Uncertainty in the Age of Fossils and the Stratigraphic Fit to Phylogenies

DIEGO POL¹ AND MARK A. NORELL²

¹CONICET, Museo Paleontológico Egidio Feruglio, Av. Fontana 140, Trelew 9100, Chubut, Argentina; E-mail: dpol@mef.org.ar ²Division of Paleontology, American Museum of Natural History, Central Park West at 79th street, New York, NY 10024, USA; E-mail: norell@amnh.org

Abstract.—The ages of first appearance of fossil taxa in the stratigraphic record are inherently associated to an interval of error or uncertainty, rather than being precise point estimates. Contrasting this temporal information with topologies of phylogenetic relationships is relevant to many aspects of evolutionary studies. Several indices have been proposed to compare the ages of first appearance of fossil taxa and phylogenies. For computing most of these indices, the ages of first appearance of fossil taxa are phylogenies. For computing most of these indices, the ages of first appearance of fossil taxa are currently used as point estimates, ignoring their associated errors or uncertainties. The effect of age uncertainty on measures of stratigraphic fit to phylogenies is explored here for two indices based on the extension of ghost lineages (MSM* and GER). A solution based on randomization of the ages of tratina is implemented, resulting in a range of possible values for measures of stratigraphic fit to phylogenies, rather than in a precise but arbitrary stratigraphic fit value. Sample cases show that ignoring the age uncertainty of fossil taxa can produce misleading results when comparing the stratigraphic fit of competing phylogenetic hypotheses. Empirical test cases of alternative phylogenies of two dinosaur groups are analyzed through the randomization procedure proposed here. [Age uncertainty; fossil record; GER; ghost lineages; MSM*; stratigraphic fit.]

Comparing the age of origination of taxa with a phylogenetic tree provides insight into the tempo and mode of the evolutionary history of a group, such as divergence age of its clades, evolutionary rates, and gaps in the fossil record (as implied by that particular tree). Several empirical measures have been proposed for assessing the fit between these ages and phylogenetic trees that include fossil taxa. These measures compare the temporal order of successive branching events with the age of appearance of terminal taxa in the stratigraphic record and are usually referred to as the stratigraphic fit to a phylogeny (Norell and Novacek, 1992; Benton and Stors, 1994; Huelsenbeck, 1994; Siddall, 1998; Wills, 1999; Pol and Norell, 2001; Pol et al., 2004). Such comparisons are frequently used to describe the stratigraphic fit of competing phylogenetic trees. Alternatively, similar comparisons have been proposed as auxiliary optimality criteria (e.g., Fisher, 1991, 1992, 1994; Wagner, 1995, 1998; Huelsenbeck and Rannala, 1997; Fox et al., 1999). Some of these procedures not only provide a measure of how well the stratigraphic appearance of terminal taxa fits their relative ordering in a phylogeny, but also provide minimal ages of divergences for every node in the tree based on temporal information in the fossil record.

The order of branching events on a phylogenetic tree is derived from hypotheses of evolutionary relationships, usually obtained from biological data alone, independent from temporal information (Norell, 1996; but see Fox et al., 1999). As noted previously, uncertainty in estimates of phylogenetic relationships could certainly affect the outcome of these metrics (Huelsenbeck, 1994; Benton et al., 1999). Consequently, some researchers have attempted to predict how these measures behave in response to erroneous phylogenetic trees through simulations (Wagner, 2000a; Wagner and Sidor, 2000).

The other kind of information used in analysis of stratigraphic fit to phylogenies is the age of fossil terminal taxa. The discussion presented here is restricted to ghost lineage metrics that have been formulated (Siddall, 1998; Wills, 1999) and used in empirical studies assuming the absence of ancestors among the terminal taxa of the phylogenetic tree (Benton and Storrs, 1994; Benton and Simms, 1995; Benton and Hitchin, 1996, 1997; Weishampel, 1996; Hitchin and Benton, 1997; Benton et al., 1999, 2000; Brochu and Norell, 2000; O'Leary, 2001; Merle and Pacaud, 2004). Under such an assumption, the only relevant temporal information for measuring the stratigraphic fit of a phylogeny is the age of the first appearance datum (FAD) of each terminal taxon in the fossil record. Analysis of stratigraphic fit to phylogenies that allow the recognition of ancestors among the terminal taxa also considers the last appearance datum (LAD) as potentially relevant information (as in stratocladistic or stratolikelihood methods, which also use stratigraphic fit as auxiliary optimality criterion to evaluate phylogenetic trees; see Fisher [1992] and Wagner [1998]). Irrespective of this distinction, the age of both FADs and LADs is generally determined through some form of chronostratigraphy. Although the ages of FADs are commonly taken as if they were raw observations (without observational errors), they are, in fact, inferences made upon observations of spatial location using a variety of assumption sets. Therefore, their ontological status is also hypothetical and, within a given temporal interval, inherently uncertain.

The impact of age uncertainty in measures of stratigraphic fit to phylogenies is explored here and a possible solution to its problems is proposed for measures of stratigraphic fit that are based on the extent of ghost lineages (sensu Norell, 1992). This procedure incorporates uncertainties in age assignment for FADs of fossil taxa.

AGE UNCERTAINTY OF FADS

In most cases the age of the FAD of a fossil taxon is inferred by determining the age of the associated sediments. Several direct and indirect methods are used for this purpose. Direct dating methods, such as radioisotopic dating, provide observations that can be used to estimate the absolute age of the rock directly. These ages typically consist of a temporal range (usually a point estimate with an associated error). Unfortunately, when such direct methods are used, suitable materials for dating are rarely associated to the exact horizon where a fossil is found. Even in the best cases they usually represent a temporal interval corresponding to dated sediments above and below the stratigraphic point of FAD of a fossil taxon (or a linear interpolation between these two dates).

Indirect dating methods are much more commonly used but certainly less precise than radioisotopic dating. These methods are based on the identification of certain changes in the sediments (e.g., paleomagnetic, lithological, or chemical composition) or on their fossil contents (e.g., biostratigraphy). This information is then used to correlate these sediments with other, betterstudied rocks referred to a temporal unit of a given temporal scale by more precise methods (e.g., radioisotopic dating). Examples of these temporal scales to which sediments are usually referred are the discrete chrons of the paleomagnetic time scale (i.e., the alternated periods of normal and reversed magnetic polarity of the Earth), the biostratigraphic zones and subzones, or the tuning to Milankovitch astronomical cycles (determined by periodic changes in the Earth's orbital eccentricity, axial tilt, and precession). Some of these time scales provide geological dates with narrow uncertainty intervals, and a great deal of chronostratigraphic research is being conducted on these subjects, significantly increasing the temporal resolution of the fossil record (e.g., Berggren et al., 1995; Shackleton et al., 1999; Hinnov, 2000; Muttoni et al., 2004). Unfortunately, some of these high-resolution time scales are not widely applicable, whereas others have a great potential but are not yet available for most terrestrial sedimentary basins (e.g., tuning lithologic cycles to Milankovitch cycles). Although future research on chronostratigraphy will likely provide more precisely constrained ages of first appearances, it is still relatively common (in particular for pre-Cenozoic continental sediments) to have stratigraphic correlations only at the level of a given chronostratigraphic stage (e.g., Norian, Toarcian, Aptian, etc.). These units correspond to subdivisions of the geological time scale that vary in their time span, although most range between 2 and 12 million years. Referring the age of an FAD to this level of temporal resolution would have an associated error that extends for the duration of the corresponding geochronologic unit.

In empirical analyses of stratigraphic fit to phylogenies, the temporal uncertainty in the age of first appearance of fossil terminal taxa is not currently considered. Usually, when using indirect methods or when the age of a fossil's FAD is bracketed by direct methods, the age of first appearance of fossil taxa is taken as the midpoint of the geochronologic temporal unit to which the sediments are referred (e.g., Weishampel, 1996; Benton and Storrs, 1994, 1996; Benton and Simms, 1995; Benton and Hitchin, 1996, 1997; Benton et al., 1999, 2000; Brochu and Norell, 2000; O'Leary, 2001). Such an approach can produce misleading results because the temporal difference between the FADs of two fossil taxa is always disregarded if they are referred to the same chronostratigraphic unit, but counted if they are found in different units. If the FADs of two fossil taxa are referred to the same geological subdivision (or in the error interval associated with a precise [e.g., radioisotopic] date), in absence of further evidence, the real age difference between these FADs could be as large as the time span of the chronostratigraphic unit or as small as zero (Fig. 1). In contrast, if the FADs of two fossil taxa are referred to two contiguous chronostratigraphic units, in absence of further evidence, the real age difference between these FADs could be as large as the temporal length of both chronostratigraphic units or could approach zero (Fig. 1).

In addition to the uncertainty in determining the age of the FAD of a fossil taxon, there is also some uncertainty in that the observed first occurrence actually represents its age of origination. Several methods exist that estimate confidence intervals on the temporal range of a fossil taxon (consequently extending its age of origination), which are based upon the distribution of gaps and occurrences of that fossil taxon in the sedimentary section (e.g., Strauss and Sadler, 1989; Marshall, 1990, 1994, 1997; Marshall et al., 1998; Foote and Raup, 1996; Foote, 1997; Solow and Smith, 1997; Wagner, 2000b; Solow, 2003). In the following discussion and examples, we will refer to the uncertainty in the determination of the age of the observed FAD of fossil terminal taxa. However, this



FIGURE 1. A hypothetical distribution of FADs of three fossil taxa through geologic time (dots = real FAD; dashed lines = inferred uncertainty interval based on correlation of sediments). The FADs of taxa 1 and 2 would be assigned the same age (chronostratigraphic unit B), whereas the FAD of taxon 3 is assigned to chronostratigraphic unit A, irrespective that the underlying age difference between the FADs of taxa 0 at 3 is less than that between 1 and 2.

uncertainty interval could be modified to incorporate the confidence intervals mentioned above.

In sum, several factors contribute to the uncertainty on estimates of the age of FADs. This uncertainty is translated into a temporal interval, the boundaries of which are usually well defined. It is therefore within these boundaries that the age uncertainty of the FADs should be considered.

Ignoring the inherent uncertainty of geological dates is not uncommon among phylogeneticists, and it is certainly not exclusive of studies dealing with stratigraphic fit of phylogenetic hypothesis. The use of geological dates of fossils for calibrating molecular clocks is usually conducted without much consideration of the uncertainty associated to the age of first appearance of a given taxon (Smith and Peterson, 2002; Brochu et al., 2004).

EFFECT OF AGE UNCERTAINTY

The effect of age uncertainty on two measures is evaluated here. Both of these are based on the magnitude of ghost lineages: the GER (Wills, 1999) and the MSM* (Siddall, 1998; Pol and Norell, 2001). Both measures also are based on the optimization of an "age character" on a phylogenetic tree (see Siddall, 1998). The age character is set by scoring each terminal taxon with a character state that represents the age of its FAD. The age character has an associated step matrix that determines the transformation costs between the character states as their pairwise temporal difference (in Mya or any other unit). The transformation costs from a younger age to an older age is set as infinite in order to assign the minimum divergence age to each node of the tree (see Pol and Norell, 2001). The age character is optimized using Sankoff parsimony (Sankoff and Rousseau, 1975) on the phylogenetic tree being evaluated to obtain its length (L_{o}) . This value (L_{o}) and the minimum and maximum possible number of steps of the age character (L_m and L_M , respectively) are used to calculate the MSM* and GER. The MSM^{*} is calculated as the consistency index $(L_m/L_o;$ Kluge and Farris, 1969) of the age character, whereas the GER can be formulated as the retention index of the age character ($[L_M - L_o]/[L_M - L_m]$). These metrics are sensitive to the temporal duration of mismatches between phylogeny and stratigraphy and therefore are those most severely affected by age uncertainty.

Two distinct effects are seen in these measures when there is age uncertainty in the FADs of terminal taxa. The first occurs when the temporal uncertainties of the FADs of fossil taxa do not temporally overlap (Fig. 2). Here, the choice of a particular set of ages from the uncertainty intervals of FADs of terminal taxa can affect the absolute value of stratigraphic fit to a given tree. However, the relative fit value of a given tree, with respect to that of other trees of the same set of taxa, will not be overturned if a different set of ages is taken from the age uncertainty intervals of their FADs (i.e., the ranking of stratigraphic fit of a set of competing trees will remain the same). Thus, this situation would be inconsequential for the comparison of trees and for significance tests associated with these measures (Siddall, 1998; Wills, 1999). The second case occurs when the uncertainty intervals associated to the age of FADs of some fossil taxa overlap. In this case, the relative stratigraphic fit of competing trees can change, depending on the set of ages taken from the uncertainty intervals of FADs of terminal taxa (except for the trivial case in which the compared uncertainty intervals are identical). Hence, the stratigraphic fit of competing trees can be ordered differently if one takes the lower limit of all uncertainty intervals of the age of FADs rather than if one takes the midpoint of all uncertainty intervals of FADs (Fig. 3). Given such a scenario, there is no rational basis for choosing between these two options (or any other set of age assignments).

These examples show that it is not possible to calculate precise measures of stratigraphic fit when the age of the FADs of terminal taxa are imprecise. Because there is no rational justification for choosing any particular set of age assignments, it is necessary to consider a range of possible values for measures of stratigraphic fit, instead of having a precise but arbitrary (and in some cases biased) metric value for each phylogenetic tree. Fortunately, the temporal uncertainties of FADs of fossil taxa can be translated into a range of possible values for these two measures of stratigraphic fit. Heretofore we will refer to the ranges of possible stratigraphic fit values (stemming from considering uncertainties on the age of FADs) as the MSM* range or the GER range. Most of the following discussion is equally applicable to both measures, but we will mostly refer to the MSM* range for the sake of simplicity.

INCORPORATING AGE UNCERTAINTY

The possible range of stratigraphic fit values (i.e., MSM^{*} range or GER range) for a given tree could be calculated given the set of uncertainty intervals associated to the ages of FADs of terminal taxa. Currently, there is no algorithm to determine both the maximum and minimum possible values of the MSM* range on a particular tree (given the age uncertainty intervals of FADs of fossil terminal taxa). Nevertheless, the maximum possible value of the MSM* range can be obtained by incorporating polymorphic scorings in the age character and optimizing it as usual with Sankoff parsimony. Unfortunately, the minimum possible value of the MSM* range cannot be garnered through the algorithms implemented in currently available software implementing Sankoff parsimony (e.g., Paup* 4b10 [Swofford, 2002], TNT [Goloboff et al., 2003]). However, when comparing the stratigraphic fit of competing phylogenetic hypotheses, just finding the maximum and minimum possible values of their MSM* ranges could be misleading or, at least, insufficient. Intuitively, if competing phylogenetic trees have similar MSM* ranges, it may seem to indicate that their stratigraphic fit is approximately equal. However, there are cases where trees with similar boundaries in their MSM* ranges can be distinguished based on their fit to the temporal information of the fossil record (see below).



FIGURE 2. Two hypothetical phylogenetic trees mapped against geological time. The fossil terminal taxa have the uncertainty intervals of their FADs plotted with a dashed line, none of which overlap in time. Three MSM* and GER values are shown, obtained considering the youngest age (MSM^*_{y} , GER_y), the oldest age (MSM^*_{o} , GER_o), and the midpoint age (MSM^*_{m} , GER_m) of each uncertainty interval of the FADs. Note that the different assignments change the raw values of these metrics but the tree on the left always has a better fit than the tree on the right.

Alternatively, the MSM* range can be approximated through randomizations of the age assignments to the FADs taken from their age uncertainty intervals (e.g., age, epoch, chron, etc.). A randomization approach to the MSM* range provides several advantages, such as testing the sensitivity of the MSM*-based ranking of trees to alternative assignments of FADs, the allowance of constraining age assignment to be either equal among coeval taxa (e.g., set of taxa found in the same horizon and locality) or to be necessarily older for some taxa with respect to others (e.g., if the relative order of their FADs is known but their age uncertainty intervals overlap on time), or the possibility of incorporating different probability distributions from which the ages are drawn in each randomization replicate.

Randomization Approaches to the MSM* Range

The randomization approach proposed here consists of performing multiple replicates, where in each replicate a precise age of first appearance is assigned to each terminal taxon (taken randomly from the uncertainty



FIGURE 3. Two hypothetical phylogenetic trees mapped against geological time. The terminal taxa have the uncertainty intervals of their FADs plotted with a dashed line. The uncertainty intervals of the FADs of some terminal taxa overlap. Three MSM* and GER values are shown, obtained considering the youngest age (MSM*_y, GER_y), the oldest age (MSM*_o, GER_o), and the midpoint age (MSM*_m, GER_m) of each uncertainty interval of FADs. Note that the different age assignments change both the raw values of these metrics and the relative fit of these two trees.

interval associated to the age of each taxon's FAD which is the imprecision of age caused by error or the duration of the referred unit of geologic time). In each replicate, the ages assigned to the terminal taxa are used to calculate the stratigraphic fit (e.g., MSM*, GER) of the phylogenetic trees being analyzed. Since the randomized assignments of FAD ages are replicated a given number of times (e.g., a thousand), an MSM* range is finally obtained for each phylogenetic tree. In addition to the MSM* range of each tree, the randomization procedure also provides the relative ranking of the topologies being analyzed in each of the replicates (according to the MSM*). The latter information is critical, because in some cases the MSM* ranges of two phylogenetic trees may be largely overlapping, but one of them can have a higher MSM* than the other tree in all replicates (i.e., for any possible combination of FAD age assignments). In these cases, we must conclude that one of the trees has a better stratigraphic fit than the other because their difference in MSM* values is always positive (see Fig. 4A), irrespective of the uncertainty in the ages of the FADs of terminal taxa. Clearly, when comparing trees, if one of them has a



FIGURE 4. Outcome of the randomization procedure (100 replicates) applied to the data and trees shown in Figures 2 and 3. The graphs show a frequency histogram of the difference in MSM* values of both trees obtained in each of the replicates of the randomization procedure, taking one of the trees as reference (MSM* difference equals MSM* of reference tree minus MSM* of compared tree). The MSM* ranges of the reference tree and the compared tree are shown above each graph. (A) MSM* difference histogram obtained during the analysis of hypothetical trees shown in Figure 2, taking tree on the left of Figure 2 as reference tree. Note that although MSM* ranges of both trees largely overlap, the MSM* difference is always positive; meaning that the reference tree had a better stratigraphic fit in all replicates. (B) MSM* difference histogram obtained during the analysis of hypothetical trees shown in Figure 3, taking tree on the left of Figure 3 as reference tree. Note that MSM* difference is always positive; meaning that the reference tree had a better stratigraphic fit in all replicates. (B) MSM* difference is positive in some replicates but negative in others; meaning that the reference tree cannot be considered to have a better stratigraphic fit than the compared tree (i.e., their MSM*-based ranking is sensitive to the ages assigned from the uncertainty interval of the FADs of terminal taxa).

higher MSM* value in some replicates but lower in others, their stratigraphic fit should be regarded as equal (Fig. 4B).

Implementation.—This procedure was implemented in the scripting language of TNT (Goloboff et al., 2003) and is available at request from the authors (or downloaded from http://research.amnh.org/~dpol/strat). This script reads the dataset containing the age character (with their associated step matrix), the topologies to compare, and the minimum and maximum ages of each taxon's FAD (i.e., their associated temporal uncertainty interval). The output consists of a list with the MSM* ranges of each tree and a detailed list with the MSM* values of each tree for each replicate. The latter list allows calculating a histogram of MSM* differences for each replicate (see Fig. 4).

This implementation includes two options for constraining the age assignments of FADs according to additional stratigraphic information on the relative order of FADs. The first option allows forcing the age assignment of some terminal taxa to be equal during each replicate (rather than assign them independently at random). This is necessary if, for example, the FADs of two or more terminal taxa are recorded at the same horizon, or if there is a high degree of certainty on the correlation of the sediments bearing these FADs. The second option allows constraining the age assignments of some terminal taxa to be necessarily older (or younger) than those of other taxa. This is useful when, for example, there is no doubt on the relative age of the FADs of two fossil taxa (e.g., one is consistently found above the other in the same sedimentary section), but the age uncertainty intervals of both FADs overlap on time (i.e., the absolute age of these sediments is rather uncertain). This option is potentially useful because the relative ages of fossil taxa are usually more precisely known than their absolute ages.

Empirical Test Cases

The randomization procedure was applied to measure the stratigraphic fit of competing hypotheses for two dinosaur groups. The first case contrasts the stratigraphic fit of two trees based on recently published hypotheses on the early radiation of Sauropodomorpha (Yates, 2003; Galton and Upchurch, 2004), one of the major groups of Dinosauria. These two trees differ markedly in their topology and evolutionary implications (Fig. 5). One of them depicts a large monophyletic Prosauropoda (composed by Late Triassic and Early Jurassic forms; node P in Fig. 5A) as the sister group of Sauropoda. The competing tree depicts Prosauropoda as a paraphyletic arrangement of taxa relative to the clade Sauropoda (Fig. 5B).



FIGURE 5. Outcome of the randomization procedure applied to two competing hypotheses of the phylogeny of basal sauropodomorph dinosaurs. The MSM* range of these trees is shown below each tree. (A) Phylogenetic relationships following Galton and Upchurch (2004), (B) phylogenetic relationships following Yates (2003), (C) frequency histogram of the difference in MSM* values obtained in each of the replicates of the randomization procedure; details of the histogram graph are as in Figure 4 (tree shown in A taken as reference tree). The MSM*difference is positive in some replicates but negative in others; meaning that none of these trees has a better stratigraphic fit than the other. Their MSM* ranking is sensitive to the temporal uncertainty in the FADs of terminal taxa. The original trees of both studies were modified (taxa not present in both analyses pruned from trees). The tree shown in B was randomly chosen from the set of most parsimonious trees (other trees produced similar results).





FIGURE 6. Outcome of the randomization procedure applied to the competing hypotheses on the evolutionary origins of birds. (A) Tree depicting the phylogenetic position of Avialae within Diapsida, the tree shows the bird lineage (circled in gray) deeply nested within Dinosauria. This tree is compared with six alternative phylogenetic positions for Avialae proposed by other authors (positions for Avialae marked with a gray dot on the tree; see Brochu and Norell, 2000, for details); (B) frequency histogram of the difference in MSM* values of the seven proposed hypotheses on the evolutionary origin of birds obtained in each of the replicates of the randomization procedure (tree depicting a dinosaur origin of birds taken as reference for MSM* differences). The MSM*difference between the reference tree and all other topologies is positive in all replicates, meaning that alternative trees score lower in their stratigraphic fit (as measured by the MSM*). The hypothesis of the dinosaur origins of birds (reference tree) has a higher MSM* than alternative hypotheses and this ranking is not sensitive to the temporal uncertainties in the FADs of terminal taxa. The MSM* ranges of the seven competing phylogenies are shown in the top left corner this graph.

Their relative ordering in stratigraphic fit, however, is strongly dependent on the age assigned to the FADs of terminal taxa and, therefore, in absence of further evidence, they must be viewed as having a similar degree of agreement with the temporal information of the fossil record (Fig. 5C).

The second test case is an analysis of trees depicting the long-standing debate surrounding the origin of birds (Fig. 6A). One of the arguments offered against the dinosaur origins of birds is that it creates a "temporal paradox," because the earliest member of Avialae (i.e., *Archaeopteryx*) is known from the Late Jurassic, whereas the dinosaur closest relatives (i.e., dromaeosaurids and other maniraptoran theropods) are known from the Cretaceous (Feduccia and Martin, 1998). The data analyzed here were taken from Brochu and Norell (2000), who made a similar comparison but used the midpoint value of the uncertainty interval associated to the age of each taxon's FAD. These authors concluded that the hypothesis depicting the dinosaurian origin of Aves has a higher stratigraphic fit than any of the proposed alternative hypotheses. Our application of the methods described here demonstrates that this conclusion is robust to the incorporation of age uncertainty in the evaluation of stratigraphic fit to phylogeny (Fig. 6B).

DISCUSSION

Previous approaches to calculating the stratigraphic fit of phylogenetic trees based on ghost lineages were derived assigning an exact age for the FAD of each terminal taxon, irrespective of the actual precision of the chronostratigraphic information on this datum. The stratigraphic fit of phylogenetic trees should be viewed as a comparison between the temporal content of two independently derived hypotheses—one of toplogy and one of age. It could be argued there is more certainty on the age estimate of a fossil taxon's FAD than on its phylogenetic placement, although this issue depends on the taxon, the method of age estimation, and several other factors. This, however, is a distinction on the precision and degree of support of these hypotheses, not on their status as observations or hypotheses. We show above that taking rough approximations of chronostratigraphic information, treating this data as precise by rounding to the mean can produce misleading results regarding the relative stratigraphic fit of alternative phylogenetic trees. These misleading conclusions will affect most applications of these indices, especially if they are considered as auxiliary optimality criteria.

The solution proposed here incorporates age uncertainty in two measures of stratigraphic fit (MSM* and GER), producing a range of possible values of these metrics for a particular phylogenetic tree (MSM* range and GER range) rather than a precise yet extremely arbitrary value. We have focused here on the effects of age uncertainty in these two measures because they are the most sensitive to differences in temporal data. Other metrics that are solely based on the relative ordering of FADs, not measuring the extent of the mismatches implied by a tree (i.e., temporal extension of ghost lineages), can also be affected but probably to a lesser degree. In particular, measures such as SCI (Huelsenbeck, 1994), SRC (Norell and Novacek, 1992), or stratocladistic approaches (Fisher, 1994) could only be affected if there is overlap between uncertainty intervals associated to the FADs of terminal taxa.

Further refinements of the method proposed here could incorporate more complex models in the age assignment process, instead of the randomized equiprobable age assignments based on the uncertainty intervals of the FADs. For instance, some age assignments could be drawn from a specified probability distribution of a particular direct dating estimate, or using probabilistic models that incorporate extensions to the observed temporal range of a fossil taxon based on confidence intervals stemming from the distribution of its occurrences in the stratigraphic record (see above). Although the methods of calculating phylogenetic trees are well understood, these and other modifications of the current implementation need to be further explored in order to integrate all the available chronostratigraphic information into better measures of stratigraphic fit to phylogenies.

ACKNOWLEDGMENTS

We would like to thank P. Makovicky and M. Siddall for useful discussions on this subject. W. Wheeler, P. Goloboff, and N. Christie-Blick provided useful comments on earlier versions of this manuscript. J. Wilson, P. Sadler, M. Benton, N. MacLeod, and R. Page provided insightful suggestions that significantly improved the quality and clarity of this manuscript. Part of this contribution was based upon work supported by the National Science Foundation under Agreement No. 0112050.

REFERENCES

- Benton, M. J., and R. Hitchin. 1996. Testing the quality of the fossil record by groups and by major habitats. Hist. Biol. 12:111–157.
- Benton, M. J., and R. Hitchin. 1997. Congruence between phylogenetic and stratigraphic data on the history of life. Proc. R. Soc. Lond. B Biol. 264:885–890.
- Benton, M. J., and M. J. Simms. 1995. Testing the marine and continental fossil records. Geology 23:601–604.
- Benton, M. J., and G. W. Storrs. 1994. Testing the quality of the fossil record: Paleontological knowledge is improving. Geology 22:111– 114.
- Benton, M. J., and G. W. Storrs. 1996. Diversity in the past: Comparing cladistic phylogenies and stratigraphy. Pages 19–24 *in* Phylogeny and biodiversity (M. Hochberg, J. Clobert, and R. Barbault, eds.). Oxford University Press, Oxford, UK.
- Benton, M. J., M. A. Wills, and R. Hitchin. 1999. Assessing congruence between cladistic and stratigraphic data. Syst. Biol. 48:580–591.
- Benton, M. J., M. A. Wills, and R. Hitchin. 2000. Quality of the fossil record through time. Nature 403:534–537.
- Berggren W. A., D. V. Kent, C. C. Swisher, and M. P. Aubry. 1995. A revised Cenozoic geochronology and chronostratigraphy. Pages 129–212 *in* Geochronology, Time scales and global stratigraphic correlation (W. A. Berggren, D. V. Kent, M. P. Aubry, J. Hardenbol, eds.). SEPM Special Publication 54.
- Brochu, C. A., and M. A. Norell. 2000. Temporal congruence and the origin of birds. J. Vertebr. Paleontol. 20:197–200.
- Brochu, C. A., C. C. Sumrall, and J. M. Theodor. 2004. When clocks (and communities) collide: Estimating time from molecules and the fossil record. J. Paleontol. 78:1–6.
- Feduccia, A., and L. D. Martin. 1998. Theropod-bird link reconsidered. Nature 391:754.
- Fisher, D. C. 1984. The Xiphosurida: archetypes of bradytely? Pages 196–213 *in* Living fossils (N. Eldredge and S. M. Stanley, eds.). Springer Verlag, New York.

- Fisher, D. C. 1991. Phylogenetic analysis and its application in evolutionary analysis. Pages 103–121 *in* Short courses in paleontology: Number 4 (N. Gilinsky and P. Singor, eds.). Paleontological Society. Knoxville, Tennessee.
- Fisher, D. C. 1992. Stratigraphic parsimony. Pages 124–129 in MacClade: Analysis of phylogeny and character evolution, version 3 (W. P. Maddison and D. R. Maddison, eds.). Sinauer Associates, Sunderland, Massachusetts.
- Fisher, D. C. 1994. Stratocladistics: Morphological and temporal patterns and their relation to phylogenetic process. Pages 133–171 *in* Interpreting the hierarchy of nature—From systematic patterns to evolutionary process theories (L. Grande and O. Rieppel, eds.). Academic Press, Orlando, Florida.
- Foote, M., and D. Raup. 1996. Fossil preservation and the stratigraphic ranges of taxa. Paleobiology 22:121–140.
- Fox, D. L., D. C. Fisher, and L. R. Leighton. 1999. Reconstructing phylogeny with and without temporal data. Science 284:1816–1819.
- Galton, P. M., and P. Upchurch. 2004. Prosauropoda. Pages 259–322 in The Dinosauria, 2nd edition (D. Weishampel, P. Dodson, and H. Osmólska, eds.). University of California Press, Berkeley, California.
- Gingerich, P. D. 1990. Stratophenetics. Pages 437–442 *in* Paleobiology: A synthesis (D. E. G. Briggs and P. R. Crowther, eds.). Blackwell Scientific Press, Oxford.
- Goloboff, P. A., J. S. Farris, and K. Nixon. 2003. T.N.T.: Tree analysis using new technology. Program and documentation, available from the authors, and at www.zmuc.dk/public/phylogeny.
- Hinnov, L. A. 2000. New perspectives on orbitally forced stratigraphy. Annu. Rev. Earth Planet. Sci. 28:419–475.
- Huelsenbeck, J. P. 1994. Comparing the stratigraphic record to estimates of phylogeny. Paleobiology 20:470–483.
- Huelsenbeck, J. P., and B. Rannala. 1997. Maximum likelihood estimation of topology using stratigraphic data. Paleobiology 23:174–180.
- Marshall, C. R. 1990. Confidence intervals on stratigraphic ranges. Paleobiology 16:1–10.
- Marshall, C. R. 1994. Confidence intervals on stratigraphic ranges: Partial relaxation of the assumption of randomly distributed fossil horizons. Paleobiology 20:459–469.
- Marshall, C. R. 1997. Confidence intervals on stratigraphic ranges with nonrandom distributions of fossil horizons. Paleobiology 23:165– 173.
- Marshall, C. R., S. Tavare, O. Will, R. D. Martin, and C. Sdoligo. 1998. Taking into account the temporal ranges of unobserved species: Reconciliation of fossil and molecular clock estimates of primate origins? Geol. Soc. Am. Abstr. Prog. 30:A326.
- Merle, D., and J. M. Pacaud. 2004. New species of Eocithara Fischer, 1883 (Mollusca, Gastropoda, Harpidae) from the Early Paleogene with phylogenetic analysis of the Harpidae. Geodiversitas 26:61–87.
- Muttoni, G., D. V. Kent, P. E. Olsen, P. DiStefano, W. Lowrie, S. Bernasconi, and F. M. Hernandez. 2004. Tethyan magnetostratigraphy from Pizzi Mondello (Sicily) and correlation to the Late Triassic Newark astrochronological polarity time scale. Geol. Soc. Am. Bull. 116:1043–1058.
- Norell, M. A. 1992. Taxic origin and temporal diversity: The effect of phylogeny. Pages 89–118 *in* Extinction and phylogeny (M. J. Novacek and Q. D. Wheeler, eds.). Columbia University Press, New York.
- Norell, M. A. 1992. 1996. Ghost taxa, ancestors, and assumptions: a comment on Wagner. Paleobiology 22:453–455.
- Norell, M. A., and M. J. Novacek. 1992. The fossil record and evolution: comparing cladistics and paleontological evidence for vertebrate history. Science 255:1690–1693.
- O'Leary, M. 2001. The phylogenetic position of cetaceans: Further combined data analyses, comparisons with the stratigraphic record

and a discussion of character optimization. Am. Zool. 41:487-506.

- Pol, D., and M. A. Norell. 2001. Comments on the Manhattan Stratigraphic Measure. Cladistics 17:285–289.
- Pol, D., M. A. Norell, and M. E. Siddall. 2004. Measures of stratigraphic fit to phylogeny and their sensitivity to tree size, tree shape, and scale. Cladistics 20:64–75.
- Sankoff, D., and P. Rousseau. 1975. Locating the vertices of a Steiner tree in arbitrary space. Math. Program. 9:240–246.
- Shackleton, N. J., I. N. McCave, and G. P. Weedon. 1999. Preface to astronomical (Milankovitch) calibration of the geological time-scale. Phil. Trans. Math. Phys. Eng. Sci. 357:1733–1734.
- Siddall, M. E. 1998. Stratigraphic fit to phylogenies: A proposed solution. Cladistics 14:201–208.
- Smith, A. B., and K. J. Peterson. 2002. Dating the time of origin of major clades: Molecular clocks and the fossil record. Ann. Rev. Earth Planet. Sci. 30:65–88.
- Solow, A. R. 2003. Estimation of stratigraphic ranges when fossil finds are not randomly distributed. Paleobiology 29:181–185.
- Solow, A. R., and W. Smith 1997. On fossil preservation and the stratigraphic ranges of taxa. Paleobiology 23:271–277.
- Strauss, D., and P. M. Sadler. 1989. Classical confidence intervals and Bayesian probability estimates for ends of local taxon ranges. Math. Geol. 21:411–427.
- Swofford, D. L. 2002. PAUP* 4b10. Phylogenetic analysis using parsimony (*and other methods). Sinauer Associates, Sunderland, Massachusetts.
- Wagner, P. J. 1995. Stratigraphic tests of cladistic hypotheses. Paleobiology 21:153–178.
- Wagner, P. J. 1996. Ghost taxa, ancestors, assumptions, and expectations: A reply to Norell. Paleobiology 22:456–460.
- Wagner, P. J. 1998. A likelihood approach for estimating phylogenetic relationships among fossil taxa. Paleobiology 24:430–449.
- Wagner, P. J. 2000a. Phylogenetic analyses and the fossil record: Tests and inferences, hypotheses and models. Pages: 341–371 *in* Deep time—Paleobiology's perspective (D. H. Erwin and S. L. Wing, eds.). Paleobiology Memoir 26 (Suppl. to No. 4). The Paleontological Society, Lancaster, Pennsylvania.
- Wagner, P. J. 2000b. Likelihood tests of hypothesized durations: Determining and accommodating biasing factors. Paleobiology 26:431– 449.
- Wagner, P. J., and C. A. Sidor. 2000. Age rank/clade rank metrics— Sampling, taxonomy, and the meaning of "stratigraphic consistency." Syst. Biol. 49:463–479.
- Weishampel, D. B. 1996. Fossils, phylogeny, and discovery: A cladistic study of the history of tree topologies and ghost lineage durations J. Vertebr. Paleontol. 16:191–197.
- Wills, M. A. 1999. Congruence between stratigraphy and phylogeny: Randomization tests and the gap excess ratio. Syst. Biol. 48:559–580.
- Yates, A. M. 2003. A new species of the primitive dinosaur *Thecodon-tosaurus* (Saurischia: Sauropodomorpha) and its implications for the systematics of early dinosaurs. J. Syst. Palaeontol. 1:1–42.
- Yin, H., W. C. Sweet, B. F. Glenister, G. Kotlyar, H. Kozur, N. D. Newell, J. Sheng, Z. Yang, and Y. D. Zakharov. 1996. Recommendation of the Meishan Section as Global Stratotype Section and Point for basal boundary of Triassic System. Newsl. Stratigr. 34:81–108.

First submitted 22 April 2005; reviews returned 20 June 2005; final acceptence 6 March 2006 Associate Editor: Norm MacLeod