

Caveats on the Use of Fossil Calibrations for Molecular Dating: A Reply to Parham and Irmis

Thomas J. Near,^{1,*} Peter A. Meylan,^{2,†} and H. Bradley Shaffer^{3,‡}

1. Department of Ecology and Evolutionary Biology and Peabody Museum of Natural History, Yale University, New Haven, Connecticut 06520;

2. Natural Sciences, Eckerd College, St. Petersburg, Florida 33711;

3. Section of Evolution and Ecology and Center for Population Biology, University of California, Davis, California 95616

*Submitted January 27, 2007; Accepted September 24, 2007;
Electronically published December 7, 2007*

Keywords: penalized likelihood, divergence time, Testudines, cross validation, molecular clock, fossil record.

In Parham and Irmis's (2008) commentary on our article "Assessing Concordance of Fossil Calibration Points in Molecular Clock Studies: An Example Using Turtles" (Near et al. 2005), these authors address three areas of concern with respect to fossil calibration analyses in general and our article in particular: (1) absolute age estimates for fossil calibration points, including the accuracy of the reporting of age information and the inclusion of stratigraphic error; (2) the use of fixed calibration points in clock calibration studies; and (3) the phylogenetic placement of fossil taxa in our article.

Before we discuss the points raised by Parham and Irmis, it is important to emphasize that the primary goal of our article was to provide a new methodological approach for incorporating uncertainty and error into molecular clock calibration estimates. In that article, we developed a new method, based on cross-validation with multiple calibration points, to identify outlier fossils that may have been inaccurately dated or phylogenetically misplaced or that had otherwise been providing inaccurate dates for clade divergences. We applied that method to turtles because they have an excellent fossil record and have been

the subject of our own previous work. Thus, our work, as well as that of Parham and Irmis, focused on identifying problems inherent in the fossil record, with respect to providing absolute time calibrations for molecular phylogenies. In our article, we provided what we view as a methodological improvement over methods available at that time. Parham and Irmis provide a commentary on certain aspects of the implementation of our method, with respect to specific turtle fossils, but they make little comment on the actual method and no reanalysis that might demonstrate how to implement their concerns. This lack of any reanalysis makes it difficult to assess the actual impact of these concerns and, therefore, to determine whether they present any substantive challenges to the method developed in our article.

When we compiled our list of fossil calibrations for our article, we considered three possible pitfalls inherent in each fossil: (1) the phylogenetic placement of the fossil taxon; (2) the geological age estimate of the site from which the fossil originated; and (3) the completeness of the fossil record and, specifically, the possibility of older undiscovered fossils for each clade that a fossil dated. The results of our efforts to identify critical fossils were ultimately presented in our 2005 article as a single short paragraph, one table, and one figure. Our distillation of the fossil information is almost certainly the source of Parham and Irmis's desire for more detail about the fossils, and we can sympathize with that desire (we had, at the time of publication, a five-page document summarizing our research on fossil reliability). However, our focus was to introduce a new method to assess the agreement among individual fossil calibrations; it was not meant to be a treatise on fossil turtles. Consequently, we provided general readers with a set of references (by no means comprehensive) that would lead them to the relevant literature, and we assumed that the truly interested paleontologists and turtle systematists would contact us or delve into that literature for additional details. We consider it unfortunate that Parham and Irmis jumped to the conclusion that our study used a less-than-rigorous treatment of the fossil information because the details of establishing the phylogenetic posi-

* E-mail: thomas.near@yale.edu.

† E-mail: meylanpa@eckerd.edu.

‡ E-mail: hbshaffer@ucdavis.edu.

tion of the 17 fossil calibrations used in our study were not included as a part of the publication.

Reported Ages of Fossil Calibration Points

Parham and Irmis provide a broad critique of our use (and perhaps the general use) of geological ages for fossil calibrations, including (a) the lack of error on geological age estimates, (b) the precision of calibration times, and (c) the failure of our citations to support the calibration ages. Parham and Irmis (2008, p. 132) criticize the lack of geological error in our calibration ages, stating, “Despite citing Graur and Martin, NEA ignored stratigraphic error in their own study.” In reading Parham and Irmis’s commentary, we were hoping to find specific estimates of geological error for fossil ages or at least a strategy to obtain those errors. However, what Parham and Irmis provide are equally imprecise values for all the fossils listed in their table 1.

Both a strategy for and the impact of incorporating such error into our analyses are unclear because our strategy uses fossils as single fixed points to test whether individual calibrations agree with one another. An alternative strategy that we did not explore in our study is using paleontological dates as hard minimum and soft maximum constraints and using Bayesian methods that allow nonuniform prior distributions of node ages (Drummond et al. 2006; Benton and Donoghue 2007). We also believe that another important direction for future development of our method is incorporating error estimates of the geological dates assigned to fossils, as well as stratigraphic density error (Marshall 1990, 1994, 1997). However, no currently available clock calibration accounts for this type of error, including ours. Instead, our article showed that divergence time analyses using one set of fossils (each treated as a point estimate) resulted in significantly disparate molecular age estimates. However, an even larger set of the fossil calibrations independently provides very consistent divergence time estimates—that is, they cross-validate each other. Our intuition is that incorporating geological error into fossil age estimates would not substantially alter our conclusions regarding either set of fossils. It should not, on average, cause consistent fossils to become inconsistent because incorporating an error estimate would tend to make a larger set of fossils statistically consistent with one another. Thus, our method should still identify a consistent, although perhaps overly stringent, set of fossils to use as multiple calibration points. For those individual fossils that provide inconsistent age estimates, the geological error would presumably have to be very large for them to actually be consistent with other ages because the decrease in our measure of disagreement drops by two orders

of magnitude as the inconsistent fossils are removed from the analysis (fig. 2 in Near et al. 2005).

Parham and Irmis note that we were overly precise in reporting the fossil dates to the nearest 100,000 years (table 1 in Near et al. 2005), when the original literature gave fossils dates in approximate terms such as “about 210 mya.” That criticism is fair enough, and we admit that it was an oversight to include the decimal point in table 1 of our article (Near et al. 2005). We were “overly precise” because we treated the molecular age estimates to the nearest 100,000 years in calculating the disagreement between fossil and molecular age estimates (fig. 1 in Near et al. 2005). However, we see no reason to suspect that this “precision” should affect the magnitude of the calculations we conducted to assess the agreement among individual calibration points, and none is provided by Parham and Irmis. We agree that we should not have included the decimal point.

Parham and Irmis (2008, p. 132) incorrectly report that 15 of the 17 absolute ages we used were not supported by the references we provided. We believe that they meant to say that 15 of 17 references provide a range of ages rather than an absolute age. The range of dates given for each fossil in their table 1 does in fact agree with the absolute age estimates used in our study, at least in the sense that our point estimates fall within their ranges.

Parham and Irmis also contend that there is no basis in the literature for two of the fossil ages we used as calibrations (nodes 4 and 9). The citation we gave for node 4 (Gaffney et al. 2001) is the description of a new genus and species of fossil turtle that includes information regarding the probable horizon of origin of the type. Because the type was a commercially collected specimen, there is a lack of specific locality data, but a referred specimen (TUTg 1798) is reported as coming from the Romualdo Member of the Santana Formation. The estimated age for the Romualdo Member is about 110 million years ago (mya; Fara et al. 2005), and we used that date in our analysis. The reference for node 9 (de Lapparent de Broin 2001) does not give an absolute age for “*Ocadia*” *crassa*; however, this taxon has presented several nomenclatorial difficulties (for a discussion of the *Ocadia*-*Mauremys*-*Chinemys* taxonomic problem, see Spinks et al. 2004), and this publication provides the best introduction to turtles that have been called “*Ocadia*” from the Eocene of the London Clay and Messel (both of which were studied by one of us, P. A. Meylan). The detailed age information in this reference is given under the taxon *Palaeochoelys*, “(including many species attributed to ‘*Ocadia*’): lower Eocene MP7–lower Miocene” (de Lapparent de Broin 2001, p. 198). Franzen (1988) gives the approximate age of 49 million years for volcanic minerals present in Eocene Lake

Messel. We rounded this to 50 million years for our analysis.

Phylogenetic Placement of Fossil Calibration Points

Parham and Irmis assume that we used a subjective methodology in placing the turtle fossils in the phylogeny. They appear to base this assumption on a search of the literature for discrepancies between published hypotheses and our phylogenetic placement of fossils. Despite their efforts, they cite only two specific examples, and as we demonstrate below, both of these discrepancies are spurious. All of the fossils in our analysis were placed in the phylogeny via maximum parsimony analyses of more than 115 morphological characters by using the computer program PAUP* (Swofford 2003). In our article (Near et al. 2005, p. 139), we stated that “placement of fossils was guided by cladistic analysis of 115 morphological characters presented in earlier studies of turtle phylogenetic relationships (Shaffer et al. 1997).” We also used four to 12 additional characters (28 characters total) in four separate analyses that allowed us to determine the best placement of the fossils within suborders (Pleurodira) or superfamilies (Chelonioidea, Trionychoidea, Testudinoidea). After completing these analyses, we discussed our phylogenetic placement of fossils extensively with E. S. Gaffney and R. Hirayama, two experts on morphology-based turtle phylogenetics.

Parham and Irmis (2008, p. 134) specifically question the phylogenetic position of *Proterochersis* as the oldest side-necked turtle (Pleurodira) in our analyses and state that “removing this calibration point could change the estimated age of the crown group by 100 million years.” We contend that the available morphological evidence supports this placement of *Proterochersis* and that there is no evidence that including or excluding that fossil would change any age estimates using our method. It certainly would change the crown group age if it were the only fossil calibration point used, but that is precisely why we used multiple calibrations.

The possibility that *Proterochersis* is not a pleurodire was originally raised after the description of *Palaeochersis* from the Triassic of Argentina (Rougier et al. 1995). *Palaeochersis* was originally described as having the pelvis sutured into the shell, resulting in the suggestion that this complex character could no longer be considered a shared derived character for the Pleurodira. However, because of its very poor state of preservation, there has always been uncertainty regarding the state of this character in *Palaeochersis*. Joyce (2007) agrees that the condition of the pelvis is unclear in *Palaeochersis* and goes on to state that this is also the case in *Proterochersis*. We disagree. Personal observation of *Proterochersis* by one of us (P. A. Meylan) and

by E. S. Gaffney in 1986 revealed evidence of sutures of the pelvis and shell in several specimens (Gaffney et al. 2006). Furthermore, Joyce (2007) does not use a character of the sacral vertebrae that we have found to be useful in assessing the phylogenetic placement of this fossil species. A condition observed in all pleurodires, including *Proterochersis*, is that thoracic vertebra 10 is incorporated into the sacrum, while in cryptodires and *Proganochelys*, it is not (Gaffney et al. 1991, 2006). Thus, there is substantial morphological evidence to justify the phylogenetic placement of *Proterochersis* used in our analyses (see Gaffney et al. 2006). In addition, Parham and Irmis contend that this fossil is not dated by any of our references. This is also incorrect. A careful reading of the article by Gaffney (1986) would have revealed these lines:

Proterochersis occurs in the Stubensandstein of southern Germany, in the Murr Valley and the vicinity of Stuttgart. It is therefore mid-Keuper and Late Triassic (Norian) in age. Some of the *Proganochelys* specimens also come from the Stubensandstein of this region, and both *Proganochelys* and *Proterochersis* appear in the fossil record at the same time. (Gaffney 1986:184)

A more precise paleontological age estimate for this fossil is derived from the observation that the Stubensandstein spans the middle Norian (Alaunian), indicating an age range of 206–211 mya (Benton 1994; Ogg 2004). We used an age of 210 mya for this fossil in our analyses.

The other example of a phylogenetically misplaced fossil that Parham and Irmis cite results from what can only be explained as a failure to correctly read or understand how fossil calibrations were used in our figure 3. They interpret our use of the Lindholemydidae to date node 13 as evidence that we consider this taxon to be “basal to *Platysternon*” (Parham and Irmis 2008, p. 134), which would indeed be novel. (We are not actually certain what they mean by “basal to *Platysternon*,” although we assume that they mean sister to *Platysternon* rather than on the direct lineage leading to it.) However, the oldest fossil in either descendent clade dates node 13, as is the case for estimating the minimal ages for all dated nodes in figure 3. The Lindholemydidae as a member of (or sister group of) the descendant clade Testudinoidea (Danilov and Parham 2005) provides a useful date for this node in the arrangement that we used without being “basal” to *Platysternon*.

Conclusions

We disagree with Parham and Irmis’s (2008, p. 134) statement that “the fossil taxa do not meet the criteria outlined for effective calibration points.” On the contrary, we feel that the calibrations we used provide a reasonable basis for investigating the integration of fossil information with

molecular phylogenies and that our final subset of consistent fossils allowed us to produce a novel chronogram for turtle diversification. We concluded that not all of the 17 fossils in our initial analysis withstood scrutiny with our cross-validation method—and that was the point of our article. Here, we have addressed the concerns raised by Parham and Irmis about specific fossil calibrations, and it is clear to us that such disagreement does not negate the use of these fossils for calibrating turtle molecular phylogenies. Three things are missing from Parham and Irmis's critique of our work that could have provided a more useful follow-up to our article: a reanalysis of the data presented in our article, a set of critiques that are free of factual error, and an alternative method for identifying misplaced or poorly dated fossil calibration points. While by no means perfect, our use of the fossil record of turtles and the method we developed offer concrete steps toward more accurately calibrated chronograms—something that is of considerable utility for all those who are interested in studying the tempo of evolutionary change across the tree of life.

Acknowledgments

M. J. Sanderson provided insightful discussion and advice on applications of fossil calibrations.

Literature Cited

- Benton, M. J. 1994. Late Triassic terrestrial vertebrate extinctions: stratigraphic aspects and the record of the Germanic Basin. *Paleontologia Lombarda*, NS, 2:19–38.
- Benton, M. J., and P. C. J. Donoghue. 2007. Paleontological evidence to date the tree of life. *Molecular Biology and Evolution* 24:26–53.
- Danilov, I. G., and J. F. Parham. 2005. A reassessment of the referral of an isolated skull from the Late Cretaceous of Uzbekistan to the stem-testudinoid turtle genus *Lindholmemyx*. *Journal of Vertebrate Paleontology* 25:784–791.
- de Lapparent de Broin, F. 2001. The European turtle fauna from the Triassic to the Present. *Dumerilia* 4:155–217.
- Drummond, A. J., S. Y. W. Ho, M. J. Phillips, and A. Rambaut. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biology* 4:699–710.
- Fara, E., A. A. F. Saraiva, D. D. A. Campos, J. K. R. Moreira, D. D. C. Siebra, and A. W. A. Kellner. 2005. Controlled excavations in the Romualdo Member of the Santana Formation (Early Cretaceous, Araripe Basin, northeastern Brazil): stratigraphic, palaeoenvironmental and palaeoecological implications. *Palaeogeography Palaeoclimatology Palaeoecology* 218:145–158.
- Franzen, J. L. 1988. Europe in the Eocene: Messel in time and space. Pages 13–15 in S. Schall and W. Ziegler, eds. *Messel: an insight into the history of life and of the earth*. Clarendon, Oxford.
- Gaffney, E. S. 1986. Triassic and early Jurassic turtles. Pages 183–187 in K. Padian, ed. *The beginnings of the age of dinosaurs*. Cambridge University Press, Cambridge.
- Gaffney, E. S., P. A. Meylan, and A. R. Wyss. 1991. A computer-assisted analysis of the relationships of the higher categories of turtles. *Cladistics* 7:313–335.
- Gaffney, E. S., D. D. A. Campos, and R. Hirayama. 2001. *Cearachelys*, a new side-necked turtle (Pelomedusoides: Bothremydidae) from the early Cretaceous of Brazil. *American Museum Novitates* 3319: 1–20.
- Gaffney, E. S., H. Tong, and P. A. Meylan. 2006. Evolution of the side-necked turtles: the families Bothremydidae, Euraxemydidae, and Araripemydidae. *Bulletin of the American Museum of Natural History* 300:1–698.
- Joyce, W. G. 2007. Phylogenetic relationships of Mesozoic turtles. *Bulletin of the Peabody Museum of Natural History* 48:3–102.
- Marshall, C. R. 1990. Confidence intervals on stratigraphic ranges. *Paleobiology* 16:1–10.
- . 1994. Confidence-intervals on stratigraphic ranges: partial relaxation of the assumption of randomly distributed fossil horizons. *Paleobiology* 20:459–469.
- . 1997. Confidence intervals on stratigraphic ranges with non-random distributions of fossil horizons. *Paleobiology* 23:165–173.
- Near, T. J., P. A. Meylan, and H. B. Shaffer. 2005. Assessing concordance of fossil calibration points in molecular clock studies: an example using turtles. *American Naturalist* 165:137–146.
- Ogg, J. G. 2004. The Triassic period. Pages 271–306 in F. Gradstein, J. Ogg, and A. Smith, eds. *A geologic time scale 2004*. Cambridge University Press, Cambridge.
- Parham, J. F., and R. B. Irmis. 2008. Caveats on the use of fossil calibrations for molecular dating: a comment on Near et al. *American Naturalist* 171:132–136.
- Rougier, G. W., M. S. de la Fuente, and A. B. Arcucci. 1995. Late Triassic turtles from South America. *Science* 268:855–858.
- Shaffer, H. B., P. Meylan, and M. L. McKnight. 1997. Tests of turtle phylogeny: molecular, morphological, and paleontological approaches. *Systematic Biology* 46:235–268.
- Spinks, P. Q., H. B. Shaffer, J. B. Iverson, and W. P. McCord. 2004. Phylogenetic hypotheses for the turtle family Geoemydidae. *Molecular Phylogenetics and Evolution* 32:164–182.
- Swofford, D. L. 2003. *PAUP*: phylogenetic analysis using parsimony (*and other methods)*. Sinauer, Sunderland, MA.

Associate Editor: John J. Wiens
Editor: Monica A. Geber