancestor of the total group Canidae. Panel (f) shows only node (3) because all three nodes synonymize due to the misplacement of living *Lycaon*+*Cuon* in the extinct Borophagines. Note that the OTUs of the ground truth tree and the Slater dataset do not overlap completely: the Slater dataset lacks the living groups *Vulpes* (true foxes) and South American *Cerdocyonina* (e.g. crab-eating fox). The ground-truth tree lacks Slater's "outgroup" OTU (the branch below node (3)). For high-resolution trees with node labels, see Supplementary Information.

Table 1. Clade features that present challenges to tip-dating methods (or any dating methods). Canidae exhibit few of the issues that may confound dating in other clades (e.g. angiosperms, mammals, birds).

Supplemental Figure 1. Top: plot of the ground-truth tree, derived from digitization of the phylogenies of Canidae published in the monographs of Wang and Tedford, using TreeRogue. Bottom: Lineages through time plot of the ground-truth tree.

Supplemental Table S1. Five Bayesian tip-dating analyses are compared to (column 1) the conclusions of Tedford & Wang. The two Slater analyses (original, and a modification repairing some unintended issues; cols. 2-3) represent MrBayes analyses under a uniform node age prior, constrained by both tip dates and many node-date priors. The third run (col. 4) shows the effect of removing the node-age calibrations, and putting flat priors on the parameters for clock rate and variation. The fourth run (col. 5) shows the drastic effect of switching to a fossilized birth-death prior. Columns 5 and 6 show Beast2 analyses with flat clock priors and BDSS and SABD tree priors, respectively. These five analyses are drawn from the 40 analyses shown in Supplemental Table S2.

Supplemental Table S2. Summary settings and results of all 40 tip-dating analyses. As the table is large, it is presented as an Excel file.

Supplemental Data Files

Canidae_traceLogs.pdf -- Trace plots of key variables for all 40 analyses.

Canidae_treeLogs.pdf -- Plots of the MCC trees for all 40 analyses.

Ground_truth_vs_estimated_node_ages.pdf -- Linear regressions showing the correlation between the ground truth and estimated node ages, for nodes shared between the ground truth tree and estimated trees.

Canidae_ground_truth.newick – The “ground-truth” tree, derived from digitization of the phylogenies of Canidae published in the monographs of Wang and Tedford, using TreeRogue.

Table_S2_TipDate_runs_v3.xlsx – Summary of all 40 variant analyses (contains Supplemental Table S2, and some associated notes and file locations)

Matzke_Wright_SuppData.zip – A zipfile of all inputs, outputs, and scripts for all analyses.

References

[1] Huelsenbeck, J.P. 1995 Performance of phylogenetic methods in simulation. *Systematic Biology* **44**, 17-48. (doi:10.1093/sysbio/44.1.17).

[2] Hillis, D., Bull, J., White, M., Badgett, M. & Molineux, I. 1992 Experimental phylogenetics: generation of a known phylogeny. *Science* **255**, 589-592. (doi:10.1126/science.1736360).

[3] Hillis, D.M. 1995 Approaches for assessing phylogenetic accuracy. *Systematic Biology* **44**, 3-16. (doi:10.1093/sysbio/44.1.3).

[4] Parham, J.F., Donoghue, P.C.J., Bell, C.J., Calway, T.D., Head, J.J., Holroyd, P.A., Inoue, J.G., Irmis, R.B., Joyce, W.G., Ksepka, D.T., Patané, J.S.L., Smith, N.D., Tarver, J.E., van Tuinen, M., Yang, Z., Angielczyk, K.D., Greenwood, J.M., Hipsley, C.A., Jacobs, L., Makovicky, P.J., Müller, J., Smith, K.T., Theodor, J.M., Warnock, R.C.M. & Benton, M.J. 2012 Best practices for justifying fossil calibrations. *Systematic Biology* **61**, 346-359. (doi:10.1093/sysbio/syr107).

[5] Pyron, R.A. 2011 Divergence time estimation using fossils as terminal taxa and the origins of Lissamphibia. *Systematic Biology* **60**, 466-481. (doi:10.1093/sysbio/syr047).

[6] Ronquist, F., Klopfstein, S., Vilhelmsen, L., Schulmeister, S., Murray, D.L. & Rasnitsyn, A.P. 2012 A total-evidence approach to dating with fossils, applied to the early radiation of the Hymenoptera. *Systematic Biology* **61**, 973-999. (doi:10.1093/sysbio/sys058).

[7] Wood, H.M., Matzke, N.J., Gillespie, R.G. & Griswold, C.E. 2013 Treating fossils as terminal taxa in divergence time estimation reveals ancient vicariance patterns in the Palpimanoid spiders. *Systematic Biology* **62**, 264-284. (doi:10.1093/sysbio/sys092).

[8] Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M.A., Rambaut, A. & Drummond, A.J. 2014 BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput Biol* **10**, e1003537. (doi:10.1371/journal.pcbi.1003537).

[9] Gavryushkina, A., Welch, D., Stadler, T. & 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).

[10] Heath, T.A., Huelsenbeck, J.P. & Stadler, T. 2014 The fossilized birth–death process for coherent calibration of divergence-time estimates. *Proceedings of the National Academy of Sciences* **111**, E2957-E2966. (doi:10.1073/pnas.1319091111).

[11] Zhang, C., Stadler, T., Klopfstein, S., Heath, T.A. & Ronquist, F. 2016 Total-evidence dating under the fossilized birth-death process. *Systematic Biology* **65**, 228-249. (doi:10.1093/sysbio/syv080).

[12] Alexandrou, M.A., Swartz, B.A., Matzke, N.J. & Oakley, T.H. 2013 Genome duplication and multiple evolutionary origins of complex migratory behavior in Salmonidae. *Molecular Phylogenetics and Evolution* **69**, 514-523. (doi:http://dx.doi.org/10.1016/j.ympev.2013.07.026).

[13] Arcila, D., Alexander Pyron, R., Tyler, J.C., Ortí, G. & Betancur-R, R. 2015 An evaluation of fossil tip-dating versus node-age calibrations in tetraodontiform fishes (Teleostei: Percomorphaceae). *Molecular Phylogenetics and Evolution* **82, Part A**, 131-145. (doi:http://dx.doi.org/10.1016/j.ympev.2014.10.011).

[14] Beck, R.M.D. & Lee, M.S.Y. 2014 Ancient dates or accelerated rates? Morphological clocks and the antiquity of placental mammals. *Proceedings of the Royal Society of London B: Biological Sciences* **281**.

[15] Bracken-Grissom, H.D., Ahyong, S.T., Wilkinson, R.D., Feldmann, R.M., Schweitzer, C.E., Breinholt, J.W., Bendall, M., Palero, F., Chan, T.-Y., Felder, D.L., Robles, R., Chu, K.-H., Tsang, L.-M., Kim, D., Martin, J.W. & Crandall, K.A. 2014 The emergence of lobsters: phylogenetic relationships, morphological evolution and divergence time comparisons of an ancient group (Decapoda: Achelata, Astacidea, Glypheidea, Polychelida). *Systematic Biology* **63**, 457-479. (doi:10.1093/sysbio/syu008).

[16] Dornburg, A., Moore, J., Beaulieu, J.M., Eytan, R.I. & Near, T.J. 2015 The impact of shifts in marine biodiversity hotspots on patterns of range evolution: Evidence from the Holocentridae (squirrelfishes and soldierfishes). *Evolution* **69**, 146-161. (doi:10.1111/evo.12562).

[17] Grimm, G.W., Kapli, P., Bomfleur, B., McLoughlin, S. & Renner, S.S. 2015 Using more than the oldest fossils: dating Osmundaceae with three Bayesian clock approaches. *Systematic Biology* **64**, 396-405. (doi:10.1093/sysbio/syu108).

[18] Huang, D., Goldberg, E.E. & Roy, K. 2015 Fossils, phylogenies, and the challenge of preserving evolutionary history in the face of anthropogenic extinctions. *Proceedings of the National Academy of Sciences* **112**, 4909-4914. (doi:10.1073/pnas.1409886112).

[19] Lee, M.S.Y., Cau, A., Naish, D. & Dyke, G.J. 2014 Morphological clocks in paleontology, and a mid-Cretaceous origin of crown Aves. *Systematic Biology* **63**, 442-449. (doi:10.1093/sysbio/syt110).

[20] Lee, M.S.Y., Cau, A., Naish, D. & Dyke, G.J. 2014 Sustained miniaturization and anatomical innovation in the dinosaurian ancestors of birds. *Science* **345**, 562-566. (doi:10.1126/science.1252243).

[21] Marx, F.G. & Fordyce, R.E. 2015 Baleen boom and bust: a synthesis of mysticete phylogeny, diversity and disparity. *Royal Society Open Science* **2**.

[22] Sharma, P.P. & Giribet, G. 2014 A revised dated phylogeny of the arachnid order Opiliones. *Frontiers in Genetics* **5**, 255. (doi:10.3389/fgene.2014.00255).

[23] Slater, G.J. 2013 Phylogenetic evidence for a shift in the mode of mammalian body size evolution at the Cretaceous-Palaeogene boundary. *Methods in Ecology and Evolution* **4**, 734-744. (doi:10.1111/2041-210X.12084).

[24] Slater, G.J. 2015 Iterative adaptive radiations of fossil canids show no evidence for diversity-dependent trait evolution. *Proceedings of the National Academy of Sciences* **112**, 4897-4902. (doi:10.1073/pnas.1403666111).

[25] Drummond, A.J. & Stadler, T. 2016 Bayesian phylogenetic estimation of fossil ages. *arXiv*, 1-23. (doi:arXiv:1601.07447v1).

[26] Matzke, N.J. 2016 The evolution of antievolution policies after *Kitzmiller versus Dover*. *Science* **351**, 28-30. (doi:10.1126/science.aad4057).

[27] Wood, H.M., Gillespie, R.G., Griswold, C.E. & Wainwright, P.C. 2015 Why is Madagascar special? The extraordinarily slow evolution of pelican spiders (Araneae, Archaeidae). *Evolution* **69**, 462-481. (doi:10.1111/evo.12578).

[28] Puttick, M.N., Thomas, G.H. & Benton, M.J. 2016 Dating placentalia: morphological clocks fail to close the molecular fossil gap. *Evolution*. (doi:10.1111/evo.12907).

[29] Tseng, Z.J., Wang, X., Slater, G.J., Takeuchi, G.T., Li, Q., Liu, J. & Xie, G. 2013 Himalayan fossils of the oldest known pantherine establish ancient origin of big cats. *Proceedings of the Royal Society of London B: Biological Sciences* **281**.

[30] Wang, X.T., Richard H. 2008 *Dogs: their fossil relatives and evolutionary history*. New York, Columbia University Press.

[31] Wang, X. 1994 Phylogenetic systematics of the Hesperocyoninae (Carnivora, Canidae). *Bulletin of the American Museum of Natural History* **221**, 1-207.

[32] Wang, X.T., Richard H.; Taylor, Beryl E. 1999 Phylogenetic systematics of the Borophaginae. *Bulletin of the American Museum of Natural History* **243**, 1-391.

[33] Tedford, R.H.W., Xiaoming; Taylor, Beryl E. 2009 Phylogenetic systematics of the North American fossil Caninae (Carnivora, Canidae). *Bulletin of the American Museum of Natural History* **325**, 1-218.

[34] Finarelli, J.A. & Flynn, J.J. 2006 Ancestral state reconstruction of body size in the Caniformia (Carnivora, Mammalia): the effects of incorporating data from the

fossil record. *Systematic Biology* **55**, 301-313. (doi:10.1080/10635150500541698).

[35] Jablonski, D. & Shubin, N.H. 2015 The future of the fossil record: Paleontology in the 21st century. *Proceedings of the National Academy of Sciences* **112**, 4852-4858. (doi:10.1073/pnas.1505146112).

[36] Slater, G.J. 2015. Data from: Iterative adaptive radiations of fossil canids show no evidence for diversity-dependent trait evolution. Dryad. Accessed May 1, 2015. http://dx.doi.org/10.5061/dryad.9qd51

[37] Matzke, N.J. 2015 BEASTmasteR: automated conversion of NEXUS data to BEAST2 XML format, for fossil tip-dating and other uses. PhyloWiki. http://phylo.wikidot.com/beastmaster

[38] Stadler, T. 2010 Sampling-through-time in birth-death trees. *Journal of Theoretical Biology* **267**, 396-404. (doi:10.1016/j.jtbi.2010.09.010).

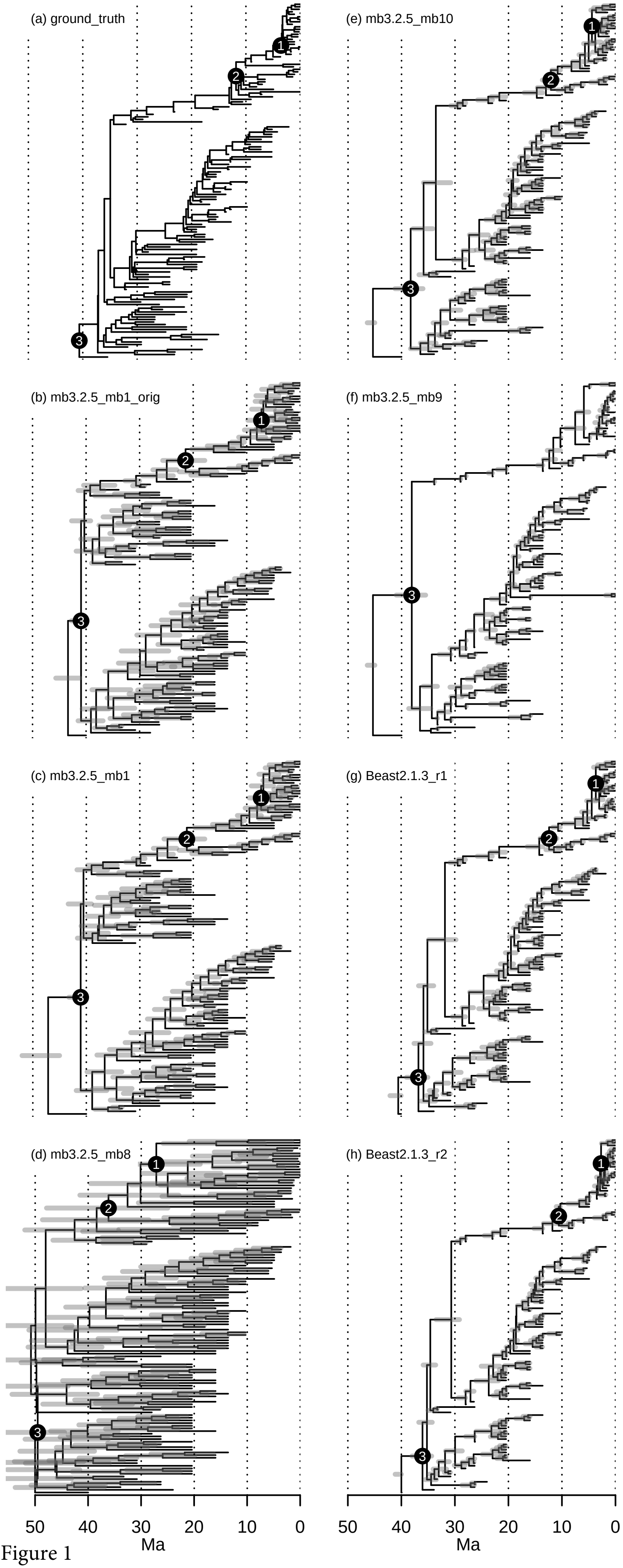

(a) ground_truth
(b) mb3.2.5_mb1_orig
(c) mb3.2.5_mb1
(d) mb3.2.5_mb8
(e) mb3.2.5_mb10
(f) mb3.2.5_mb9
(g) Beast2.1.3_r1
(h) Beast2.1.3_r2
1
2
3
50
40
30
20
10
0
Ma

Figure 1

Table 1. Clade features that present challenges to tip-dating methods (or any dating estimation methods. Canidae exhibit few of the issues that may confound dating in other clades (e.g. angiosperms, mammals, birds).

Clade features that make tip-dating challenging	Examples	Canidae
Clade occupies widely disparate niches (habitats, body form, etc.)	angiosperms, mammals	Clade in about the same ecological niche (carnivore)
Clade spans a mass extinction and post-extinction diversification	mammals, birds	Approximately constant macroevolutionary regime
Clade has a massive worldwide radiation	angiosperms, mammals, birds	Mostly endemic to a single region (North America) for most of Canidae history
Fossils have few characters	angiosperms (pollen), bivalves	Fossils have many characters (100+)
Fossils episodic or scarce near possible clade origin	placentals, angiosperms, Cambrian arthropods	Fossils preserved continuously throughout clade history (40-0 Ma)
Possibly major changes in preservability	angiosperms (woody vs. herbaceous); Cambrian phyla (soft vs. hard parts); hominids (forest vs. savanna habitats)	Approximately constant preservability
Likely changes in molecular/ morphological rate (due to major changes in body size, population size, growth rate, etc.)	angiosperms (woody vs. herbaceous, annuals vs. perennials)	Moderate change
Available coded fossils represent only a small proportion of total known diversity	E.g. O'Leary et al. (2012) placental dataset	Coded fossil diversity greatly exceeds extant diversity

Supplemental Material for Matzke and Wright (2016), "Ground truthing tip-dating methods using fossil Canidae reveals major differences in performance"

Supplemental Methods

Ground-truth tree. The ground-truth tree was digitized using TreeRogue [1], with judgment calls resolved in favour of preserving Wang and Tedford's depictions of divergence times. The source figures were, specifically, Figure 65 of Wang (1994) [2]; Figure 141 of Wang et al. (1999) [3]; and Figure 66 of Tedford et al. (2009) [4]. Digitization resolution was <1 my, undoubtedly more precise than either the ground-truth estimate or any Bayesian inference.

A plot of the tree, and a lineages-through-time plot, are available in Supplemental Figure 1. The Newick file is Canidae_ground_truth.newick (Supplemental Data).

Data. Morphological characters and dates came from the published matrix of Slater (2015) [5], specifically the Dryad repository [6] containing a NEXUS file with both morphology and MrBayes commands. Slater's matrix synthesized and updated the matrices published in the monographs by Wang and Tedford.

The data (characters and tip-dates) were left unchanged, except that no non-North American species were removed, and taxon "outgroup" was removed. Slater used last-occurrence dates for the tip-date of each species; for the purposes of tip-dating, this decision might be suboptimal, because a specimen bearing characters may sample from anywhere in a species' time-range. However, the nature of OTUs in tip-dating analyses is a complex question not yet addressed in the literature (Matzke and Irmis, this volume). Therefore, Slater's tip dates were retained for purposes of simplicity and direct comparability. Slater also used extensive node calibrations to represent the stratigraphic first occurrences of many taxa; these are reasonable given Slater's goal (fitting models of trait evolution), but node calibrations obscure the differences between tip-dating methods and so they were deleted from most analyses here.

MrBayes analyses. The "fossilized birth-death" (FBD) tree prior available in MrBayes starting with version 3.2.2 is closely analogous to the Birth-Death-Serial-Sampling (BDSS) tree prior available in Beast2, so we use BDSS to describe these runs. In MrBayes 3.2.5, sampling ancestors in BDSS became available

The 6 focal analyses presented in the main text are mb1 (Slater's original uniform node age prior analysis including node date calibrations, with some corrections), mb8 (uniform node age prior, no node dates, flat priors on clock parameters, uniform(45,100) prior on the root age), mb9 (mb8 but with SABD tree prior and flat priors on speciation, extinction, and sampling rate), and mb10 (mb9 but BDSS, i.e. disallowing sampled ancestors via the command "prset samplestrat = fossiltips;").

These focal analyses were selected out of a much larger collection of analyses (40 total) that were run while experimenting with modifications of the original

Slater (2015) NEXUS file. A summary of the variant inputs, and the results, is presented in Supplemental Table S2. Apart from the issues surrounding the uniform node age prior versus birth-death tree priors, which are the topic of the main text, Appendix 1 identifies other issues in the Slater (2015) NEXUS file, and the MrBayes implementation and documentation.

All MrBayes analyses ran for 5 million generations (sampling every 2500), with 2 runs, 4 chains, and default temperatures. The only exceptions were 3 analyses run with 4 runs and varying temperatures, aimed at improving the unsatisfactory topological result of the original mb9 run (see below). Typically this was sufficient for convergence as assessed in Tracer plots and MrBayes output statistics. For a few non-focal runs (Supplemental Table S2) it was not. Sometimes this was due to improper settings, and sometimes likely due to more fundamental issue with MrBayes's tree-searching abilities (notably in the case of BDSS analyses). As the purpose of many of the runs was to determine negative effects of certain versions, models, or settings (i.e. including problems with convergence), no effort was made to improve sampling further. The Perl scripts burntrees and catmb [7] were used to extract the last 50% of each tree sample and convert to dated trees.

Attempting to improve MrBayes SABD inference. The topological result of the original mb9 SABD run was unsatisfactory in that wild dogs () and the dhole () came out as a clade within the extinct Borophagines. It was suggested that adding more runs and varying temperatures might help. We attempted this by changing the number of runs from the default 2 to 4, and then running three analyses with different temperatures: 9a (4 runs, temperature=0.05), 9b (4 runs, temperature=0.5, the default), and 9c (4 runs, temperature=1). These runs are numbered 36a, 36b, and 36c in Supplemental Table S2.

Beast2 analyses. BEASTmasteR was used to construct the XML files for two Beast2 tip-dating analyses; the R scripts and Excel settings files are available in SM. The first analysis (r1) used a BDSS tree prior [8]; the second (r2) used SABD [9, 10]. Flat priors were used for each major parameter (mean and SD of the lognormal relaxed clock; and birth, death, and serial sampling rates). Rho (proportion of living species sampled) was fixed to 1, as it is not statistically identifiable if left free [11]. All runs in both programs used a single morphology partition, with an M*k*v model correcting for the ascertainment bias against invariant characters [12, 13] and gamma-distributed rate variation with 4 rate categories.

The Beast2 analyses were run for 50 million generations (sampling every 25000). For all runs, TreeAnnotator was used to choose the Maximum Clade Credibility (MCC) tree and calculate node-date HPDs (95% highest posterior densities) and bipartition posterior probabilities (PP). Burntrees [7] was used to process MrBayes outputs for input into TreeAnnotator.

BEASTmasteR [14] and custom R scripts (Supplemental Data) were used to plot all MrBayes and Beast2 MCC trees, as well as the trace plots for all key parameters (SI), and to extract parameter and node-date estimates of interest.

To assess the overall estimate of topology in each analysis, the topological symmetric distance (treedist function; phangorn R package; [15] between the dated MCC tree and the MCC tree from the undated (mb2) analysis was calculated and compared to the distribution of distances between trees in the mb2 post-burnin treecloud [16].

For the six focal analyses, an additional test was run to see how well the methods could predict selected tip dates [17]. Four tips were selected from across the tree (*Canis ferox*, 3.5 Ma; *Epicyon haydeni*, 5.3 Ma; *Leptocyon gregorii*, 23 Ma; *Hesperocyon gregarious*, 30.8 Ma). For each tip and focal analysis, the settings file was modified to change the tip date to a uniform(0,100) prior. After the MCMC run, the sampled tip date was extracted from the post-burnin tree sample and plotted as a histogram.

Supplemental Results

Rate parameters. Estimates of rate parameters in the focal analyses are consistent with the dating results, in that analyses with the youngest node age estimates have the highest clock, speciation, and sampling rates. The parameter describing the mean of relaxed clock branchwise rate variation (IGRvar for MrBayes, clockSD for Beast2) was inferred with similar precision across all analyses, despite uninformative priors, suggesting no special effort is needed to determine the prior for clock relaxation (clock models with autocorrelated rates may be different; [18]). The uncorrected Slater analysis does show the effect of the IGRvar prior used there (the intention was a diffuse prior, but the setting used forced a close-to-strict clock instead; see Appendix 1); however, the downstream effect on the analysis was minimal.

Summary results of all 40 variant analyses are presented in Supplemental Table S2 (supplemental Excel file). Trace plots of key variables for all 40 analyses are available in Supplemental Data (file *Canidae_traceLogs.pdf*). Plots of the MCC trees for all 40 analyses are also available (*Canidae_treeLogs.pdf*).

MrBayes SABD topology and convergence issues. The disagreement between the focal MrBayes BDSS and SABD (m10 and m9) analyses about the position of the wild dogs (*Lycaon pictus*) and dhole (*Cuon javanicus*) is puzzling, because both analyses place this clade with a sister with posterior probability 1.0; they just disagree on whether that sister is *Xenocyon texanus* (in Caninae) or *Desmocyon thomsoni* (in the extinct Borophaginae). It is true that in undated analyses (mb2), the clade has a relatively long morphological branch, and this may make it more difficult to place; however, the fact that Beast2 analyses under both BDSS and SABD place the clade unambiguously with *Xenocyon* suggest the primary issue may lie with MrBayes's implementation of the MCMC search for SABD. Design of the MCMC operators used to search tree and parameter space is very complex, and sampling ancestors adds another later of complexity, e.g. because the tree may be changing dimensionality as fossils switch from being side-branches to direct ancestors. Bayesian estimation of direct ancestry is a very new (and potentially revolutionary) development in paleontology and phylogenetics.

Variant MrBayes mb9 analyses with 4 runs and altered temperatures did improve the situation for MrBayes SABD in some respects. Both mb9a and mb9b (4 runs each, temperatures 0.05 and the default 0.5, respectively) did retrieve *Lycaon+Cuon* inside the Caninae. The analysis mb9c (temperature 1) did not, again placing this clade in the Borophagines. However, even the more successful MrBayes SABD analyses still exhibited difficulties (Supplemental Table S2). Convergence failed for mb9a and mb9c, with the standard deviation of split frequencies never declining below 0.1(0.17 and 0.19, respectively). For mb9b, this metric was better, but still marginal (0.05). For comparison, mb10 (MrBayes BDSS) achieved 0.026 in the same number of generations.

In addition, the percentage Robinson-Foulds topological difference from the undated MrBayes analysis (mb2) also indicated issues. The average within-mb2 difference was 20.8%. The average difference between r1 and mb2 was 36.2%, and between r2 and mb2 it was 33.8%. For mb10, the difference was 33.1%. For mb9, mb9a, mb9b, and mb9c, by contrast, the differences were 56.2%, 43.8%, 41.5%, and 56.9%, respectively.

It is encouraging that the Beast2 implementation inferred the conventional relationship for *Lycaon+Cuon*, and also that it estimated several *Leptocyon* OTUs as direct ancestors of Caninae, confirming the conclusions of Tedford et al. (2009).

Supplemental References

[1] Matzke, N.J. 2013 TreeRogue: R code for digitizing trees. https://stat.ethz.ch/pipermail/r-sig-phylo/2010-October/000816.html

[2] Wang, X. 1994 Phylogenetic systematics of the Hesperocyoninae (Carnivora, Canidae). *Bulletin of the American Museum of Natural History* **221**, 1-207.

[3] Wang, X.T., Richard H.; Taylor, Beryl E. 1999 Phylogenetic systematics of the Borophaginae. *Bulletin of the American Museum of Natural History* **243**, 1-391.

[4] Tedford, R.H.W., Xiaoming; Taylor, Beryl E. 2009 Phylogenetic systematics of the North American fossil Caninae (Carnivora, Canidae). *Bulletin of the American Museum of Natural History* **325**, 1-218.

[5] Slater, G.J. 2015 Iterative adaptive radiations of fossil canids show no evidence for diversity-dependent trait evolution. *Proceedings of the National Academy of Sciences* **112**, 4897-4902. (doi:10.1073/pnas.1403666111).

[6] Slater, G.J. 2015. Data from: Iterative adaptive radiations of fossil canids show no evidence for diversity-dependent trait evolution. Dryad. Accessed May 1, 2015. http://dx.doi.org/10.5061/dryad.9qd51

[7] Nylander, J.A.A. 2014. Burntrees v.0.2.2.: Perl script for manipulating MrBayes tree and parameter files. Accessed May 1, 2015. https://github.com/nylander/Burntrees

[8] Stadler, T. 2010 Sampling-through-time in birth-death trees. *Journal of Theoretical Biology* **267**, 396-404. (doi:10.1016/j.jtbi.2010.09.010).

[9] Gavryushkina, A., Welch, D., Stadler, T. & 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).

[10] Heath, T.A., Huelsenbeck, J.P. & Stadler, T. 2014 The fossilized birth–death process for coherent calibration of divergence-time estimates. *Proceedings of the National Academy of Sciences* **111**, E2957-E2966. (doi:10.1073/pnas.1319091111).

[11] Stadler, T. & Yang, Z. 2013 Dating phylogenies with sequentially sampled tips. *Systematic Biology* **62**, 674-688. (doi:10.1093/sysbio/syt030).

[12] Lewis, P.O. 2001 A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology* **50**, 913-925. (doi:10.1080/106351501753462876).

[13] Felsenstein, J. 1992 Phylogenies from restriction sites: a maximum-likelihood approach. *Evolution* **46**, 159-173. (doi:10.2307/2409811).

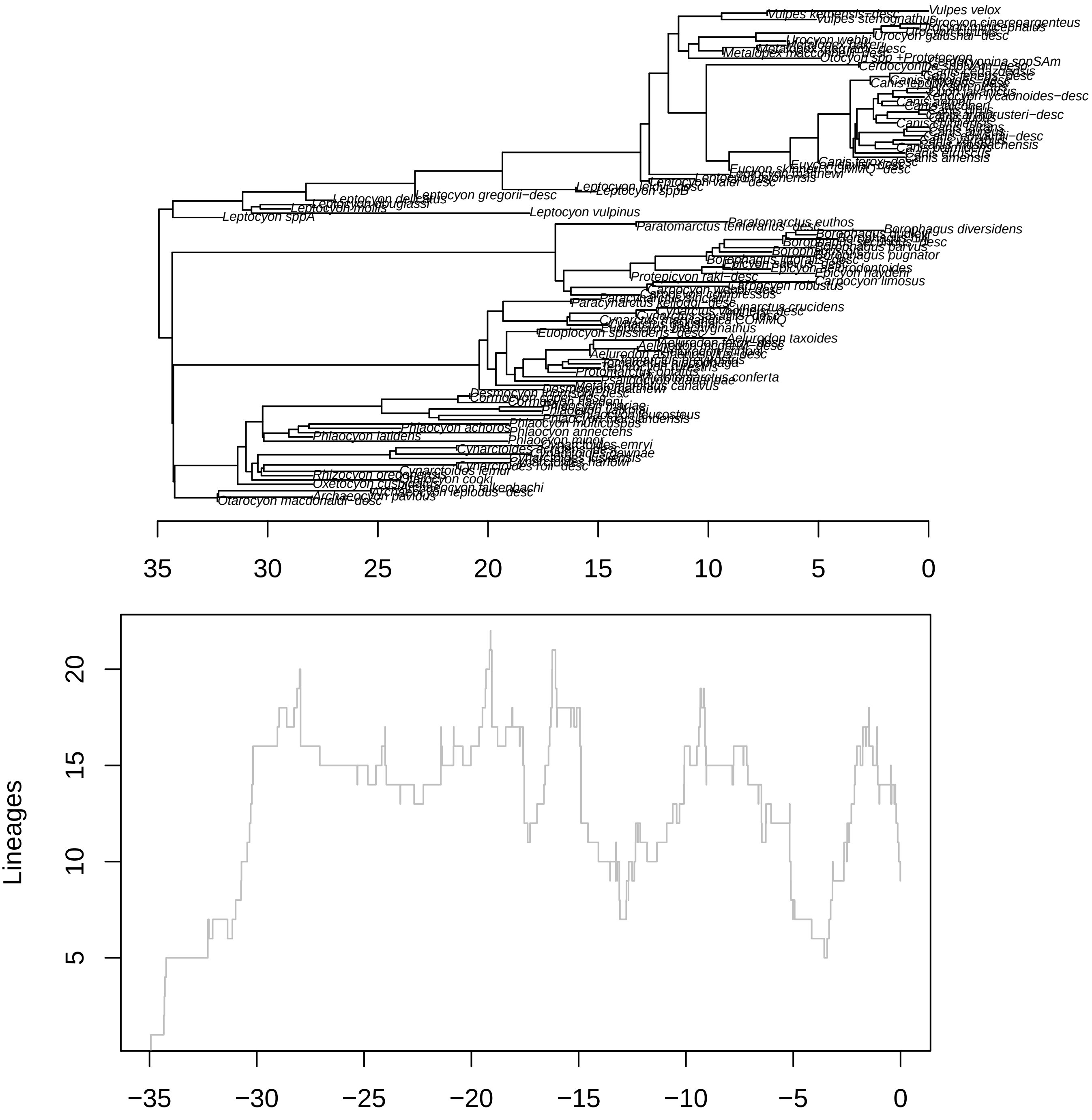
[14] Matzke, N.J. 2015 BEASTmasteR: automated conversion of NEXUS data to BEAST2 XML format, for fossil tip-dating and other uses. PhyloWiki. http://phylo.wikidot.com/beastmaster

[15] Schliep, K.P. 2011 phangorn: phylogenetic analysis in R. *Bioinformatics* **27**, 592-593. (doi:10.1093/bioinformatics/btq706).

[16] Matzke, N.J., Shih, P.M. & Kerfeld, C.A. 2014 Bayesian analysis of congruence of core genes in *Prochlorococcus* and *Synechococcus* and implications on horizontal gene transfer. *PLoS ONE* **9**, e85103. (doi:10.1371/journal.pone.0085103).

[17] Drummond, A.J. & Stadler, T. 2016 Bayesian phylogenetic estimation of fossil ages. *arXiv*, 1-23. (doi:arXiv:1601.07447v1).

[18] Ronquist, F., Klopfstein, S., Vilhelmsen, L., Schulmeister, S., Murray, D.L. & Rasnitsyn, A.P. 2012 A total-evidence approach to dating with fossils, applied to the early radiation of the Hymenoptera. *Systematic Biology* **61**, 973-999. (doi:10.1093/sysbio/sys058).



Supplemental Figure 1

Table S1. Five Bayesian tip-dating runs are compared to (column 1) the conclusions of Tedford & Wang. The two Slater analyses (original, and a modification repairing some unintended issues; cols. 2-3) represent MrBayes analyses under a uniform node age prior, constrained by both tip dates and many node-date priors. The third run (col. 4) shows the effect of removing the node-age calibrations, and putting flat priors on the parameters for clock rate and variation. The fourth run (col. 5) shows the drastic effect of switching to a fossilized birth-death prior. Columns 5 and 6 shown Beast2 runs with flat clock priors and BDSS and SABD tree priors, respectively. These five runs are drawn from the 40 analyses shown in Supplemental Table 1.

Estimates	**Analysis**	**Tedford & Wang, manually time-scaled cladogram**	**mb1_orig: Slater (2015) MrBayes analysis, original**	**mb1: Slater (2015) MrBayes analysis, corrected**	**mb8: MrBayes analysis, unconstrained, uniform node age prior**	**mb10: MrBayes analysis, unconstrained, fossilized-BDSS node age prior**	**mb9: MrBayes analysis, unconstrained, fossilized-SABD node age prior**	**r1: Beast2, unconstrained, BDSS tree prior**	**r2: Beast2, unconstrained, SABD tree prior**	**Notes**
node ages	age(crown *Canis*)	3.2	7.3 (5.2,9.8)	7.4 (5.1,9.6)	27.5 (16.9,39.1)	4.4 (3.3,5.5)	37.1 (35.5,41.0)	3.7 (2.9,4.8)	2.7 (2.1,3.3)	4,6
	age(crown Caninae)	11.7	21.4 (17.9,25.8)	21.2 (17.8,25.5)	38.9 (29.0,54.0)	12.1 (10.7,13.7)	37.1 (35.5,41.0)	12.4 (10.9,14.1)	10.6 (10.3,12.0)	3,5
	age(Canidae)	40-36	41.0 (40.4,41.8)	41.0 (40.4,43.3)	49.0 (44.4,66.8)	38.3 (36.0,41.3)	37.1 (35.5,41.0)	36.8 (35.0,39.0)	36.1 (34.2,38.0)	2
	prior on root age	-	gamma(1,1) expect: 1 (0.024, 3.6)	offsetexp(45, 50) expect: 50 (45.1,63.6)	uniform(45,100)			none		
	age(root)	41.5	43.4 (41.2,45.5)	47.1 (45.0,52.0)	49.5 (45.0,67.7)	45.3 (45.0,46.4)	45.3 (45.0,46.5)	40.6 (40.0,42.1)	40.0 (40.0,41.2)	1
topology	Percent toplogical distance to undated MrBayes tree (mean betw. undated trees = 31.0% +/- 8.3%)		33.8%	32.3%	40.8%	33.1%	56.2%	36.2%	33.8%	
	crown Caninae monophyletic?	y	y	y	y	y	n	y	y	
	((Can., Boro.), Hesp.)?	y	n	n	y	y	n	y	y	
clock	model	-	IGR	IGR	IGR	IGR	IGR	ucld	ucld	
	prior	-	lognorm(-6,0.1) expectation: 0.0025 (0.002,0.003)		truncated normal(0.0025,0.1) expectation: 0.08 (0.003,0.23)			uniform(0,10)		
	clock rate	-	0.00309 (0.00201,0.00438)	0.0094 (0.0071,0.012)	0.0045 (0.0025,0.006)	0.019 (0.015,0.024)	0.024 (0.019,0.030)	0.038 (0.025,0.053)	0.052 (0.033,0.075)	
	variation prior	-	exp(126.887); expectation: 0.0079 (0.00019,0.029)	uniform(0.0001,200)				uniform(0,10)		
	among-branch variation parameter	-	0.00955 (0.00479,0.0154)	0.028 (0.018,0.039)	0.034 (0.021,0.046)	0.024 (0.015,0.033)	0.029 (0.018,0.043)	1.20 (0.98,1.42)	1.24 (1.03,1.47)	7
	among-site variation gamma parameter	-	1.03 (0.19, 1.92)	1.45 (0.87,2.02)	1.44 (0.86, 2.06)	1.38 (0.87,2.0)	1.335 (0.78,1.88)	1.16 (0.66,1.70)	1.14 (0.61,1.62)	
tip dates	*Canis ferox*	3.5	fixed(3.5)	4.4 (0.2,9.3)	10.4 (0.2,32.8)	2.8 (0.5,5.1)	3.9 (2.7,4.8)	2.9 (0.5,5.2)	2.3 (0.4,4.2)	8,9
	Epicyon haydeni	10-5.3	fixed(5.332)	3.5 (0.4,7.9)	4.7 (0.2,17.1)	4.9 (1.3,7.4)	4.1 (1.8,5.8)	5.4 (1.9,8.8)	5.0 (2.1,7.9)	
	Leptocyon gregorii	24.4-23	fixed(23)	23.3 (15.4,28.5)	32.3 (14.9,49.5)	21.8 (17.3,26.2)	23.6 (21.3,26.5)	20.1 (14.3,26.7)	15.4 (12,20.4)	
	Hesperocyon gregarius	37.2-30.8	fixed(30.8)	35.7 (28,40.5)	32.2 (11.4,43.7)	33.7 (30.4,36.2)	37 (30.9,43.0)	33.3 (30.1,35.7)	33.0 (30.3,34.8)	
tree	prior (all 3)	-	-	-	-	unif(0,10)	unif(0,10)	unif(0,10)	unif(0,10)	
	speciation	-	-	-	-	0.37 (0.041,0.79)	0.48 (0.043,0.96)	0.48 (0.29,0.73)	0.65 (0.36,1.08)	
	extinction	-	-	-	-	0.33 (0.037,0.71)	0.43 (0.39,0.48)	0.187 (0,0.53)	0.33 (0,0.90)	
	sampling	-	-	-	-	0.29 (0.033,0.63)	0.13 (0.08,0.18)	0.27 (0.16,0.38)	0.30 (0.15,0.44)	
Suppl. Table	run #	-	3	31	35	37	36	1	2	
	code	-	mb3.2.5, mb1_orig	mb3.2.5, mb1	mb3.2.5, mb8	mb3.2.5, mb10	mb3.2.5, mb9	r1	r2	

Notes

1 Wang & Tedford (2008) date from Fig. 6.6, Arctoidea outgroup
40 Ma: Wang & Tedford (2008), Fig. 6.6, origin of Hesperocyoninae;

2 36 Ma: Tedford & Wang (2008), Fig. 7.1, divergence of Hesperocyoninae and Borophaginae+Caninae

3 Tedford et al. (2009), Fig. 66

4 Tedford et al. (2009), Fig. 66

5 For mb9, excluding Cuon/Lycaon, crown Caninae date is 11.6 (10.3,12.8)

6 For mb9, excluding Cuon/Lycaon, crown *Canis* date is 5.9 (5.0,7.1)

7 The branch-rate variation parameters are not equivalent between MrBayes IGR and Beast2 ucld relaxed clocks.

8 The time ranges in column 1 are the stratigraphic range of each species as indicated in Fig. 65 of Wang (1994; Hesperocyon), Fig. 141 of Wang et al. (1999; Epicyon), and Figure 66 of Tedford et al. (2009; Canis and Leptocyon).

9 The fixed() statements in column 2 are the tip ages used by Slater (2015) for these taxa. To estimate the tip date (next 4 columns), these were changed to uniform(0,100).

Supplemental Table S2. Summary settings and results of all 40 tip-dating runs.

Working directory	num	SD split frequencies	MrBayes convergence warning	Average ESS for Tree Length (TL) {1}	arithmetic mean marginal LnL	harmonic mean marginal LnL	orig_code	Program	Focal analyses name in main text	Run	title	Slater node age calibrations used?	Tree Prior (MrBayes's fossilized birth-death is similar to Beast2 BDSS; when ancestors are sampled, similar to Beast2 SABD)	NEXUS tree/root age prior	Reported tree/root age prior	Expected mean of actual used tree/root age prior (in my) {2}	estimated root age	SD root age	coefficent of variation in root age	mean CV node ages	mean CV non-constrained node ages	mean posterior prob	mean PP uncalibrated nodes	mean PP calibrated nodes	branch-length RF distance to undated RF tree	topological RF distance to undated MrBayes MCC tree (max=260)	Percentage RF difference to MrBayes MCC tree (max=260)
/d	1	NA	NA	705	-2891.8	nr	r1	Beast2.1.3	r1	r1	Beast2.1.3_r1	n	BDSS	unif	unif	flat	40.54	0.531	0.013	0.079	0.079	0.739	0.702	0.961	0.075	94	36.2%
/d	2	NA	NA	1801	-2894.1	nr	r2	Beast2.1.3	r2	r2	Beast2.1.3_r2	n	SABD	unif	unif	flat	40	0.296	0.007	0.068	0.068	0.769	0.736	0.977	0.086	88	33.8%
/d	3	0.01702	n	53.47	-2868.2	-3036	mb1	3.2.5		mb1_orig	mb3.2.5_mb1_orig	y	unif	offsete	gamma	1	43.38	1.133	0.026	0.108	0.115	0.687	0.635	1	0.061	88	33.8%
/d	4	0.02375	n	19.1	-2823	-2892.3	mb2	3.2.5		mb2_orig	mb3.2.5_mb2_orig	NA	nonclock	NA	NA	NA	1.667	1.051	0.631	0.711	0.704	0.707	0.662	0.998	0.085	54	20.8%
/d	5	0.02579	n	729.68	-2827.1	-2891.4	mb2_mf	3.2.5		mb2_mfixed	mb3.2.5_mb2_mfix	NA	nonclock	NA	NA	NA	0.935	0.11	0.117	0.283	0.296	0.704	0.657	0.997	0.054	64	24.6%
/d	6	0.11341	n	16.51	-2933.4	-3263.5	mb3	3.2.5		mb3_orig	mb3.2.5_mb3_orig	y	fossil (BDSS)	offsete	gamma	1	41.87	0.847	0.02	0.049	0.053	0.732	0.687	1	0.081	112	43.1%
/d	7	0.13779	n	16.36	-2928.3	-3210.5	mb4	3.2.5		mb4_orig	mb3.2.5_mb4_orig	y	fossil (BDSS)	offsete	gamma	1	41.38	0.76	0.018	0.061	0.06	0.733	0.688	1	0.082	128	49.2%
/d	8	0.21449	y	13.72	-2947.3	-3145.5	mb5	3.2.5		mb5_orig	mb3.2.5_mb5_orig	n	fossil (BDSS)	unif(45	gamma	flat	40.45	0.301	0.007	0.032	0.032	0.615	0.584	0.844	0.087	130	50.0%
/d	9	0.02561	n	34.76	-2874.6	-3142.5	mb6	3.2.5		mb6_orig	mb3.2.5_mb6_orig	n	unif	unif(45	gamma	flat	40.56	0.26	0.006	0.122	0.122	0.693	0.646	0.969	0.062	80	30.8%
/d	10	0.0209	n	32.76	-2870.4	-3047	mb7	3.2.5		mb7_orig	mb3.2.5_mb7_orig	n	unif	unif(45	gamma	flat	40.6	0.301	0.007	0.121	0.122	0.682	0.634	0.966	0.061	82	31.5%
/d	11	0.02535	n	12.8	-2856.9	-2972.8	mb8	3.2.5		mb8_orig	mb3.2.5_mb8_orig	n	unif	unif(45	gamma	flat	40.31	0.199	0.005	0.09	0.091	0.682	0.633	0.976	0.062	78	30.0%
/d	12	0.15611	y	7.98	-2868.4	-2986.1	mb9	3.2.5		mb9_orig	mb3.2.5_mb9_orig	n	fossil (BDSS)	unif(45	gamma	flat	40.45	0.247	0.006	0.056	0.061	0.658	0.622	0.889	0.079	108	41.5%
/d	13	0.02032	n	112.15	-2872.8	-2973.1	mb1	3.2.5		mb1_root_unif	mb3.2.5_mb1_root	y	unif	unif(45	unif(45	1	43.6	1.482	0.034	0.11	0.118	0.698	0.647	1	0.06	88	33.8%
/d	14	0.01937	n	80.05	-2879	-2970.1	mb1_fix	3.2.5		mb1_fixed_root	mb3.2.5_mb1_fixed	y	unif	unif(45	unif(45	1	48.29	2.446	0.051	0.109	0.115	0.692	0.64	1	0.063	76	29.2%
/d	15	0.02364	n	619.39	-2821.6	-2886.2	mb2	3.2.5		mb2_root_unif	mb3.2.5_mb2_root	NA	nonclock	NA	NA	NA	0.924	0.117	0.127	0.392	0.573	0.673	0.622	0.995	0.022	52	20.0%
/d	16	0.07696	n	231.35	-2942.4	-3061.8	mb3	3.2.5		mb3_root_unif	mb3.2.5_mb3_root	y	fossil (BDSS)	unif(45	unif(45	1	45.61	0.625	0.014	0.061	0.063	0.843	0.817	1	0.077	96	36.9%
/d	17	0.12259	y	61.42	-2944.8	-3047.7	mb4	3.2.5		mb4_root_unif	mb3.2.5_mb4_root	y	fossil (BDSS)	unif(45	unif(45	1	45.71	0.5	0.011	0.064	0.065	0.752	0.71	1	0.079	122	46.9%
/d	18	0.29434	y	139.4	-2992.5	-3075.3	mb5	3.2.5		mb5_root_unif	mb3.2.5_mb5_root	n	fossil (BDSS)	unif(45	unif(45	flat	45.52	0.518	0.011	0.063	0.067	0.608	0.579	0.854	0.092	160	61.5%
/d	19	0.02602	n	185.65	-2877.6	-2971.2	mb6	3.2.5		mb6_root_unif	mb3.2.5_mb6_root	n	unif	unif(45	unif(45	flat	45.75	0.566	0.012	0.134	0.132	0.663	0.613	0.984	0.06	90	34.6%
/d	20	0.03121	n	116.62	-2876.3	-2980.1	mb7	3.2.5		mb7_root_unif	mb3.2.5_mb7_root	n	unif	unif(45	unif(45	flat	45.31	0.327	0.007	0.14	0.139	0.67	0.621	0.981	0.063	84	32.3%
/d	21	0.02328	n	220.3	-2875.2	-2962.9	mb8	3.2.5		mb8_root_unif	mb3.2.5_mb8_root	n	unif	unif(45	unif(45	flat	45.29	0.321	0.007	0.128	0.128	0.677	0.628	0.984	0.061	80	30.8%
/d	22	0.20434	y	209.42	-2907.7	-3019.6	mb9	3.2.5		mb9_root_unif	mb3.2.5_mb9_root	n	fossil (BDSS)	unif(45	unif(45	flat	45.37	0.378	0.008	0.055	0.058	0.832	0.812	0.999	0.093	148	56.9%
/d	23	0.02656	n	50.54	-2863.4	-2956	mb6	3.2.0		mb6	mb3.2.0_mb6	n	unif	unif(45	Expone	1	42.63	1.709	0.04	0.168	0.168	0.627	0.567	0.983	0.068	84	32.3%
/d	24	0.02106	n	42.18	-2854.8	-2961.7	mb7	3.2.0		mb7	mb3.2.0_mb7	n	unif	unif(45	Expone	1	42.96	2.056	0.048	0.224	0.232	0.614	0.552	0.984	0.068	102	39.2%
/d	25	0.02575	n	70.94	-2851.1	-2944.2	mb8	3.2.0		mb8	mb3.2.0_mb8	n	unif	unif(45	Expone	1	40.85	0.765	0.019	0.163	0.164	0.646	0.588	0.988	0.065	94	36.2%
/d	26	0.01992	n	57.62	-2862.8	-2960.7	mb5	3.2.2		mb5	mb3.2.2_mb5	n	fossil (BDSS)	unif(45	unif(45	flat	45.52	0.62	0.014	0.097	0.097	0.733	0.69	0.985	0.066	72	27.7%
/d	27	0.02354	n	40.29	-2855.2	-2978.8	mb6	3.2.2		mb6	mb3.2.2_mb6	n	unif	unif(45	unif(45	flat	57.6	6.452	0.112	0.211	0.211	0.614	0.553	0.981	0.066	106	40.8%
/d	28	0.02445	n	42.99	-2860.9	-2950.2	mb7	3.2.2		mb7	mb3.2.2_mb7	n	unif	unif(45	unif(45	flat	62.51	7.168	0.115	0.226	0.228	0.622	0.562	0.983	0.068	98	37.7%
/d	29	0.02333	n	112.41	-2862	-2953.3	mb8	3.2.2		mb8	mb3.2.2_mb8	n	unif	unif(45	unif(45	flat	48.61	4.156	0.085	0.207	0.207	0.603	0.541	0.993	0.066	98	37.7%
/d	30	0.02638	n	51.86	-2846.4	-2932.7	mb9	3.2.2		mb9	mb3.2.2_mb9	n	fossil (BDSS)	unif(45	unif(45	flat	45.4	0.413	0.009	0.084	0.085	0.738	0.698	0.975	0.073	84	32.3%
/d	31	0.02061	n	297.8	-2867.5	-2955.4	mb1	3.2.5	mb1	mb1	mb3.2.5_mb1	y	unif	offsetexp(45,5		50	47.14	1.791	0.038	0.109	0.116	0.701	0.651	1	0.062	84	32.3%
/d	32	0.24734	y	114.53	-2956.3	-3099.8	mb5	3.2.5		mb5	mb3.2.5_mb5	n	fossil (BDSS)	unif(45	unif(45	flat	45.48	0.513	0.011	0.061	0.065	0.684	0.646	0.974	0.091	156	60.0%
/d	33	0.0253	n	26.35	-2890.7	-3014.6	mb6	3.2.5		mb6	mb3.2.5_mb6	n	unif	unif(45	unif(45	flat	69.93	7.61	0.109	0.212	0.212	0.624	0.564	0.985	0.063	114	43.8%
/d	34	0.02943	n	41.03	-2892.3	-2993.1	mb7	3.2.5		mb7	mb3.2.5_mb7	n	unif	unif(45	unif(45	flat	70.6	9.531	0.135	0.222	0.222	0.649	0.592	0.988	0.066	106	40.8%
/d	35	0.0214	n	212.33	-2887.3	-2984.5	mb8	3.2.5	mb8	mb8	mb3.2.5_mb8	n	unif	unif(45	unif(45	flat	49.52	5.778	0.117	0.199	0.199	0.647	0.592	0.979	0.062	106	40.8%
/d	36	0.0573	n	247.13	-2927.4	-3001.6	mb9	3.2.5	mb9	mb9	mb3.2.5_mb9	n	fossil (SABD)	unif(45	unif(45	flat	45.36	0.362	0.008	0.041	0.042	0.786	0.761	1	0.09	146	56.2%
/d	37	0.02651	n	239.27	-2872.6	-2946.5	mb10	3.2.5	mb10	mb10	mb3.2.5_mb10	n	fossil (BDSS)	unif(45	unif(45	flat	45.34	0.36	0.008	0.078	0.077	0.744	0.705	0.994	0.071	86	33.1%

(Note: These three analyses were added after pre-publication reviews suggested altering the number of runs and/or temperature to improve MrBayes search/convergence in the MrBayes 3.2.5 mb9 (BDSS) run, which tended to place wild dogs in the exti

Working directory	num	SD split frequencies	MrBayes convergence warning	Average ESS for Tree Length (TL) {1}	arithmetic mean marginal LnL	harmonic mean marginal LnL	orig_code	Program	Focal analyses name in main text	Run	title	Slater node age calibrations used?	Tree Prior (MrBayes's fossilized birth-death is similar to Beast2 BDSS; when ancestors are sampled, similar to Beast2 SABD)	NEXUS tree/root age prior	Reported tree/root age prior	Expected mean of actual used tree/root age prior (in my) {2}	estimated root age	SD root age	coefficent of variation in root age	mean CV node ages	mean CV non-constrained node ages	mean posterior prob	mean PP uncalibrated nodes	mean PP calibrated nodes	branch-length RF distance to undated RF tree	topological RF distance to undated MrBayes MCC tree (max=260)	Percentage RF difference to MrBayes MCC tree (max=260)
/d	36a	0.16942	y	183.54	-2899.8	-3037.9	mb9a	3.2.5	mb9a	mb9a	mb3.2.5_mb9a		fossil (SABD)	unif(45	unif(45	flat	45.35	0.334	0.007	0.045	0.047	0.918	0.908	0.978	0.086	114	43.8%
/d	36b	0.05277	n	236.64	-2903.1	-3008.1	mb9b	3.2.5	mb9b	mb9b	mb3.2.5_mb9b		fossil (SABD)	unif(45	unif(45	flat	45.34	0.385	0.008	0.053	0.047	1.009	1.014	0.977	0.087	108	41.5%
/d	36c	0.18948	y	221.65	-2917.6	-3012.7	mb9c	3.2.5	mb9c	mb9c	mb3.2.5_mb9c		fossil (SABD)	unif(45	unif(45	flat	45.37	0.39	0.009	0.062	0.066	0.824	0.803	0.999	0.093	148	56.9%

Clock Rate Prior	Expected mean of clock rate prior (in expected changes/my)	Estimated clock rate	SD of estimated clock rate	Relaxed clock model	Prior on variance of branch rate variability (relaxed clock)	Expected mean of branch rate variance parameter	Estimated mean of branch rate variance parameter	SD of estimated mean of branch rate variance parameter	Relative rate setting for morphology partition	Crenelation observed in trace plots	outgroup setting	ingroup (Canidae) constrained to be monophyletic?	outgroup in correct position?	Wild dog+Dhole in correct position?	Modifications beyond 9 base models	**Base model**
unif	flat	0.038	0.008	UCLD	unif	flat	1.202	0.11399	fixed	n	none	y	y	y		BEASTmasteR setup BDSS model, uniform priors, no node calibrations
unif	flat	0.053	0.011	UCLD	unif	flat	1.236	0.11382	fixed	n	none	y	y	y		BEASTmasteR setup SABD model, uniform priors, no node calibrations
lognorr	0.0025	0.003	8E-04	IGR	exp(12	0.0079	0.01	0.00321	variabl	cor {3}	y	y	y	y		Slater original dating analysis; changed only to include non-North America living species
NA	NA	NA	NA	NA	NA	NA	NA	NA	variabl	y	y	y	y	y		Slater original non-dating analysis; changed only to include non-North America living species
NA	NA	NA	NA	NA	NA	NA	NA	NA	fixed	n	y	n	y	y		mb2, but relrate param m fixed to 1
lognorr	0.0025	0.003	1E-03	IGR	exp(12	0.0079	0.007	0.00246	variabl	y	y	y	y	y		mb1 but with fossilized-BD tree prior
lognorr	0.0025	0.003	9E-04	IGR	exp(12	0.0079	0.007	0.00266	variabl	y	y	y	y	y		mb3 but with priors on tree parameters explicitly set
lognorr	0.0025	0.004	0.001	IGR	exp(12	0.0079	0.005	0.00252	variabl	y	y	n	y	y		mb4 but with node calibrations removed, root date prior set to uniform(45,100)
lognorr	0.0025	0.003	8E-04	IGR	exp(12	0.0079	0.009	0.003	variabl	y	y	n	n?	y		mb5; node calibrations removed, root date prior set to uniform(45,100); tree prior back to uniform
lognorr	0.0025	0.003	8E-04	IGR	uniforn	flat	0.01	0.00352	variabl	y	y	n	n?	y		mb5; node calibrations removed, root date prior set to uniform(45,100), IGRprior set to igrvarpr=uniform(0.0001, 200)
flattish	flat	0.373	0.554	IGR	uniforn	flat	0.856	1.34177	variabl	y	y	n	n?	y		m7; vague prior on clockrate
flattish	flat	0.493	0.433	IGR	uniforn	flat	0.578	0.51918	variabl	y	y	n	y	y		m8; but with BDSS
lognorr	0.0025	0.004	3E-04	IGR	exp(12	0.0079	0.013	0.00206	variabl	cor	y	y	y	y	add treeagepr=uniform(45,50) to force non-gamma(1,1) tree age prior	
lognorr	0.0025	0.004	3E-04	IGR	exp(12	0.0079	0.013	0.00235	fixed	n	y	y	y	y	"	
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	y	NA	
lognorr	0.0025	0.004	4E-04	IGR	exp(12	0.0079	0.009	0.00239	variabl	y	y	y	y	y	"	
lognorr	0.0025	0.004	4E-04	IGR	exp(12	0.0079	0.008	0.00131	variabl	y	y	y	y	y	"	
lognorr	0.0025	0.005	3E-04	IGR	exp(12	0.0079	0.008	0.00111	variabl	y	y	n	y	**n**	"	
lognorr	0.0025	0.004	3E-04	IGR	exp(12	0.0079	0.015	0.00249	variabl	y	y	n	n	y	"	
lognorr	0.0025	0.004	3E-04	IGR	uniforn	flat	0.015	0.0027	variabl	y	y	n	n	y	"	
flattish	flat	0.009	0.001	IGR	uniforn	flat	0.027	0.00561	variabl	y	y	n	n	y	"	
flattish	flat	0.026	0.003	IGR	uniforn	flat	0.034	0.00595	variabl	y	y	n	y	**n**	"	
lognorr	0.0025	0.003	3E-04	IGR	exp(12	0.0079	0.04	0.00717	fixed	n	y	y	y	y	offsetexp(45,50) treeagepr, m fixed, ingroup monophyletic	
lognorr	0.0025	0.003	3E-04	IGR	uniforn	flat	0.052	0.01145	fixed	n	y	y	y	y	"	
flattish	flat	0.006	9E-04	IGR	uniforn	flat	0.067	0.015	fixed	n	y	y	y	y	"	
lognorr	0.0025	0.004	4E-04	IGR	exp(12	0.0079	0.011	0.00262	fixed	n	y	y	y	y	unif treeagepr, m fixed, ingroup monophyletic	
lognorr	0.0025	0.003	2E-04	IGR	exp(12	0.0079	0.047	0.00914	fixed	n	y	y	y	y	"	
lognorr	0.0025	0.003	2E-04	IGR	uniforn	flat	0.063	0.01555	fixed	n	y	y	y	y	"	
flattish	flat	0.004	7E-04	IGR	uniforn	flat	0.065	0.01447	fixed	n	y	y	y	y	"	
flattish	flat	0.02	0.002	IGR	uniforn	flat	0.042	0.00883	fixed	n	y	y	y	y	"	
flattish	flat	0.009	0.001				0.028	0.00519						y		
lognorr	0.0025	0.005	3E-04	IGR	exp(12	0.0079	0.007	0.00112	fixed	n	y	y	y	**n**	"	
lognorr	0.0025	0.003	2E-04	IGR	exp(12	0.0079	0.029	0.00435	fixed	n	y	y	y	y	"	
lognorr	0.0025	0.003	3E-04	IGR	uniforn	flat	0.033	0.00557	fixed	n	y	y	y	y	"	
flattish	flat	0.005	9E-04	IGR	uniforn	flat	0.035	0.00653	fixed	n	y	y	y	y	"	
flattish	flat	0.024	0.003	IGR	uniforn	flat	0.03	0.006	fixed	n	y	y	y	**n**	"	prset samplestrat = random; [SABD tree prior]
flattish	flat	0.019	0.002	IGR	uniforn	flat	0.024	0.00473	fixed	n	y	y	y	y	"	mb9 but prset samplestrat = fossiltips; (no sampled ancestors)

nct Borophagines.)

flattish	flat	0.026	0.003	IGR	uniforn	flat	0.036	0.00701	fixed	n	y	y	y	y	unif treeagepr, m fixed, ingroup monophyletic; changed to 4 runs, same default temperature (0.5)
flattish	flat	0.025	0.003	IGR	uniforn	flat	0.032	0.00627	fixed	n	y	y	y	y	unif treeagepr, m fixed, ingroup monophyletic; changed to 4 runs, lower temperature (0.05)
flattish	flat	0.028	0.004	IGR	uniforn	flat	0.035	0.00709	fixed	n	y	y	y	**n**	unif treeagepr, m fixed, ingroup monophyletic; changed to 4 runs, higher temperature (1)

Appendix 1: Issues with the MrBayes dating analysis of Slater (2015).

In setting up variant MrBayes analyses (Supplemental Table S2), a number of issues became apparent with the NEXUS file of the original Slater (2015) analysis. These are detailed below in order to help aid future MrBayes analyses, and in some cases to suggest improvements in the MrBayes code or documentation. These issues do not appear to greatly alter the dating results of Slater (2015), due to the large number of tip- and node-date constraints in that analysis (compare Figure 1b: mb2.3.5_mb1_orig; and Figure 1c: mb3.2.5_mb1), but they did cause major issues for analyses without node-date constraints.

The example NEXUS file being examined is canidae.nex, downloaded May 2015, and re-downloaded (unchanged) in April 2016 from: http://datadryad.org/bitstream/handle/10255/dryad.73273/canidae.nex?sequence=1 .

A file correcting the issues identified below, but otherwise maintaining the intended analysis of Slater (2015) (uniform node age prior, node date constraints, etc.) is file "canidae_all_issues_fixed.nex", located in directory mb_3.2.5b_add_ingroup/mb1/ of the Supplemental Data file "Matzke_Wright_SuppData.zip."

Issue 1: Root node date calibration

The NEXUS file includes a variety of node-date calibrations, including an offsetexp(min=45, mean=50) calibration for the root node:

Line 433 of canidae.nex:

```
calibrate root=offsetexponential(45, 50); [mean = 50, median = 48.5, 95% upper = 60]
```

Unfortunately, this date prior on the root node appears to be ignored by MrBayes. This can be confirmed by inspecting Slater (2015)'s Figure S2, where the age of the root is approximately 42 Ma, despite the fact that the root node constraint has a hard minimum of 45 Ma.

The only hint that MrBayes is ignoring the root calibration is the following warning message:

```
================================================================
WARNING: Constraint 'root' refers only to deleted taxa
         and will be disregarded
================================================================
```

In the screen output of the MrBayes run, this warning is easy to miss, as it is hidden amongst many other warnings of this type:

```
=============================================================
WARNING: There is one character incompatible with the specified
         coding bias. This character will be excluded.
=============================================================
```

The second warning is due to a character in the data matrix being invariant. Another reason that the first warning can be missed is that the warning is inaccurate (the taxa were not deleted).

It appears that, with the root node date calibration ignored, and with no tree age prior (treeagePr setting) given, the MrBayes dating analyses default to a tree height prior with a gamma(1,1) distribution. The message output to the screen at runtime is:

```
=============================================================
Tree age has a Gamma(1.00,1.00) distribution
=============================================================
```

This distribution is equivalent to an exponential(1) distribution. This suggests that the real prior being used on the age of the root is informative, and has a mean of 1 mya. This provides a ready explanation of the young date of the root node – the root node date is estimated to be just below the oldest tip (the outgroup, fixed to 40 Ma).

Issue 2: Prior on clock rate variability

The NEXUS file includes this prior for the igrvar parameter (igrvar = variance parameter for the gamma distribution on branchwise rate variability, for independent branch rates).

Line 468 of canidae.nex:

```
prset igrvarpr=exp(126.887) ; [a vague prior]
```

This parameter describes the expected variance given a branchlength in expected amounts of change. igrvar is multiplied by the each branchlength to give the expected variability.

The comment suggests the Exponential(126.887) prior as "a vague prior". We can see how a user could think this, given the language in the igrvar documentation:

MrBayes > help prset (in MrBayes 3.2.5) gives:

```
  Igrvarpr -- This parameter allows you to specify a prior on the variance
              of the gamma distribution from which the branch lengths are
              drawn in the independent branch rate (IGR) relaxed clock
              model. Specifically, the parameter specifies the rate at
```

which the variance increases with respect to the base rate of the clock. If you have a branch of a length corresponding to 0.4 expected changes per site according to the base rate of the clock, and the igrvar parameter has a value of 2.0, then the effective branch length will be drawn from a distribution with a variance of 0.4*2.0.

You can set the parameter to a fixed value, or specify that it is drawn from an exponential or uniform distribution:

```
prset igrvarpr = fixed(<number>)
prset igrvarpr = exponential(<number>)
prset igrvarpr = uniform(<number>,<number>)
```

For backward compatibility, 'ibrvarpr' is allowed as a synonym of 'igrvarpr'.

(This text is also found in commref_mb3.2.txt in the MrBayes 3.2.x download)

However, elsewhere in MrBayes, the exponential distribution is generally interpreted such that the input parameter for exp() is the exponential rate parameter, λ, and the expectation of the mean is $\beta = 1/\lambda$. Thus, the expectation of an Exponential(126.887) distribution is 1/126.887=0.00788. Thus, instead of a vague prior on branchwise rate variation, this prior essentially mandates a strict clock.

Our interpretation is confirmed by examining the inference of the estimated mean of branch rate variance parameter under MrBayes runs where the igrvar parameter has been changed (Supplemental Table S2).

Issue 3. Relative rate prior (ratepr)

In the NEXUS file, the relative rate prior (ratepr) is set to "variable":

Line 458 of canidae.nex:

```
prset applyto=(all) ratepr = variable;
```

This setting creates a parameter, m{1}, representing the relative rate of the morphology partition compared to other partitions (DNA, RNA, etc.) under a common overall clock model. However, canidae.nex is a morphology-only dataset and only has 1 partition. MrBayes does not identify this situation and fix m{1} to "fixed". Instead, it attempts to estimate this relative rate along with the clock rate and clock variability. This creates poor mixing due to non-identifiability, and "crenelations" in the MCMC trace of parameters. Page 3 (analysis mb1_orig) of the Supplemental Data file Canidae_traceLogs.pdf shows these crenelations: the MCMC trace jumps to one value, samples around that value for a while, and then jumps to a much different value. Later in the chain, it discretely jumps towards the original value, and the cycle repeats. This behaviour leads to low ESS values and bimodal parameter estimates.

In Slater's highly-constrained original analysis, the effect on other inferences is not particularly noticeable and presumably makes little difference. However, it becomes a major issue for mixing and parameter estimation as node constraints are removed.

Issue 4. Rate prior on the morphological clock

In the NEXUS file, the relative rate prior (ratepr) is set to "variable":

```
Line 469 of canidae.nex:

prset clockratepr = lognorm(-6,0.1);
```

Slater set a tight prior on the morphology clock rate. The lognorm(-6, 0.1) distribution has a mean in real space of 0.0025 changes/my, and an SD of 0.00025. This is a user decision rather than a problem, and it is clearly mentioned in Slater (2015).

It may be, however, that the decision for a strongly informative prior on the clock rate was made in part in order to “make the analysis behave,” due to problems caused by the uniform node age prior, and perhaps some of the other issues mentioned in this appendix. We note that tip-dating analyses with BDSS-type tree models function very well even with broad, uninformative priors on the rate of the morphological clock (Supplemental Table S2).

Issue 5. Outgroup, and specifying the outgroup

The outgroup taxon, named "outgroup" in Slater's analysis, is identified as the outgroup in canidae.nex:

```
Line 459 of canidae.nex:

outgroup 1;
```

Taxon 1 is the outgroup OTU. However, in MrBayes dating analyses, it appears that the outgroup setting is ignored. This highlights a fundamental difference between undated and dating analyses. In undated analyses, all trees are formally unrooted, and rooting via an outgroup can take place during or after the phylogenetic inference. Thus, in the original, non-dating versions of MrBayes, the “outgroup” option was simply a convenience for the user, unless the outgroup consisted of multiple OTUs, in which case it serves as a topology constraint.

However, in a dating analysis, all sampled trees are always rooted, whether or not the user has decided on an outgroup. Furthermore, the mechanics of specifying an outgroup are more complex. Merely declaring an OTU an outgroup, or declaring an outgroup clade to be monophyletic, will not necessarily do the job. After all, a clade that is forced to be monophyletic could still be deeply nested inside the ingroup, unless something prevents this.

The simplest way to force the outgroup to be the earliest-branching group in a dating analysis is to set up a node constraint specifying that the ingroup is monophyletic. This could be programmed into the MrBayes outgroup command, but at the time of writing, it was not. In the case of Slater (2015)'s canidae.nex, it happens that there is a node constraint named "Canidae" and includes all living and fossil Canidae in the analysis. This constraint is used in the original Slater analysis, so the effect of the MrBayes outgroup problem is not noticed until the node constraints are removed; in this situation, some uniform clock tip-dating analyses fail to put the outgroup in the outgroup position (Supplemental Table S2). Fossilized BD analyses seem to put the outgroup in the correct position even without any constraints (Figure 1).

Issue 6. Typos in some OTU names.

Comparison with the "ground truth" tree manually digitized from the monographs of Tedford and Wang identified several likely typos in Canidae.nex (in fairness, comparison also revealed a number of typos in the draft ground truth tree; these are corrected in the final version). The correct spellings were double-checked via google and comparison to the monographs.

Typo	**Corrected**
Cynarctoides_accridens	Cynarctoides_acridens
Phlaocyon_marshlandensis	Phlaocyon_marslandensis
Paracynarctus_sinclari	Paracynarctus_sinclairi
Rhizocyon_oreganensis	Rhizocyon_oregonensis
Cynarctoides_gawanae	Cynarctoides_gawnae
Protomarctus_opatus	Protomarctus_optatus
Urocyon_galushi	Urocyon_galushai
Urocyon_citronus	Urocyon_citrinus