OXFORD UNIVERSITY PRESS | Systematic Biology

Phylogenetic stability, tree shape, and character compatibility: a case study using early tetrapods

Journal:	*Systematic Biology*
Manuscript ID	USYB-2014-243.R1
Manuscript Type:	Regular Manuscript
Date Submitted by the Author:	n/a
Complete List of Authors:	Bernardi, Massimo; MUSE - Museo delle Scienze, Geology and Palaeontology Angielczyk, Kenneth; Field Museum of Natural History, Integrative Research Center Mitchell, Jonathan; University of Michigan Ruta, Marcello; University of Lincoln, School of Life Sciences
Keywords:	Character compatibility, Tree balance, Tree distance, Diversification shifts, Tetrapods , Terrestrialization, Paleozoic, Mesozoic

Phylogenetic stability, tree shape, and character compatibility: a case study using early tetrapods

Massimo Bernardi[1,2], Kenneth D. Angielczyk[3], Jonathan S. Mitchell[4], and Marcello Ruta[5]

[1] *MuSe – Museo delle Scienze, Corso del Lavoro e della Scienza, 3, 38122 Trento, Italy.*

[2] *School of Earth Sciences, University of Bristol, Wills Memorial Building, Queens Road, Bristol, BS8 1RJ, United Kingdom.*

[3] *Integrative Research Center, Field Museum of Natural History, 1400 South Lake Shore Drive, Chicago, IL 60605-2496, USA.*

[4] *Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48103, USA*

[5] *School of Life Sciences, Joseph Banks Laboratories, University of Lincoln, Green Lane, Lincoln LN6 7DL, United Kingdom.*

* Corresponding author

Massimo Bernardi, MuSe – Museo delle Scienze, Corso del Lavoro e della Scienza, 3, 38122 Trento, Italy

Email: massimo.bernardi@muse.it

Phone: +39 0461 270344

Abstract

Phylogenetic tree shape varies as the evolutionary processes affecting a clade change over time. In this study, we examined an empirical phylogeny of fossil tetrapods during several time intervals, and studied how temporal constraints manifested in patterns of tree imbalance and character change. The results indicate that the impact of temporal constraints on tree shape is minimal and highlights the stability through time of the reference tetrapod phylogeny. Unexpected values of imbalance for Mississippian and Pennsylvanian time slices strongly support the hypothesis that the Carboniferous was a period of explosive tetrapod radiation. Several significant diversification shifts (i.e., lineage multiplication events) take place in the Mississippian and underpin increased terrestrialization among the earliest limbed vertebrates. Character incompatibility is relatively high at the beginning of tetrapod history, but quickly decreases to a relatively stable lower level, relative to a null distribution based on constant rates of character change. This implies that basal tetrapods had high, but declining, rates of homoplasy early in their evolutionary history, although the origin of Lissamphibia is an exception to this trend. The time slice approach is a powerful method of phylogenetic analysis and a useful tool for assessing the impact of combining extinct and extant taxa in phylogenetic analyses of large and speciose clades.

Keywords:

Character compatibility, Tree balance, Tree distance, Diversification shifts, Tetrapod

Terrestrialization, Paleozoic, Mesozoic

Phylogeny reconstruction is a cardinal component of modern evolutionary biology because it provides the fundamental framework for investigating the dynamics of evolutionary processes, including tempo and mode of change and models of group diversification. Tree shape may be substantially altered by different régimes of character and taxon inclusion/exclusion, and by different character coding, ordering and weighting schemes. As a result, much interest surrounds phylogenetic stability, namely "the tendency for clades that are resolved by an analysis to continue to be resolved when either the data or the analytical method is altered" (Davis et al. 1993, p. 188). Numerous methods are now available for measuring cladistic stability (e.g., Felsenstein 1985; Bremer 1988; Goloboff 1991; Kållersïo et al. 1992; Davis 1993; Faith and Ballard 1994; Farris et al. 1996; Gatesy 2000), that is the amount of statistical support for tree nodes. However, a particularly relevant aspect of stability in a paleontological context is the impact of taxa from different time intervals on phylogenetic resolution. Because of such factors as genetic saturation (e.g., Felsensein 1978; Huelsenbeck and Hillis 1993) and morphological exhaustion (Wagner 2000a), later-evolving taxa might erode phylogenetic signal among early-evolving taxa. Thus, it is important to investigate whether phylogenetic stability (as defined above) remains constant with the addition of later-evolving taxa, or whether it changes over clade history. As a metaphor (Peter J. Wagner, personal communication, 2014), imagine a systematist living in the Pennsylvanian. How accurately could they reconstruct the phylogeny of tetrapods using just the taxa in that time period? Would the accuracy of their tree improve if they included both "contemporaneous" taxa and, say, "fossil" taxa from an earlier interval (e.g.

Devonian)? What would a phylogeny look like from the standpoint of a systematist living in the Permian, in terms of accuracy and stability? The significance of these questions goes beyond the specific arrangement of taxa on the tree. Thus, factors such as the rate of character state changes and the potential of later evolving characters to “erode” the signal of earlier evolving characters should also be considered. Beginning with the work of the 'Woods Hole Group' of paleontologists (Raup et al. 1973; Gould et al. 1977; Schopf 1979; see summaries by Slowinski and Guyer 1989; Mooers and Heard 1997; Huss 2009), tree shape has been used to analyze the tempo and mode of cladogenetic events (e.g., Savage 1983; Heard 1992; Guyer and Slowinski 1993; Mooers and Heard 1997; Chan and Moore 2002; Good-Avila et al. 2006; Heath et al. 2008). Despite the important initial role of paleontologists, some subsequent work has focused on phylogenies of extant taxa only (although see Harcourt-Brown et al. 2001; Harcourt-Brown 2002). This “neontological bias” is reflected by the fact that some recent applications of diversification shift analyses to paleontological trees (e.g., Ruta et al. 2007; Lloyd et al. 2008; Botha-Brink and Angielczyk 2010) required modifications of available methods to fit better the nature of fossil data (see also Tarver and Donoghue 2011; Brocklehurt et al. 2015), even though the importance of fossil data has become widely recognized (e.g., time-calibrating trees: Stadler 2010; Parham et al. 2011; Didier et al. 2012). Harcourt-Brown (2002) suggested that analysis of tree balance at different time intervals in a group's history could provide insight into diversification patterns, but there has been little additional work on this topic. Here, we build on Harcourt-Brown's (2002) study by examining changes in tree shape imparted by taxon addition during successive time intervals, and discuss the implications of those changes. We focus on three complementary aspects of tree shape: (1) stability, i.e., the retrieval of identical mutual relationships among taxa when new taxa are added to an existing data matrix; (2) balance, i.e., a measure of how symmetrical or asymmetrical a tree is; and (3) distribution of diversification shifts, i.e., occurrences of significant changes in rates of lineage splitting through time. In addition, we use character compatibility (e.g., Camin and Sokal 1965; Le Quesne 1969, 1982; Estabrook et al. 1976a,b; Meacham and Estabrook 1985) to examine how the structure of the data matrix yielding the trees of interest changes through successive time intervals. Empirical work has shown that addition fossils may alter hypotheses of relationships based on extant taxa only (e.g., Gauthier et al 1988; Cobbett et al. 2007), and simulation studies have revealed that such altered relationships may improve phylogenetic estimates (e.g., Huelsenbeck 1991; Wagner 2000b; Wagner and Sidor 2000), a conclusion that has been backed up by real case studies (e.g., Cunningham et al. 1988). To build on the metaphor of systematists living at different times in the past (see above), strictly extant taxa are simply one particular case of 'contemporaneous' taxa (i.e., taxa from a single time slice). Fossil-based phylogenies allow us to look at different sets of "contemporaneous" taxa, and permit comparisons between "contemporaneous only" vs. "fossil+contemporaneous" taxon sets. For the present work, we chose Ruta and Coates's (2007) phylogeny of early tetrapods (the limbed vertebrates). The monophyly of tetrapods is well established (Gaffney 1979; Panchen and Smithson 1987; Carroll 1991; Clack 2000, 2012). Early tetrapods consist of those limbed vertebrate groups that branch from the tetrapod stem and from the stems of each of the two major extant tetrapod radiations, the lissamphibians and the amniotes. Our use of early tetrapods is justified by the fact that their fossil record is extensive and diverse (Clack, 2012). Furthermore, there is renewed interest in the origin of limbed vertebrates and the patterns and processes underpinning terrestrialization. Notably, the origin of tetrapods represents the most recent of the major evolutionary transitions that led to the establishment of a fundamentally novel animal body plan (Clack 2002a, 2012). We emphasize that there is no agreement on the mutual relationships of various early tetrapod groups and on their affinities with either lissamphibians or amniotes. Although the debate is ongoing (for recent reviews and commentaries, see Anderson 2008 and Marjanović and Laurin 2013), it has little or no relevance to this paper, because we are more concerned with the issues of tree stability and its interpretation than we are with the specific implications of one hypothesized tetrapod phylogeny or another. The present contribution offers a set of protocols that can be used to validate some or all of the main conclusions presented here in light of future, more encompassing studies. In that respect, our approach should be seen as purely exploratory and the results from our investigation ought to be considered exclusively in light of the original findings in Ruta and Coates (2007). In summary, we chose Ruta and Coates (2007) because the taxon sample in that study is large enough to allow us to investigate clade stability over a relatively long time interval. We are aware that the study in question is neither the sole hypothesis of tetrapod interrelationships nor an exhaustive treatment of taxa. We also note that the lissamphibian radiation appears to be conspicuous only in the Mesozoic, and remains modest at the beginning of that era (Marjanović and Laurin 2014), so its impact is trivial for the case study presented here.

METHODS

Time Slicing and Phylogenetic Analyses

Harcourt-Brown (2002) examined changes in tree shape over a 28 myr time period using a foraminiferan tree. The tree was divided into a series of 500,000 year intervals. For any given interval, the relationships of taxa were derived from the original tree based on the presence of only those taxa that were present in that interval, and after manually pruning taxa outside that interval. Our approach also considers taxa that occur in specific time intervals, but differs from Harcourt-Brown's (2002) study because we ran separate phylogenetic analyses for each interval. Specifically, we explored changes in tree shape, relative to the original tree topology, not only through manual taxon pruning, but also by subjecting the taxa present in a given interval to a parsimony analysis.

The phylogenetic data set of Ruta and Coates (2007) includes 102 early tetrapod taxa coded for 339 characters (Nexus File #320 in the Paleobiology Database http://www.paleobiodb.org/cgi-bin/bridge.pl?a=viewNexusFile&nexusfile_no=320). Our reference topology is a relatively well-resolved strict consensus of 324 MPTs (1584 steps, CI = 0.22, RI = 0.67, RC = 0.15) resulting from a maximum parsimony analysis of all taxa. Taxa were assigned to five time intervals: Devonian (D), Mississippian (M), Pennsylvanian (P), Permian (R), and Mesozoic (Z) (see Fig. 1; Table 1). As early tetrapod diversity is unevenly distributed through time, a finer temporal subdivision would have resulted in intervals with low or no diversity, for which it would be difficult to construct a meaningful phylogeny, as well as intervals with disproportionately high diversity. As an additional simplification, we did not take into account differences in stratigraphic ranges within each time interval (e.g., Brocklehurst et al. 2015). The ranges of five taxa (*Edops*, *Chenoprosopus*, *Isodectes*, *Stegotretus*, *Diploceraspis*) cross the boundary between two intervals (Pennsylvanian-Permian) either because of uncertain age assignments or because of separate occurrences in adjacent intervals. Those taxa were treated as belonging to both intervals (see Appendix 1 for stratigraphic ranges of all ingroup taxa).

Our time slicing procedure yielded five non-cumulative data sets (hereafter referred to as “extant”), each consisting of taxa that occur solely in a specific interval (i.e., D, M, P, R, Z), as well as four cumulative data sets (hereafter referred to as “fossil+extant”), each consisting of taxa in any given interval plus all taxa occurring in preceding intervals (i.e., D+M, D+M+P, D+M+P+R, D+M+P+R+Z). The 'extant' trees can be likened to neontological phylogenies. Cumulative addition of intervals is likened to the total evidence practice of systematists who consider both extant and fossil taxa simultaneously. We excluded all characters that were uninformative in any given interval (both “extant” and“fossil+extant”), and we conducted maximum parsimony analysis using PAUP* v. 4.0b10 (Swofford 2003) on each of the nine data sets using the tree search protocol of Ruta et al. (2003a) (specifically, parsimony ratchet; see also Quicke et al. 2001). Multistate characters were left unordered. Although ordering may be recommendable in some cases, for instance when alternative states could plausibly be arranged in a morphocline sequence (e.g., Grand et al. 2013), we decided to impose minimum constraints on the relationships among states (i.e., the costs of transformations between non-adjacent states were left identical and equiprobable). Following the phylogenetic analyses, we computed a strict consensus topology for each interval. Finally, we compared the resulting nine consensus trees (hereafter, "re-analyzed trees") with reference consensus trees (hereafter, "pruned trees"). These pruned trees were obtained by manually pruning the strict consensus of Ruta et al. (2007) in MacClade v. 4.08 (Maddison and Maddison 2003), such that only taxa present in a given interval were retained.

The rationale behind this approach is that the taxa present in the pruned trees have the same mutual relationships as in the strict consensus. Conversely, the re-analyzed trees are built from smaller matrices obtained after removal of taxa from the original matrix; these smaller matrices may yield trees that differ from those obtained via the pruning procedure. Comparisons between the pruned trees and the re-analyzed trees allow us to determine the impact of taxon pruning on the topology of a "temporally driven" subsampled tree.

Measures of Tree Distance

To assess clade stability after applying time slicing, we examined the congruence between the pruned trees and the re-analyzed trees for each interval. Congruence between trees was assessed with two Tree Distance Metrics (TDMs): the *Partition Metric* (PM) and the *Triplets Based Distance Metric* (δ_{TM}s) (Page 1993) using Do not Conflict (DC) and Explicitly Agree (EA) distance criteria (Estabrook et al. 1985). These metrics represent trees as sets of simpler structures (e.g. partitions; triplets) and use different metrics to assess the similarity of those structures. EA only considers partitions that are both resolved and of the same type in order to represent similarities between trees, whereas DC also includes partitions that do not explicitly represent conflicts (Estabrook et al., 1985). The calculation of these metrics is easy compared to other metrics, such as transformation metrics (Boorman and Oliver 1973) and was carried out in Component Lite v. 0.1 (Page 1997; see Janzen et al. 2002; Pisani et al. 2007; Wollenberg et al. 2007 for recent similar studies). In addition, these metrics offer the advantage of being fairly intuitive and are appropriate for comparisons among tree topologies generated using a variety of methods, such as parsimony and manual pruning , as expounded above (but see Grand et al. 2013 for novel methods of tree shape comparisons). Because the various time slice trees have differing numbers of taxa, we followed Pisani's (2002) recommendations in applying normalized variants of the δ_{PM} and δ_{TMs} values, using two normalizing factors:

$$\varphi_r(\delta_{PM}) = 2\mathrm{n} - 4 \quad (1)$$

$$\varphi_R(\delta_{TMs}) = \frac{[n(n-1)(n-2)]}{6} \quad (2)$$

where n is the number of taxa in a given time slice. Normalized values vary between 0 and 1, and all of the trees we examined were rooted. We also subtracted normalized δ_{PM} and δ_{TMs} values from 1 to obtain indices of congruence relative to the 'true' target topology of Ruta and

Coates (2007). We used randomization tests to assess the statistical significance of the observed TDMs. The distribution of random simulated trees followed the Equal-Rates Markov (ERM) model (Simberloff et al. 1981; also see below), and we generated the null distribution by sampling all possible binary trees at random with 100 replications.

Finally, we used three parsimony-based tests to examine whether the re-analyzed trees fit the time sliced data sets better than the topologies generated by pruning the Ruta and Coates (2007) tree: the Kishino-Hasegawa test (Kishino and Hasegawa 1989), Templeton's (1983) implementation of the Wilcoxon signed-ranks test, and the Winning-sites (sign) test (Prager and Wilson 1988). In brief, the Kishino-Hasegawa test asks whether the steps from trees A and B represent two different normal distributions; Templeton's test examines whether the ranked steps from trees A and B represent two different distributions; and the Winning-sites test asks whether significantly more than half of the characters favor one tree. All three tests are implemented in PAUP* v. 4.0b10 (Swofford 2003), and we set the level of significance (α) at 0.05. The use of these tests in parsimony-based analyses has been criticized on the basis of circularity and violation of the null hypothesis (Goldman et al. 2000; Smith 2010) because the trees compared should be specified prior to the phylogenetic analysis, not after (as is usually the case). However, we consider them to be useful heuristic tools to examine the differences in tree length of various topologies given the data at hand, even if they lack true statistical rigor.

Analysis of Balance

Two parameters that are frequently used to describe the shape of a cladogram are balance – i.e. the degree of symmetry – and branch length – i.e. the expected amount of change between branching events, usually expressed in terms of number of character-state changes (Sanderson and Donoghue 1996). Here, we focus on tree balance (but see Hey 1992; Brown 1994). Balance is intuitive and easily interpreted (Harcourt-Brown 2001), and numerous indices have been proposed to measure it (Sackin 1972; Colless 1982; Shao and Sokal 1990; Heard 1992; Kirkpatrick and Slatkin 1993; Fusco and Cronk 1995; Rogers 1996; Mooers and Heard 1997; McKenzie and Steel 2000; Purvis et al. 2002). Here, we use Colless' index (*Ic*), as modified by Heard (1992), to measure balance. *Ic* is defined as:

$$Ic=\frac{\sum_{allinternalnodes} \left| T_R - T_L \right|}{\left[(n-1)(n-2)/2 \right]} \quad (3)$$

In a tree of *n* taxa, for every interior node the number of terminal taxa subtended by the right hand branch (T_R) and the number subtended by the left hand branch (T_L) are counted (Heard 1992). *Ic* is then calculated using (3): the normalizing factor bounds the values so they range from 0 (in the case of perfect balance) to 1 (in the case of complete imbalance). *Ic* is easy to calculate, its behavior is well known, and it gives normalized results that are comparable across all trees. Ideally, *Ic* should rely on a complete set of taxa (e.g., all taxa known to belong to a clade). As our case study phylogeny includes only a subset of taxa, the *Ic* values should be considered as if calculated on a whole-taxon topology, i.e. the 102-taxon sample in Ruta and Coates (2007) would represent the 'total' target topology.

We calculated *Ic* after the polytomies in strict consensus trees were resolved using the software SymmeTREE (Moore and Chan 2005). In SymmeTREE, the range of most and least symmetric dichotomous outcomes is approximated through the random resolution of polytomies using different underlying branching models. We used the taxon-size sensitive (TSS) equal-rates Markov algorithm because it is most conservative with respect to the null hypothesis that there was no significant diversification rate variation leading to unbalanced phylogenies (see Chan and Moore 2005 for further discussion), with 100,000 random resolutions generated for each tree. Because SymmeTREE assumes all polytomies to be soft, any genuine hard polytomies will be resolved (Chan and Moore 2005). We estimated *Ic* for the series of randomly resolved phylogenies as the arithmetic mean of the confidence intervals with upper (U_b) and lower bounds (L_b) corresponding to the tail probabilities for the 0.025 and 0.975 frequentiles, respectively.

We compared the observed indices with those associated with the equal-rates Markov (ERM) null model (Yule 1924). This model is based on a pure-birth (Markovian) branching process (usually bifurcation instead of budding cladogenesis) in which speciation and extinction rates are equally likely across all lineages (Simberloff et al. 1981; see Kirkpatrick and Slatkin 1993, Rogers 1994, Rogers 1996, Heard 1996). The ERM model as originally proposed is now often labeled as ERM-TS (equal-rates Markov time slice) model in order to distinguish it from the ERM-TI (equal-rates Markov time-inclusive) model proposed by Harcourt-Brown et al. (2001). Under the ERM-TS model, all branches have an equal chance of splitting at any time, and no probability of extinction is considered (Slowinski and Guyer 1989; Mooers and Heard

1997; Harcourt-Brown et al. 2001). Conversely, lineages under the ERM-TI model have an equal probability of splitting or extinction in each time step (Harcourt-Brown et al. 2001). Rogers (1994, 1996) calculated expected values of *Ic* for trees of varying taxon number under the ERM-TS model by growing trees by random branching and artificially terminating them after a given number of branching events in order to simulate the clade at a given time slice. Harcourt-Brown et al. (2001) demonstrated that the ERM-TS model is in fact relevant only to taxa from a single time slice (i.e., equivalent to neontological trees), and it is not applicable to cases where taxa have been selected from different time intervals, as in paleontological phylogenies. In order to deal with trees including taxa from multiple time slices, Harcourt Brown et al. (2001) introduced the ERM-TI model and they showed that the balance distribution of paleontological phylogenies fits the ERM-TI model extremely well. For both ERM models, as the number of terminal taxa increase, both the expected value of *Ic* and its standard deviation decrease very rapidly (Fig. 2). This is because the addition of taxa to the tree will, on average, increase balance because the proportion of completely imbalanced topologies will be much lower (Rogers 1996).

Given the different proprieties of ERM-TS and ERM-TI null models, we carried out two different kinds of comparisons of our tree balance data: 1) single time slices were treated in the same fashion as neontological phylogenies. Following Harcourt-Brown et al. (2001), we compared the value of *Ic* for these time slices to that expected from the ERM-TS model. 2) cumulatively added time slices were treated in the same fashion as paleontological phylogenies; we compared the value of *Ic* for these time slices to that expected from the ERM-TI model.

Diversification shifts

Although a number of non-biological factors can affect tree balance (Guyer and Slowinski 1991; Minelli et al. 1991; Fusco and Cronk 1995; Mooers 1995; Mooers et al. 1995; Heard and Mooers 1996; Huelsenbeck and Kirkpatrick 1996; Rannala et al. 1998; Pybus and Harvey 2000; Purvis and Agapow 2002; Huelsenbeck and Lander 2003), the analysis of balance is of intrinsic interest because it can provide insight into macroevolutionary patterns (Farris 1976; Slowinski and Guyer 1989; Heard 1992). Thus, asymmetric phylogenies are expected in cases where sister lineages diversify at different rates, whereas symmetric ones are expected when diversification rates are roughly equal across lineages (Kirkpatrik and Slatkin 1993). Based on these expectations, methods have been developed that use tree shape to infer shifts in diversification rates (Chan and Moore 2002, 2005; Moore et al. 2004), and these topology-based methods have been used in several contexts (e.g., McKenna and Farrell 2006; Ruta et al. 2007; Lloyd et al. 2008; Botha-Brink and Angielczyk 2010). Because the nature of the speciation process has been shown to be intrinsically stochastic (e.g., Raup et al. 1973; Gould et al. 1977), it is necessary to distinguish between chance variation in cladogram shape from variation which requires deterministic explanation when using topology-based methods for identifying diversification shifts (Chan and Moore 2002; see Mooers and Heard 1997 for a review), so the methods compare observed results to those obtained from a null model of random speciation.

Our analysis of diversification shifts focused on the pruned trees, particularly those showing cumulative addition of taxa over the five time slices. We carried out the tests with SymmeTREE (Moore and Chan 2005), which uses the equal-rates Markov (ERM) random-branching model (Yule 1924) as null model. This software performs several whole-tree tests on the relative diversity of all internal nodes of a given tree generalizing individual ERM nodal probabilities *P* (4) as:

$$P = \frac{2l}{N-1} \quad (4)$$

where *N* is the number of species from two sister groups, each consisting of *l* and *r* species, and where *l* is the number of species in the less diverse sister group (Chan and Moore 2002). *P* thus corresponds to the probabilities of having nodes with the observed level of asymmetry in the descendent lineages.

We also investigated temporal distribution of the diversification shift statistic (Δ_1 values in the SymmeTREE output), and of statistically significant ($p \leq 0.05$) and informative ($0.05 < p < 0.1$) shifts ($p\ \Delta_1$ values in the SymmeTREE output) across time slices. This statistic measures the difference in likelihood ratios between the inclusive and the nested node of a three-taxon statement under homogeneous and heterogeneous diversification models (for calculations, see Moore et al. 2004). We used ghost lineages and range extensions from the complete tetrapod phylogeny to date nodes in the time slice trees based on the following two rules. First, the minimum age of a node is taken to coincide with the age of the oldest taxon in the group subtended by that node. Second, if a taxon is present in a more recent time slice than the time slice considered, and if it forms the sister group to an older species or clade, then the range extension of that taxon in the time slice considered was taken to represent an occurrence *de facto* (i.e., the taxon was considered as if it was present).

After assigning ages to each internal node, we grouped Δ_1 values according to their ages, and we then compared Δ_1 value clusters within each time slice (e.g., Devonian values compared with Mississippian values within the D+M time slice) and across cumulatively added time slices (e.g., Devonian values in the D+M time slices compared with Devonian values in the D+M+P time slice) to determine whether diversification rates were significantly higher in particular time slices. We used one-way analysis of variance (ANOVA) to determine whether there was significant variation in diversification rates. In cases where significant variation was present we conducted pairwise comparisons between slices using Tukey's Honestly Significant Differences (HSD) test on pairwise comparisons of time slices to determine which time slices had significantly different rates. Since the distribution of our samples was unknown, we also ran non-parametric Wilcoxon Two-Sample tests on pairwise comparisons of time slices.

Character Compatibility

The previous tests focus on the topological effects of conducting phylogenetic analyses using taxa in single time slices or several time slices, but they do not provide information on potential changes in the structure of the underlying data matrices that presumably are responsible for those effects. Here we use character compatibility to determine how the structure of the character matrix changes from time slice to time slice, and with the cumulative addition of time slices. Two characters are compatible if a cladogram exists on which they can be optimized without homoplasy (Camin and Sokal 1965; Le Quesne 1969), and methods for deducing compatibility based on character state distributions without examining trees are available for several types of data, including binary and ordered multistate characters (Estabrook and Landrum 1975; McMorris 1975; Estabrook et al. 1976a, 1976b, Estabrook and McMorris 1980; Day et al. 1998). Compatibility has been used for several purposes in the context of phylogenetic studies (Meacham and Estabrook 1985;Wilkinson 2001). Our interest in compatibility stems from the fact that it can provide insight into the amount of homoplasy and hierarchical structure present in a given data set (Alroy 1994; Day et al. 1998), particularly because characters that change relatively infrequently tend to have higher compatibilities than those that change more frequently (O'Keefe and Wagner 2001).

We analysed compatibility on our “extant” and “fossil+extant” trees for each period using R (https://cran.r-project.org/; see Dryad repository for code and data). We excluded polymorphic codings from each of the “extant” and “fossil+extant” data sets and all invariant characters. With these modifications, the analyzed data sets ranged in size from six to 102 taxa and from 78 to 318 characters.

To put the incompatibilities in context, we simulated a null distribution for each period using the following procedure. First, we time-calibrated a complete tree using the cal3 method of Bapst (2013), where rates were arbitrarily chosen to keep the root age in the Devonian. Second, we randomly placed 1584 character changes along this phylogeny with the constraint that each character in the data matrix changed at least once; the probability of a character changing on a particular branch was proportional to the length of that branch. Third, we segmented this random tree into the different time slices (D, M, P, R, Z, D+M, D+M+P, D+M+P+R, D+M+P+R+Z) and computed the number of incompatible characters for each. Finally, we repeated this entire procedure 100 times to generate null distributions of incompatibility counts for each time bin.

Another set of experiments was devised to assess which two taxa are most incompatible in the data set, and was introduced to make sense of the particularly unstable position of one terminal taxon, *Lethiscus*, and one pair of sister taxa, Adelospondyli + *Acherontiscus*. For this experiment, we computed all of the possible pairs of taxa and removed them from the dataset, then compared the number of incompatibilities in the resulting datasets. This allowed us to compare all pairs with *Lethiscus* to all pairs without *Lethiscus*, so that we could assess whether *Lethiscus* had an unusually strong effect on incompatibility.

RESULTS

Phylogenetic Analyses and Measures of Tree Distance

Parameters of the re-analyzed trees can be found in Appendix 2. Results from comparisons using the Partition Metric (δ_{PM}) and the Triplets Based Distance Metrics (δ_{TMs}), as well as the results of the randomization tests, are summarized in Table 2. DC and EA δ_{TMs} returned nearly identical results, with only the comparisons between the Permian and Mesozoic time slices producing noteworthy differences (Permian: $\mathrm{DC}_{\mathrm{normalized}} = 0.12$, $\mathrm{EA}_{\mathrm{normalized}} = 0.33$; Mesozoic:

$DC_{normalized}$ = 0.00, $EA_{normalized}$ = 0.23). Because the DC and EA values generally agree, we calculated their means and focus on those in the following discussion and plots. Results of comparisons between trees obtained by cumulative addition of time slices are plotted in Figure 3, with stability quantified as an index of congruence (1 - δ_{PM} normalized values) plotted against time. Comparisons between single time slices are shown in Figure 4. Sample analyses in which we arbitrarily assigned the five taxa that cross the Pennsylvanian-Permian boundary to one of the two time slices did not show significantly different results.

The addition of taxa to the data set by means of cumulative addition of time slices results in a sigmoidal pattern for δ_{PM} (Fig. 3A), with an increase in congruence in the Mississippian, a plateau in the Pennsylvanian, an increase again in the Permian, and a new plateau in the Mesozoic. δ_{TMs} show a smoother pattern, with little difference between topologies through time (Fig. 3B). We also obtained different results for the two TDMs when we compared single time slices, with δ_{PM} distances emphasizing differences between trees. Devonian time slices were identical using both TDMs. No clear correlation links single time slice comparisons to the pattern of growth shown by the cumulative addition of time slices through time (Figs. 3-4).

Results of the three parsimony-based analyses conducted on the various sub-sets of the data matrix are shown in Table 3. In all cases but the Permian, both the re-analyzed “extant” trees and the re-analyzed “fossil+extant” trees fit the data significantly better than the pruned trees in all tests. However, very few taxa were relocated in the D+M (Fig. 5) and D+M+P (Fig. 6) phylogenies relative to the original consensus trees for the entire data matrix: the aïstopod

Lethiscus and the Adelospondyli + *Acherontiscus* clade were particularly unstable, and the position of the Pennsylvanian temnospondyl *Capetus* was resolved within other temnospondyls in the D+M+P tree. The phylogeny for the D+M+P+R data set (Fig. 7) was also nearly identical to the pruned tree, implying only a minor change very close to the tips of the tree (specifically, the positions of *Eoscopus* and *Platyrhinops* appear resolved within temnospondyls).

Analysis of Balance

Comparisons between *Ic* values generated under the null models and mean *Ic* values for the "extant" and the "fossil+extant" phylogenies (following random resolution of polytomies) are presented in Figures 8 and 9. *Ic* values for "fossil+extant" phylogenies (Fig. 8) all fall within the 95% confidence interval of the expected values under the ERM-TI model, but all are more imbalanced than expected. The phylogenies for the D+M and D+M+P data sets are more imbalanced than those of the successive cumulative data sets, with the balance of the D+M+P+R+Z phylogeny (*Ic* = 0.31) being the closest to the balance expected from the null model (*Ic* = 0.25).

The distribution of *Ic* values for the "extant" phylogenies (Fig. 9) show three different patterns through time when compared with values expected from the ERM-TS model. The Devonian phylogeny (i.e., 6 taxa, fully pectinate topology, *Ic* = 1) falls within the 95% confidence interval derived from the null model. However, we urge caution in interpreting these results because, with so few Devonian taxa in the phylogeny, it would be impossible to detect shifts, no matter how heavily reshuffled the taxa are. Both Carboniferous phylogenies fall outside the 95% confidence interval, being more imbalanced. The Permian and Mesozoic phylogenies fall well within the confidence interval.

Diversification Shifts

Table 3 shows the distribution of diversification shifts through time. No diversification shift was observed in the Devonian, but when successive time slices were cumulatively added, diversification shifts occurred at nodes dating to the Devonian, Mississippian and Pennsylvanian. No diversification shift was found among Permian and Mesozoic nodes, regardless of whether "extant" or "fossil+extant" intervals were considered. All shifts found in one time slice were retrieved for corresponding nodes when successive (i.e., more recent) time slices were added (Supplementary Data). Results of ANOVAs on variation in diversification shifts within each time slice are also presented in Table 3: p values indicate that the distribution of diversification shifts is not uniform through time (except for the D+M interval, but see discussion of shifts below).

The post hoc Tukey's HSD test did not find significant differences in any pairwise comparison, though this may be due to small variance differences between samples. Results of the Wilcoxon Two-Sample test are shown in Table 4. Pairwise comparisons found statistically significant differences in the rates of diversification between Devonian and Pennsylvanian, and between Mississippian and Pennsylvanian, in all the time slices where shifts of those ages were detected. There were no differences in diversification rate between Devonian and Mississippian in any interval (but see discussion of diversification shifts below).

Several statistically significant and informative $p\ \Delta_1$ values were found in the analysis (Table 4, Fig. 1), and all shifts were recovered in the same locations when successive time slices were added. No shifts were found when only the Devonian time slice was analyzed. Simulation of one branching event at random within this set of taxa did not lead to retrieval of significant shift within this tree. Note, however, that shifts along pectinate trees are more likely to occur with increasing numbers of taxa. In short, the tree might have to attain a certain threshold size before a shift can be recognized. Shifts D and C (letters correspond to labels in Fig. 1) show informative $p\ \Delta_1$ values when recovered for the first time during cumulative addition time slices D+M and D+M+P, respectively. Successive addition of time slices increased the statistical support for the shifts at nodes D and C ($p\ \Delta_1$ values < 0.05). Nine out of ten significant and informative shifts are located in the Carboniferous, seven of which are observed in the Mississippian. One shift is located at the boundary between the Devonian and Carboniferous.

Character Compatibility

Results from the character compatibility analysis are shown in Table 5 and Fig. 10. The total number of incompatibilities increases through time, because of novel taxon additions as progressively more recent time slices are added. Addition of more recent taxa are expected to increase incompatibility, e.g. due to introduction of conflicting states (e.g., reversals; losses) compared with earlier taxa. For the “extant” trees, the observed incompatibilities within each interval fall well within the null distribution that is expected given random character changes along the tree. However, for the “fossil+extant” trees, the observed incompatibilities are greater than the null distribution for the early bins (D+M and D+M+P) and substantially less than the null distribution for the latest time bins (D+M+P+R, D+M+P+R+Z). The fact that the observed incompatibility is higher than expected early on suggests rapid and sustained exhaustion of character states, with the later decrease suggesting introduction of new characters that are less homoplastic. The single bin results imply that, for those data sets, incompatibility does not increase more quickly (or slowly) than expected for the size of the datasets. We interpret the asymptotic shape of the increase as being due to the size of the datasets (i.e., in terms of number of taxa).

Experiments of removal of all taxon pairs from the matrix revealed that the stem frog *Triadobatrachus* and putative stem amniote *Caerorhachis* are the pair that, when removed, produce the most compatible overall dataset. *Triadobatrachus* shares several “absence” characters with various groups of early tetrapods. *Caerorhachis* shows a mosaic of primitive and derived characters, and its position relative to the dichotomy between amphibians and amniotes is particularly unstable (Clack 2012). Both taxa also receive a large number of unknown scores for several characters, due to inapplicable and unknown conditions. These results bear on our discussion of the unstable placements of *Lethiscus* and the Adelospondyli + *Acherontiscus* clade (see below; Figure 11).

DISCUSSION

Phylogenetic Analyses and Measures of Tree Distance

The Partition Metric analysis and the parsimony based tests highlighted important differences between time slices, which might indicated that our ability to reconstruct early tetrapod phylogeny changed over time. However, a detailed comparison between time slices and the results of the Triplets Based Distance Metrics showed that only minor topological changes occur through time and between single time slices. In general, most clades are extremely stable through time, with only two particularly unstable taxa (*Lethiscus* and the clade Adelospondyli + *Acherontiscus*) causing the observed differences. Therefore, the unstable placement of some tetrapods in Ruta and Coates' (2007) phylogeny in the re-analyzed trees may be better explained as a result of matrix properties and particular features of the taxa in question than a significant change in our ability to accurately reconstruct phylogeny at different points in the clade's history.

Lethiscus is a highly specialized long-bodied tetrapod without traces of limbs or girdles, and with a highly fenestrated skull that has lost most of the dermal cover and cheek bones (Milner 1994). Because of this unusual body plan, *Lethiscus* was coded with 222 out of 339 (65.5%) inapplicable (or unknown) entries in Ruta and Coates' (2007) data matrix. Coded characters concentrate in the skull table; in the postcranial skeleton, only few vertebral characters were coded, mostly concerning ornamental features. *Lethiscus* occupies a fairly derived position among D+M tetrapods in the pruned tree (Fig. 5a). However, in the tree resulting from re-analysis of D+M taxa only, *Lethiscus* appears on the tetrapod stem, in close proximity to a clade including (Adelospondyli + *Acherontiscus*) and the colosteid *Greererpeton* (note that *Acherontiscus* has been suggested to be an immature or paedomorphic adelospondyl; Ruta et al. 2003a, and references therein). From the D+M+P slice onward, *Lethiscus* clusters invariably with aïstopods; in Ruta and Coates' (2007) original analysis it is the most basal aïstopod, a position corroborated by several other analyses (e.g., Anderson 2001; Anderson et al. 2003; Ruta et al. 2003a). The joining of *Lethiscus* and Adelospondyli in the D+M tree likely reflects the fact that adelospondyls, like aïstopods, have elongated bodies, highly modified skulls with orbits placed far anteriorly on the skull (Clack 2002a) and no limbs (Ruta et al. 2003a). Unlike *Lethiscus* and other aïstopods, however, adelospondyls retained putative primitive characters such as a sculptured dermal skull roof and holospondylous vertebrae (Carroll 2001). Therefore, the unstable position of *Lethiscus* probably stems from a combination of missing data and homoplasy. It also emphasizes the potential impact of inadequate taxonomic sampling on phylogeny reconstruction (e.g., Cantino 1992; Wheeler 1992; Wheeler et al. 1993; Wiens 1998; Prendini 2001), and indicates that this can result from analyzing taxa from only a single time slice (such as would be the case for an extant taxa only analysis).

The other unstable clade is Adelospondyli + *Acherontiscus*. When we analyzed the time slice D+M+P, the Mississippian clade encompassing the adelogyrinids *Adelospondylus*, *Adelogyrinus*, and *Dolichopareias*, and the acherontiscid *Acherontiscus* moved from a stem group tetrapod position (where it is retrieved in all other time slices) to a total group amniote position as sister group of Nectridea (Fig. 6). This change presumably highlights the paucity of characters of adelospondyls that are uniquely shared with one or more specific tetrapod groups, as well as the highly divergent morphology of these animals. Adelospodyls display a mixture of (suggested) primitive and derived characters such as a temporal notch, relatively simple ribs, large dermal bones, and skull features reminiscent of those of Colosteidae (see Panchen and Smithson 1987). In other respects, such as the vertebral construction, they resemble lepospondyls such as microsaurs and lysorophids (Clack 2002a; Ruta et al. 2003a), and Ruta et al. (2003a) reconstructed adelospondyls nested within lepospondyls.

In this context, the Partition Metric distances and the results of the parsimony-based tests appear to sharpen what are in fact small differences between the pruned and re-analyzed time slice trees, creating a spurious pattern of conflict. In contrast, the use of the δ_{TMs} portrayed the phylogeny as very stable through time. Poor performance of δ_{PM} was previously noted by Penny and Hendy (1985), who showed that under certain conditions the Partition Metric can portray two trees differing solely in the position of few or even one taxon as maximally different. Our results for the parsimony based tests can be explained by the fact that the changes to the trees in question cause a great reshuffling of character states depending on the number of times features related to an elongate, limbless body plan are hypothesized to have evolved, despite the overall similarity of the rest of the topologies.

The time slice approach also may provide useful insight for helping resolve relationships among taxa in the face of saturation/character state exhaustion. For example, consider the

Pennsylvanian temnospondyl *Capetus* (Fig. 6), which possesses primitive features that are ubiquitous among other temnospondyls and autapomorphic characters of its own. Recent analyses have provided some improvement over the incertae sedis taxonomic status originally assigned to *Capetus* by Sequeira and Milner (1993) (e.g., Carroll 2001; Ruta et al. 2003a, b; Laurin and Soler-Gijón 2006; Ruta et al. 2007). Ruta and Coates' (2007) consensus tree placed *Capetus* in a polytomy within Temnospondyli. When we analyzed the D+M+P time slice, which includes only contemporaries of *Capetus* and older taxa, the position of *Capetus* was well-resolved: however, *Capetus* is positioned closer to amphibamids than to cochleosaurids, a result which is obviously at odds with our current understanding of this taxon. In succeeding time slices, this resolution is lost because new taxa with superficially similar but likely homoplastic morphologies are added to the analysis. This type of signal loss likely accounts for the unresolved position of Embolomeri + Eoherpetontidae among total group amniotes in the D+M+P time slice (Fig. 6).

Analysis of Balance

The cumulative time slice trees in our data set are all more imbalanced than expected under the null model. Many previous studies have found that published phylogenies reconstructed from empirical data are more imbalanced than predicted under the ERM model (Guyer and Slowinski 1991; Heard 1992; Mooers 1995; Purvis and Agapow 2002; Holman 2005; Blum and François 2006; Heath et al. 2008), but all these studies used the ERM-TS as their null. According to

Harcourt-Brown et al.'s (2001, p. 199) ERM-TI model, "the range of balances generated by [the model] is so great that it is unlikely that any tree will fall outside two standard deviations of the expected value, even if significant evolutionary processes that affect tree shape occurred". Therefore, the fact that two cumulative time slices (i.e., D+M and D+M+P; see Fig. 8) deviate strongly from the null expectation is surprising. This result appears to stem from the Devonian taxa included in the analysis. Although we did not compare the Devonian time slice to the ERM-TI null model (since single time slices were compared with the ERM-TS model), the Devonian taxa are noteworthy because they occupy a very basal position on the tree and appear in a fully pectinate topology. They also form a relatively large proportion of included taxa in the D+M and D+M+P time slices, causing the phylogenies for these slices to be relatively imbalanced. The addition of more recent taxa in subsequent time slices increases the balance of the cladogram, bringing the *Ic* closer to that predicted by the ERM-TI model. The fact that the Devonian portion of the tree is so imbalanced is certainly due to a sampling artifact (i.e., paucity of taxa), and a meaningful assessment of the contribution of these taxa to tree shape must await inclusion of new Devonian tetrapods.

Ic values for 'extant' time slices differ from those of the 'fossil+extant' time slices because the tree shape for each time slice is independent from those of preceding time slices. As explained above, 'extant' time slices are also more similar to neontological phylogenies, making comparison to the ERM-TS model appropriate. Previous studies suggested that neontological phylogenies are, on average, more imbalanced than expected under ERM-TS model (see Mooers and Heard 1997 for a detailed discussion). However, we found that only the two Carboniferous time slices (i.e., Mississippian and Pennsylvanian) fall outside the 95% confidence interval of the expected value (Fig. 9). This difference implies that diversification rates, extinction rates, or preservation potential were variable among tetrapod lineages at the time, such that relatively imbalanced trees are obtained when preceding and succeeding taxa are excluded from consideration.

Uneven diversification rates and preservation potentials among Carboniferous tetrapod lineages would not be surprising. The Carboniferous Period saw the first expansion of tetrapods into fully terrestrial habitats, as well as their radiation into a wealth of body forms and ecologies (Clack 2002a). Two subsequent major episodes of diversification within these tetrapods (a stem-lissamphibian radiation, and a stem-amniote radiation; Ruta et al. 2006) also probably occurred during this time span (although uncertainty exists related to different proposed hypotheses of relationships among early tetrapods). Numerous Carboniferous forms retained aquatic adaptations but also possessed characters associated with increased terrestriality (Holmes 1980; Clack 2002b), perhaps related to the acquisition of progressively more terrestrial habitats (Ruta et al. 2003a). Morphological variation was demonstrated to have been unusually high in the early history of tetrapods (e.g., Shubin et al. 2004), and the frequencies of character change also appear to have been high (Ruta et al. 2006). Finally, new food resources for early tetrapods were represented by the increasing diversity of land plants and various groups of invertebrates, arthropods in particular (DiMichele and Hook 1992; Shear and Selden 2001).

Diversification shifts

Cumulative addition of time slices did not affect the relative locations of the significant and informative diversification shifts found in the reference cladogram. This behavior is interpreted, in the context of testing the impact of temporal constraints on phylogenetic analysis, as highlighting a point of stability in the reference phylogeny.

Most of the shifts are Mississippian in age, and W-statistics and Δ_1 values for these nodes are consistent as taxa from succeeding time slices are added. We found no shifts in the Permian and Mesozoic, although few nodes of these ages are present (e.g., the temnospondyl sample in Ruta and Coates (2007) includes some of the best-known Permian and Carboniferous members of the clade but omits for practical reasons the large, mostly Mesozoic stereospondyl radiation). There were also no shifts in diversification when only the Devonian time slice was analyzed, which is not surprising because the topology of this tree is fully pectinate and the number of taxa is small. When successive time slices are added, we found six Devonian shifts in diversification, one for each internal Devonian node, all of which are located at the very base of the tree. Shifts at these nodes are the result of a tree topology in which a few taxa arranged in a pectinate fashion join a more complex topology with dozens of branching linages, and may be spurious because of low taxon sampling among the basal lineages. Even though the Δ_1 values for the Devonian nodes are very low, and far from being statistically significant ($p = 0.48$), they cause the Wilcoxon Two Sample Test to fail in identifying a statistically significant difference between Devonian and

Mississippian diversification shifts. Once again, this may reflect paucity of finds from these time intervals as well as inclusion of only the best documented taxa from relevant groups. Even when the test was conducted with $\Delta_1 = 0$ for the Devonian nodes, it yielded a statistically informative ($W = 375$, $p = 0.0672$) difference between Devonian and Mississippian diversification shifts. An ANOVA test (that for two samples is a simple Student's *t*-test) also did not identify a significant difference ($F = 0.8526$, $p = 0.3600$). Although these results call into question the reality of the Devonian diversification shifts, we suggest that this hypothesis be tested in the future by adding other basal taxa, such as osteolepiform fish, to the analysis.

Of the remaining diversification shifts, Shift A (Fig. 1) captures the initial post-Devonian radiation of early tetrapods. Notably, this event appears to separate two different faunas because Devonian and Carboniferous taxa are not interleaved on the tree. The Carboniferous radiation appears to have been a unique episode in early tetrapod evolutionary history, with Ruta and Coates (2007) suggesting that all post-Devonian tetrapods may have evolved from a single taxon. The unusual nature of this diversification event is supported by our results because it is found in all time slices and has the largest magnitude of all identified. Various schemes of tetrapod phylogeny – beginning with Coates (2006) – have presented the Devonian *Tulerpeton* as the sister taxon to all post-Devonian tetrapods. However, this scenario has been challenged repeatedly (e.g. Ruta and Bolt 2006), and *Tulerpeton* has been placed at times in a basal position relative to whatcheeriids, a primarily Carboniferous clade (e.g. Lombard and Bolt 1995; Clack and Finney 2005; Warren 2007).

Shift B occurs within stem tetrapods. This shift is of particular interest because it highlights an increase in speciation rate among forms that are still primarily aquatic, although it is close to the first primarily terrestrial radiation represented by *Whatcheeria* and *Pederpes* (Ruta and Clack 2006).

Shift C is placed after *Eucritta*, a taxon showing a mixture of characters found in baphetids, temnospondyls and anthracosaurids that subtends the origin of the total groups Amphibia and Amniota. This shift began the radiation among early tetrapods that eventually led to the emergence of the modern tetrapod fauna (i.e., true amniotes and amphibians, and their closest relatives).

Shift D is associated with the origin of total group amniotes. Its position is important because it is soon after their divergence from total group amphibians. This radiation is also strictly related to the invasion of land: basal amniotes show a pattern of character acquisition that reflects progressive, rapid adaptation to life on land (Ruta and Clack 2006), even if certain amniote lineages returned to the aquatic environment in the group's early history (e.g., some seymouriamorphs, microsaurs, nectrideans) (e.g. see Bossy and Milner 1998).

Shifts E, F and G characterize radiations among stem amniotes that are associated with important skeletal modifications for terrestrial life (see below). Shifts E and F are close to the base of crown amniotes, and subtend the origin of very terrestrially adapted groups such as the Seymouriamorpha and Diadectomorpha. Notably, diadectomorphs include some of the earliest high-fiber tetrapod herbivores (Hotton et al. 1996; Sues and Reisz 1998; Reisz and Sues 2000;

Reisz, 2006). Shift G characterizes the radiation of Lepospondyli, a highly diverse and heterogeneous clade showing extreme variability in features such as the number of vertebrae, pattern and/or number of cranial bones, and presence/absence of limbs.

Shift H pinpoints a diversification that takes place within Nectridea, between the basal scincosaurid *Scincosaurus* and the holospondyls. Scincosaurids previously were interpreted as showing primitive conditions within the nectrideans, especially with respect to many cranial features (Bossy and Milner 1998). Their box-like skulls are very different from the diversity of shapes that evolved within the other members of the clade, and may represent the basic shape that gave rise to the urocordylids, aïstopods and the more derived nectrideans. *Scincosaurus* was probably amphibious (Milner 1980), with the subsequent diversification producing both aquatic (e.g., urocordylideans) and semi-aquatic/terrestrial forms (e.g., aïstopods) (Laurin et al. 2004).

Shifts I and J are well within the total groups Amphibia and Amniota, and highlight radiations among two important groups of early tetrapods: Temnospondyli (the most abundant and diverse of all groups of basal tetrapods; Ruta et al. 2003a) and Microsauria (an extremely diverse group that reached a position of dominance during the Carboniferous; Carroll and Gaskill, 1978). Interestingly, shift I subtends the radiation of many amphibamid temnospondyls, a successful group of miniaturized dissorophoids showing predominantly terrestrial adaptations (Laurin et al. 2004). Since these shifts are both in derived positions within the tree it will be important to test if they are retrieved in the same positions when taxonomic sampling is increased.

Results of the diversification shift analysis presented here strongly support the view that the Mississippian, particularly the early Mississippian, was a time of extensive radiation among basal tetrapods, with all of the major lineages originating at this time and accompanied by increasing morphological disparity (Garcia et al. 2006). However, the 30-million-year hiatus ('Romer's Gap') at the base of the Carboniferous fossil record hinders our understanding of how this event unfolded (Coates and Clack 1995). Similarly, all of the Mississippian localities represent aquatic environments only (with the notable exception of the hot spring volcaniclastic sediments of East Kirkton; Rolfe et al. 1994) increasing the unevenness of our knowledge. Although new discoveries are slowly filling this gap in the history of early tetrapods (e.g., Clack 2002b) indirect analyses such as our diversification shift analysis still have a key role to play in developing hypotheses about this diversification that can be tested as new fossil data are collected.

Character Compatibility

The results of our character compatibility analysis show that levels of character compatibility (and therefore homoplasy) did not differ significantly from the amounts expected by random character stae changes distributed across the trees. We suspect that this is an artifact of the exclusion of many taxa from the individual time slices. The resulting loss of information makes it difficult to accurately assess when apparent compatibilities between characters are genuine versus when spurious compatibilities arise because the number and polarities of character state changes in various lineages are obscured by inadequate taxon sampling. In the cumulative trees, however, incompatibility (and thus homoplasy) was high relative to null expectations early in the history of tetrapods, but much lower later in their history. Our null distributions assume an equal rate of character change through time, so these offsets could be explained by declining rates of morphological evolution. High early rates produce higher-than-expected levels of incompatibility, possibly reflective of character state exhaustion (i.e., new taxa are more likely to possess novel combinations of exisiting characeter states than genuinely new states; see also Wagner, 2000a), whereas lower rates produce lower-than-expected levels of incompatibility. Angielczyk and Ruta (2012) found that skull shape disparity was strongly correlated with diversity in Permo-Carboniferous temnospondyls, implying that newly evolved species possessed distinctive new skull shapes, perhaps reflecting a similar underlying process for generating morphological novelty.

The change in sampling through time may also generate a signal of lower-than-expected homoplasy later in the history of tetrapods. Radiations such as the lepospondyls and lissamphibians are poorly sampled in our phylogeny. If these tetrapod subclades underwent their own dramatic diversifications with high rates of character change, but have fewer characters scored in our dataset, the lower rates of homoplasy later could be due to declining rates only in the overall backbone of the phylogeny.

CONCLUSIONS

Analysis of Ruta and Coates' (2007) early tetrapod phylogeny using a time slice approach found that our ability to reconstruct relationships among basal tetrapod clades remains almost unchanged through time. A few shifts in the position of some taxa and clades through time are explained in terms of matrix properties and particular features of the considered taxa.

The time slice approach (i.e., the dissection of the phylogenetic analysis into time bins) permits a better understanding of the relationships of a group at particular intervals in its evolutionary history, and shows considerable potential for exploring the behavior of phylogenies when taxa are added in the chronological order in which they appear in the fossil record. A particularly significant case in the application of this method is represented by the addition of fossils to data sets of extant organisms. Studies of the mutual positions of taxa at specific time intervals may thus be of interest to paleontologists working on groups with long evolutionary histories, and to neontologists studying clades with extensive fossil records. Changes in the position of taxa through time may be due to the effects imparted by the subsequent history of a group (e.g., convergent evolution of similar character suites), although methodological problems (i.e., matrix properties) cannot be ruled out. The methodology highlights the importance of comprehensive taxon and character inclusion in phylogenetic analysis, as well as detailed consideration of how particular taxa and characters influence the results of the phylogenetic analysis. However, it is important to note that such analyses will need to consider the idiosyncracies of the underlying data set(s), so the extrapolation of general observations may be difficult. As suggested by Harcourt-Brown (2002), the time-slice approach provides a valuable quantitative method for identifying non-stochastic evolutionary processes, with the unique possibility of pinpointing particular times in a clade's history during which taxa underwent non-random speciation events. Cumulative addition of time slices and comparisons of resulting topologies with those generated by an ERM-TI model extends this technique and may provide important clues about the macroevolutionary history of a clade. However, the wide 95% confidence interval around tree balance index values prevent statistical support for non-randomness in most topologies. For this reason, non-stochastic distributions of balance through time slices are considered to be informative and have clearly interpretable meanings in our example of tetrapod phylogeny.

Specifically, tree balance in Carboniferous time slices was found to differ from expected values, emphasizing the remarkable diversification that tetrapods underwent at this time (Clack 2012).

The overall stability of early tetrapod phylogeny is supported by the analysis of shifts in diversification; addition of successive (i.e., more recent) time slices does not affect the distribution of diversification shifts. The Mississippian was found to be a time of explosive radiation in early tetrapod evolution, corroborating previous analyses. The importance of the Early Carboniferous likely stems from ecological and environmental factors, especially increasing terrestrialization of the first limbed vertebrates.

Character incompatibility was high in the early history of tetrapods, but declined to be much lower than expected relative to a null distribution based on a single rate of character change. This result is consistent with models of evolutionary radiations where clades have higher rates early in their history, producing greater than expected homoplasy. Subsequent radiations, such as the origin of lissamphibians in the Triassic, may “reset” the rates and produce new, novel forms with higher rates in new regions of character space. Removing *Triadobatrachus* and *Caerorhacis* from the data matrix produced the largest decrease in the total number of incompatibilities (compared to all other pairs of taxa).

ACKNOWLEDGEMENTS

We thank D. Pisani, J.E. Tarver, and S.J. Braddy (University of Bristol) for discussion and advice, A. Minelli (University of Padova) and M. Menegon (MUSE, Trento) for helpful comments. Reviews by P. Wagner, M. Laurin and an anonymous reviewer, and comments by Editor F. Anderson and Associate Editor B. Dayrat greatly improved the quality of this paper.

REFERENCES

Alroy J. 1994. Four permutation tests for the presence of phylogenetic structure. Syst. Biol. 43:430–437.

Anderson J S. 2001. The phylogenetic trunk: Maximal inclusion of taxa with missing data in an analysis of the Lepospondyli. Syst. Biol. 50:170–193.

Anderson J.S. 2008. The origin(s) of modern amphibians. Evol. Biol. 35:231–247.

Anderson J.S., Carroll R.L., Rowe T.B. 2003. New information on *Lethiscus stocki* (Tetrapoda: Lepospondyli: Aïstopoda) from high-resolution computed tomography and a phylogenetic analysis of Aistopoda. Can. J. Earth Sci. 40:1071–1083.

Angielczyk K.D., Ruta M. 2012. The roots of amphibian morphospace: a geometric morphometric analysis of Paleozoic temnospondyls. Fieldiana; Life Earth Sci. 5:40–58.

Benjamini Y., Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J.R. Stat. Soc. S. B 57:289–300.

Blum M.G.B., François O. 2006. Which random processes describe the Tree of Life? A large-scale study of phylogenetic tree imbalance. Syst. Biol. 55:685–691.

Boorman S.A., Oliver D.C. 1973. Metrics on spaces of finite trees. J. Math. Psych. 10:26–59.

Bossy K.V.H., Milner A.C. 1998. Order Nectridea. In: Carroll R.L., Bossy K.A., Milner A.C., Andrews S.M. and Wellstead C.F. editors. Encyclopedia of Paleoherpetology - Lepospondyli. Stuttgart: Gustav Fischer Verlag. p. 73–131.

Botha-Brink J., Angielczyk K.D. 2010. Do extraordinarily high growth rates in Permo-Triassic dicynodonts (Therapsida, Anomodontia) explain their success before and after the end-Permian extinction? Zool. J. Linn. Soc. 160:341–365.

Bremer K. 1988. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. Evolution 42:795–803.

Brocklehurst N., Ruta M., Müller J., Fröbisch J. 2015. Elevated extinction rates as a trigger for diversification rate shifts: early amniotes as a case study. Sci. Rep. 5:17104.

Brown J.K.M. 1994. Probabilities of evolutionary trees. Syst Biol 43:78–91.

Camin J.H., Sokal R.R. 1965. A method for deducing branching sequences in phylogeny. Evolution 19:311–326.

Cantino P.D. 1992. Evidence for a polyphyletic origin of the Labiatae. Ann Miss Bot Gard 79:361–379.

Carroll R.L. 1991. The origin of reptiles. In: Schultze H.-P. and Trueb L., editors. Origins of the higher groups of tetrapods: Controversy and consensus. Ithaca: Cornell University Press, p. 331–353.

Carroll R.L. 2001. The origin and early radiation of terrestrial vertebrates. Journal of Paleontology 75:1202–1213.

Carroll R.L., Gaskill P. 1978. The order Microsauria. Memoirs of the American. Phil Soci 126:1–211.

Chan K.M.A., Moore B.R. 2002. Whole-tree methods for detecting differential diversification rates. Syst Biol 51:855–865.

Chan K.M.A. 2005. SymmeTREE: Whole-tree analysis of differential diversification rates. Bioiformatics 21:1709–1710.

Clack J.A. 2000. The origin of tetrapods. In: Heatwole H., Carroll R., editors. Amphibian biology, Vol. 4, Palaeontology: The evolutionary history of amphibians. Chipping Norton: Surrey Beatty & Sons, p. 979–1029.

Clack J.A. 2002a. Gaining ground: The origin and early evolution of tetrapods. Bloomington: Indiana University Press.

Clack J.A. 2002b. An early tetrapod from 'Romer's Gap'. Nature 418:72–76.

Clack J.A. 2012. Gaining ground: The origin and early evolution of tetrapods (2nd Edition). Bloomington: Indiana University Press.

Clack J.A., Ahlberg P.A., Finney S.M., Dominguez A.P., Robinson J., Ketcham R.A. 2003. A uniquely specialized ear in a very early tetrapod. Nature 425:65–69.

Clack J.A., Finney S.M. 2005. *Pederpes finneyae*, an articulated tetrapod from theTournaisian of Western Scotland. J. Syst. Palaeo. 2:311–46.

Coates M.I. 1996. The Devonian tetrapod *Acanthostega gunnari* Jarvik: postcranial anatomy, basal tetrapod interrelationships and patterns of skeletal evolution. Trans. R. Soc. Edinburgh: Earth Env. Sci. 87:363–422.

Coates M.I., Clack J.A. 1995. Romer's gap: Tetrapod origins and terrestriality. Bull. Mus. Natio. Hist. Natur. (Paris) 17:373–388.

Cobbett A., Wilkinson M., Wills M. A. 2007. Fossils impact as hard as living taxa in parsimony analyses of morphology. Syst. Biol. 56:753–766.

Cohen K.M., Finney S.M., Gibbard P.L., Fan J.-X. 2013. The ICS International Chronostratigraphic Chart. Episodes 36:199–204.

Colless D.H. 1982. Review of: Phylogenetics: The theory and practice of phylogenetic systematics. Syst. Zool. 31:100–104.

Colless D.H. 1996. A further note on symmetry of taxonomic trees. Syst. Biol. 45:385–395.

Cunningham C.W., Zhu H., Hillis D. M. 1998. Best-fit maximum-likelihood models for phylogenetic inference: empirical tests with known phylogenies. Evolution 52:978 – 987

Davis J.I. 1993. Character removal as a means for assessing stability of clades. Cladistics 9:201–210.

Davis J.I., Frohlich M.W., Soreng R.J. 1993. Cladistic characters and cladogram stability. Syst. Bot. 18:188–196.

Day W.H.E., Estabrook G.F., McMorris F.R. 1998. Measuring the phylogenetic randomness of biological data sets. Syst. Biol. 47:604–616.

Didier G., Royer-Carenzi M., Laurin, M. 2012. The reconstructed evolutionary process with the fossil record. J. Theor. Biol. 315:26–37.

DiMichele W.A., Hook R.W. 1992. Paleozoic terrestrial ecosystems. In: Behrensmeyer A.K., Damuth J.D., DiMichele W.A., Potts R., Sues H.-D. and Wing S. editors. Terrestrial ecosystems through time: Evolutionary paleoecology of terrestrial plants and animals. Chicago: Chicago University Press, p. 205–235.

Estabrook G.F., Johnson C.S.Jr., McMorris F.R. 1976a. A mathematical foundation for the analysis of cladistic character compatibility. Math. Biosci. 29:181–187.

Estabrook G.F., Johnson C.S.Jr., McMorris F.R. 1976b. An algebraic analysis of cladistic characters. Disc. Math. 16:141–147.

Estabrook G.F., Landrum L. 1975. A simple test for the possible simultaneous evolutionary divergence of two amino acid positions. Taxon 25:609–613.

Estabrook G.F., McMorris F.R. 1980. When is one estimate of evolutionary relationships a refinement of another? J. Math. Biol. 10:367–373.

Estabrook G.F, McMorris F.R., Meacham C.A. 1985. Comparison of undirected phylogenetic trees based on subtrees of four evolutionary units. Syst. Zool. 34:139–200.

Faith D.P., Ballard J.W.O. 1994. Length differences and topology-dependent tests: A response to Kållersïo et al. Cladistics 10:57–64.

Farris J.S. 1976. Expected asymmetry of phylogenetic trees. Syst. Zool. 25:196–198.

Farris J.S., Albert V.A., Kållersïo M., Lipscomb D. 1996. Parsimony jackknifing outperforms neighbour-joining. Cladistics 12:99–124.

Felsenstein J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. Syst. Zool. 27:401–410.

Felsenstein J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39:783–791.

Fusco G., Cronk Q. C. B. 1995. A new method for evaluating the shape of large phylogenies. J.

Theor. Biol. 175:235–243.

Gaffney E.S. 1979. Tetrapod monophyly: A phylogenetic analysis. Bull. Carnegie Mus. Nat. Hist. 13:92–105.

Garcia W.J., Storrs G.W., Greb S.F. 2006. The Hancock County tetrapod locality: A new Mississippian (Chesterian) wetlands fauna from western Kentucky (USA). In: Greb S.F. and DiMichele W.A., editors. Wetland through time. The Geological Society of America Special Paper 399, p. 155-168

Gatesy J. 2000. Linked branch support and tree stability. Syst. Biol. 49:800–807.

Gauthier J.A., Kluge A.G., Rowe T. 1988. Amniote phylogeny and the importance of fossils. Cladistics 4:105–209.

Goldman N., Anderson J.B., Rodrigo A.G. 2000. Likelihood-based tests of topologies in phylogenetics. Syst. Biol. 49:652–670.

Goloboff P.A. 1991. Homoplasy and the choice among cladograms. Cladistics 7:215–232.

Good-Avila S.V., Souza V., Gaut B.S., Eguiarte L.E. 2006. Timing and rate of speciation in *Agave* (Agavaceae). Proc. Nat. Acad. Sci. USA 103:9124–9129.

Gould S.J., Raup D.M., Sepkoski J. J. Jr., Schopf T.J.M., Simberloff D.S. 1977. The shape of evolution: A comparison of real and random clades. Paleobiology 3:23–40.

Grand A., Corvez A., Duque Velez L.M., Laurin M. 2013. Phylogenetic inference using discrete characters: performance of ordered and unordered parsimony and of three-item statements. Biol. J. Linn. Soc. 110:914–930.

Guyer C., Slowinski J.B. 1991. Comparison of observed phylogenetic topologies with null expectations among three monophyletic lineages. Evolution 45:340–350.

Guyer C., Slowinski J. B. 1993. Adaptive radiation and the topology of large phylogenies. Evolution 47:253–263.

Harcourt-Brown K. G. 2001. Phylogenetic tree shape with special reference to the phylogeny of the Cretaceous globotruncanid Foraminifera. Ph.D. Thesis. Dept. Earth Sciences, University of Bristol (UK).

Harcourt-Brown K.G., Pearson P.N., Wilkinson M. 2001. The imbalance of paleontological trees. Paleobiology 27:188–204.

Harcourt-Brown K.G. 2002. Tree balance, time slices, and evolutionary turnover in Cretaceous planktonic foraminifera. Syst. Biol. 51:908–916.

Heard S.B. 1992. Patterns in tree balance among cladistic, phenetic and randomly generated phylogenetic trees. Evolution 46:1818–1826.

Heard S.B. 1996. Patterns in phylogenetic tree balance with variable and evolving speciation rates. Evolution 50:2141–2148.

Heard S.B., Mooers A.Ø. 1996. Imperfect information and the balance of cladograms and phenograms. Syst. Biol. 45:115–118.

Heath T.A., Zwickl D.J., Junhyong K., Hillis D.M. 2008. Taxon sampling affects inferences of macroevolutionary processes from phylogenetic trees. Syst. Biol. 57:160–166.

Hey J. 1992. Using phylogenetic trees to study speciation and extinction. Evolution 46:627–640.

Holman E.W. 2005. Nodes in phylogenetic trees: The relation between imbalance and number of descendent species. Syst. Biol. 54:895–899.

Holmes R. 1980. *Proterogyrinus scheelei* and the early evolution of the labirinthodont pectoral limb. In: Panchen A.L., editor. The terrestrial environment and the origin of land vertebrates. The Systematics Association Special Volume 15. New York: Academic Press, p. 351–376

Hotton N. III, Olson E. C., Beerbower R. 1996. Amniote origins and the discovery of herbivory. In: Sumida S. and Martin K.L.M., editors. Amniote Origins. New York: Academic Press, p. 207-264.

Huelsenbeck J.P. 1991. When are fossils better than extant taxa in phylogenetic analysis? Syst. Zool. 40:458–469.

Huelsenbeck J.P., Hillis D.M. 1993. Success of phylogenetic methods in the four-taxon case. Syst. Biol. 42:247–264.

Huelsenbeck J.P., Kirkpatrick M. 1996. Do phylogenetic methods produce trees with biased shapes? Evolution 50:1418–1424.

Huelsenbeck J.P., Lander K.M. 2003. Frequent inconsistency of parsimony under a simple model of cladogenesis. Syst. Biol. 52:641–648.

Huss J. 2009. The shape of evolution: the MBL model and clade shape. In: Sepkoski D. and Ruse M., editors. The paleobiological revolution: essays on the growth of modern paleontology. Chicago: University of Chicago Press, p. 326-345.

Janzen F.J., Krenz J.G., Haselkorn T.S., Brodie E. D. Jr., Brodie E.D. 2002. Molecular phylogeography of common garter snakes (*Thamnophis sirtalis*) in western North America: Implications for regional historical forces. Mol. Ecol. 11:1739–1751.

Kirkpatrick M., Slatkin M. 1993. Searching for evolutionary patterns in the shape of a phylogenetic tree. Evolution 47:1171–1181.

Kishino H., Hasegawa M. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. J. Mol. Evol. 29:170–179.

Kållersïo, M., J. S. Farris, A. G. Kluge, and C. Bult. 1992. Phylogenetic analysis of the Cichorioideae (Asteraceae), with emphasis on the Mutisieae. Annals of the Missouri Botanical Garden 79:416–427.

Laurin M., Girandot M., Loth M.-M. 2004. The evolution of long bone microanatomy and lifestyle in lissamphibians. Paleobiology 30:589–613.

Le Quesne W.J. 1969. A method of selection of characters in numerical taxonomy. Syst. Zool. 18:201–205.

Le Quesne W.J. 1982. Compatibility analysis and its applications. Zool. J. Linn. Soc. 74:267–275.

Lloyd G.T., Davis K.E., Pisani D., Tarver J.E., Ruta M., Sakamoto M., Hone D.W.E., Jennings R., Benton M.J. 2008. Dinosaurs and the Cretaceous terrestrial revolution. Proc. R. Soc. B: Biol. Sci. 275:2483–2490.

Lombard E., Bolt J.R. 1995. A new primitive tetrapod *Whatcheeria deltae* from the Lower

Carboniferous of Iowa. Palaeontology 38:471–494.

Maddison D.R., Maddison W.P. 2003. MacClade 4: Analysis of phylogeny and character evolution. Version 4.0. Sinauer Associates, Sunderland, Massachusetts.

Maddison W.P. 1989. Reconstructing character evolution on polytomous cladograms. Cladistics 5:365–377.

Marjanović D., Laurin M. 2013. The origin(s) of extant amphibians: a review with emphasis on the lepospondyl hypothesis. Geodiversitas 35:207–272.

Marjanović, D., Laurin M. 2014. An updated palaeontological timetree of lissamphibians, with comments on the anatomy of Jurassic crown-group salamanders (Urodela). Historical Biology 26:535–550.

McKenna D.D., Farrell B.D. 2006. Tropical forests are both evolutionary cradles and museums of leaf beetle diversity. Proc. Nat. Acad. Sci. USA 103:10947–10951.

McKenzie A., Steel M. 2000. Distributions of cherries for two models of trees. Math. Biosci. 164:81–92.

McMorris F.R. 1975. Compatibility criteria for cladistic and qualitative taxonomic characters. In: Estabrook B.F. editor. Proceedings of the Eighth Annual Freeman International Congress on Numerical Taxonomy. San Francisco: W. H. Freeman, p. 399-415.

Meacham C.A., Estabrook G.F. 1985. Compatibility methods in systematics. Ann. Rev. Ecol. Syst. 16:431–446.

Milner A.C. 1980. The tetrapod assemblage from Nýrany, Czechoslovakia. In: Panchen A.L.

editor. The terrestrial environment and the origin of land vertebrates. The Systematics Association Special Volume 15. New York: Academic Press, p. 439-496.

Milner A.C. 1994. The aïstopod amphibians from the Viséan of East Kirkton, West Lothian, Scotland. Trans. R. Soc. Edinburgh 84:363–368.

Minelli A., Fusco G., Sartori S. 1991. Self-similarity in biological classifications. BioSystems 26:89–97.

Mooers A.Ø. 1995. Tree balance and tree completeness. Evolution 49:379–384.

Mooers A.Ø., Heard S.B. 1997. Inferring evolutionary processes from phylogenetic tree shape. Quat. Rev. Biol. 72:31–54.

Mooers A.Ø., Page R.D.M., Purvis A., Harvey P.H. 1995. Phylogenetic noise leads to unbalanced cladistic tree reconstructions. Syst. Biol. 44:332–342.

Moore B.R., Chan K.M.A. 2005. SymmeTREE: An application for performing whole-tree tests of diversification rate variation. Version 1.0.

Moore B.R., Chan K.M.A., Donoghue M.J. 2004. Detecting diversification rate variation in supertrees. In: Bininda-Emonds O. R. P. editor, Phylogenetic Supertrees: Combining Information to Reveal the Tree of Life. Dordrecht: Kluwer Academic, p. 487-533

O'Keefe F.R., Wagner P.J. 2001. Inferring and testing hypotheses of cladistic character dependence by using character compatibility. Syst. Biol. 50:657–675.

Page R.D.M. 1993. Component Version 2.0 Software and User's Guide. Biogeography and Conservation Laboratory. The Natural History Museum, London (UK).

Page R.D.M. 1997. Component lite, v0.1. Distributed by the author. Department of Zoology, University of Glasgow (UK).

Panchen A.L., Smithson T.R. 1987. Character diagnosis, fossils and the origin of tetrapods. Biol. Rev. Cambridge Phil. Soc. 62:341–438.

Parham, J. F., Donoghue, P.C.J., Bell, C.J. et al. 2011. Best practices for justifying fossil calibrations. Syst. Biol. 61:346–359.

Penny D., Hendy M.D. 1985. The use of trees comparison metrics. Syst. Zool. 34:75–82.

Pisani D. 2002. Comparing and combining data and trees in phylogenetic analysis. Ph.D. Thesis. Dept. Earth Sciences, University of Bristol (UK).

Pisani D., Benton M.J., Wilkinson M. 2007. Congruence of morphological and molecular phylogenies. Acta Biotheoretica 55:269–281.

Prager E.M., Wilson A.C. 1988. Ancient origin of lactalbumin from lysozyme: Analysis of DNA and amino acid sequences. J. Mol. Evol. 27:326–335.

Prendini L. 2001. Species or supraspecific taxa as terminals in cladistic analysis? Groundplans versus exemplars revisited. Syst. Biol. 50:290–300.

Purvis A., Agapow P. -M. 2002. Phylogeny imbalance: Taxonomic level matters. Syst. Biol. 51:844–854.

Purvis A., Katzourakis A., Agapow P.-M. 2002. Evaluating phylogenetic tree shape: Two modifcations to Fusco and Cronk's method. J. Theor. Biol. 214:99–103.

Pybus O.G., Harvey P.H. 2000. Testing macro-evolutionary models using incomplete molecular phylogenies. Proc. R. Soc. Lon. B: Biol. Sci. 267:2267–2272.

Quicke D.L.J., Taylor J., Purvis A. 2001. Changing the landscape: A new strategy for estimating large phylogenies. Syst. Biol. 50:60–66.

Rannala B., Huelsenback J.P., Yang Z., Nielsen R. 1998. Taxon sampling and the accuracy of large phylogenies. Syst. Biol. 47:702–710.

Raup D.M., Gould S.J., Shopf T.J.M., Simberloff D. S. 1973. Stochastic models of phylogeny and the evolution of diversity. J. Geol. 81:525–542.

Reisz R.R. 2006. Origin of dental occlusion in tetrapods: Signal for terrestrial vertebrate evolution. J. Vert. Paleo. 26:114A–114A.

Reisz R.R., Sues H. D. 2000. Herbivory in Late Paleozoic and Triassic terrestrial vertebrates. In: H. D. Sues editor, Evolution of Herbivory in Terrestrial Vertebrates. Cambridge: Cambridge University Press, p. 9–41.

Rogers J.S. 1994. Central moments and probability distribution of Colless' coefficient of tree imbalance. Evolution 48:2026–2036.

Rogers J.S. 1996. Central moments and probability distribution of three measures of phylogenetic tree imbalance. Syst. Biol. 45:99–100.

Rolfe W.D.I., Clarkson E.N.K., Panchen A. L. 1994. Volcanism and early terrestrial biotas. Trans. R. Soc. Edinburgh: Earth Sci. 84.

Ruta M., Bolt J.R. 2006. A reassessment of the temnospondyl amphibian *Perryella olsoni* from the Lower Permian of Oklahoma. Earth Environ. Sci. Trans. R. Soc. Edinburgh 97:113–165.

Ruta M, Clack J. A. 2006. A review of *Silvanerpeton miripedes*, a stem amniote from the Lower Carboniferous of East Kirkton, West Lothian, Scotland. Trans. R. Sci. Edinburgh: Earth Sci. 97:31–63.

Ruta M., Coates M.I. 2007. Dates, nodes and character conflict: Addressing the lissamphibian origin problem. J. Syst. Pal. 5:69–122.

Ruta M., Coates M. I., Quicke D. L. J. 2003a. Early tetrapod relationships revisited. Biol. Rev. 78:251–345.

Ruta M., Jeffery J.E., Coates M.I. 2003b. A supertree of early tetrapods. Proc. R. Soc. Lon. B: Biol. Sci. 270:2507–2516.

Ruta M., Wagner P.J., Coates M.I. 2006. Evolutionary patterns in early tetrapods. I. Rapid initial diversification followed by decrease in rates of character change. Proc. R. Soc. Lon. B: Biol. Sci. 273:2107–2111.

Ruta M., Pisani D., Lloyd G.T., Benton M. J. 2007. A supertree of Temnospondyli: cladogenetic patterns in the most species-rich group of early tetrapods. Proc. R. Soc. Lon. B: Biol. Sci. 274:3087–3095.

Sackin M.J. 1972. 'Good' and 'Bad' phenograms. Syst. Zool. 21:225–226.

Sanderson M. J. 1994. Reconstructing the history of evolutionary processes using maximum likelihood. In: Fambrough D.M., editor. Molecular evolution of physiological processes. New York: Rockefeller University Press, p.13-26.

Sanderson M. J., Donoghue M.J. 1996. Reconstructing shifts in diversification rates on phylogenetic trees. Trends Ecol. Evol. 11:15–20.

Savage H. M. 1983. The shape of evolution: Systematic tree topology. Biol. J. Linn. Soc. 20:255–244.

Schopf T.J.M. 1979. Evolving paleontological views on deterministic and stochastic approaches. Paleobiology 5:337–352.

Sequeira S.E.K., Milner A.R. 1993. The temnospondyl amphibian *Capetus* from the Upper Carboniferous of the Czech Republic. Palaeontology 36:657–680.

Shao K., Sokal R. R. 1990. Tree balance. Syst. Zool. 39:266–276.

Shear W.A., Selden P.A. 2001. Rustling in the undergrowth: Animals in early terrestrial ecosystems. In: Gensel P.G. and Edwards D. editors. Plants invade the land: Evolutionary and environmental perspectives. New York: Columbia University Press, p. 29–51.

Shubin N.H., Daeschler E.B., Coates M.I. 2004 The early evolution of the tetrapod humerus. Science 304:90–93.

Sigurdsen T., Bolt J.R. 2010. The Lower Permian amphibamid *Doleserpeton* (Temnospondyli: Dissorophoidea), the interrelationships of amphibamids, and the origin of modern amphibians. J. Vert. Paleont. 30:1360–1377.

Simberloff D., Heck K.L., McCoy E.D., Connor E. F. 1981. There have been no statistical tests of cladistic biogeographic hypotheses. In: Nelson G. and Rosen D. E., editors. Vicariance in biogeography: A critique. New York: Columbia University Press, p. 40–63

Slowinski J.B., Guyer C. 1989. Testing the stochasticity of patterns of organismal diversity: An improved null model. Am. Nat. 142:1019–1024.

Smith N.D. 2010. Phylogenetic analysis of Pelecaniformes (Aves) based on osteological data: Implications for waterbird phylogeny and fossil calibration studies. PLOSOne 5(10):e13354.

Stadler T. 2010. Sampling-through-time in birth-death trees. J. Theor. Biol. 267:396–404.

Sues H.-D., Reisz R.R. 1998. Origin and early evolution of herbivory in tetrapods. Trends Ecol. Evol. 13:141–145.

Swofford D.L. 2003. PAUP*: Phylogenetic Analysis Using Parsimony (* and other methods). Version 4.0b10. Sinauer Associates, Sunderland.

Tarver J.E., Donoghue P.C.J. 2011. The trouble with topology: phylogenies without fossils provide a revisionist perspective of evolutionary history in topological analyses of diversity. Syst. Biol. 60:700–712.

Templeton A. 1983. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. Evolution 37:221–244.

Wagner P.J. 2000a. Exhaustion of cladistic character states among fossil taxa. Evolution 54:365–386.

Wagner P.J. 2000b. The quality of the fossil record and the accuracy of phylogenetic inferences about sampling and diveristy. Syst. Biol. 49:65–86.

Wagner P.J, Ruta M., Coates M.I. 2006. Evolutionary patterns in early tetrapods. II. Differing constraints on available character space among clades. Proc. R. Soc. B: Biol. Sci. 273:2113–2118.

Wagner, P.J., Sidor C.A. 2000. Age rank : clade rank metrics - sampling, taxonomy, and the meaning of “stratigraphic consistency”. Syst. Biol. 49:463–479.

Warren A. 2007. New data on *Ossinodus pueri*, a stem tetrapod from the Early Carboniferous of Australia. J. Vert. Paleont. 27:850–862.

Wheeler W.C. 1992. Extinction, sampling, and molecular phylogenetics. In: Novacek M. and Wheeler Q., editors. Extinction and phylogeny. New York: Columbia University Press, p. 205–215.

Wheeler W.C., Cartwright P., Hayashi C. 1993. Arthropod phylogeny: A combined approach. Cladistics 9:1–39.

Wiens J.J. 1998. The accuracy of methods for coding and sampling higher-level taxa for phylogenetic analysis: A simulation study. Syst. Biol. 47:397–413.

Wilkinson M. 1997. On phylogenetic relationships within *Dendrotriton* (Amphibia: Caudata: Plethodontidae): is there sufficient evidence? Herpet. J. 7:55–65.

Wilkinson M. 2001. PICA 4.0: software and documentation. Department of Zoology, The Natural History Museum, London.

Wollenberg K.C., Glaw F., Meyer A., Vences M. 2007. Molecular phylogeny of Malagasy reed frogs, *Heterixalus*, and the relative performance of bioacoustics and color-patterns for resolving their systematics. Mol. Phyl. Evol. 45:14–22.

Yule G.U. 1924. A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis. Phil. Trans. R. Soc. Lon. B: Biol. Sci. 213:21–87.

Appendix 1: Taxonomic sample (from Ruta and Coates 2007)

Taxa		**Period/s**
Unranked basal taxa	*Acanthostega gunnari* Jarvik 1952	D
	Ichthyostega stensioei Säve-Söderbergh 1932	D
	Tulerpeton curtum Lebedev 1984	D
	Ventastega curonica Ahlberg et al. 1994	D
	Caerorhachis bairdi Holmes and Carroll 1977	M
	Crassigyrinus scoticus Watson 1929	M
	Eucritta melanolimnetes Clack 1998.	M
	Ossinodus pueri (Warren and Turner 2004)	M
	Pederpes finneyae (Clack 2002)	M
	Westlothiana lizziae Smithson and Rolfe 1990	M
	Whatcheeria deltae Lombard and Bolt 1995	M
Acherontiscidae	**Acherontiscidae**	
	Acherontiscus caledoniae Carroll 1969b	M
Adelospondyli	**Adelogyrinidae**	
	Adelospondylus watsoni Carroll 1967	M
	Adelogyrinus simorhynchus Watson 1929	M
	Dolichopareias disjectus Watson 1929	M
Aïstopoda	**Lethiscidae**	
	Lethiscus stocki Wellstead 1982	M
	Ophiderpetontidae	
	Oestocephalus amphiuminum Cope 1868	P
	Phlegethontiidae	
	Phlegethontia linearis Cope 1871	P
Baphetidae	**Baphetidae**	
	Baphetes kirkbyi Watson 1929	P
	Megalocephalus pachycephalus (Barkas 1873)	P
Colosteidae	**Colosteidae**	
	Colosteus scutellatus (Newberry 1856)	P
	Greererpeton burkemorani Romer 1969	M
Diadectomorpha	**Diadectidae**	
	Diadectes absitus Berman et al. 1998	R
	Orobates pabsti (Berman et al. 2004)	R
	Tseajaia campi (Moss 1972)	R
	Limnoscelidae	
	Limnoscelis paludis Williston 1911	R
	Solenodonsaurus janenschi Broili 1924	P
Embolomeri	**Anthracosauridae**	
	Anthracosaurus russelli Huxley 1863	P
	Silvanerpeton miripedes (Clack 1994)	M

	Archeriidae	
	Archeria crassidisca (Cope 1884)	R
	Eogyrinidae	
	Pholiderpeton attheyi (Watson 1926)	P
	Pholiderpeton scutigerum Huxley 1869	P
	Eoherpetontidae	
	Eoherpeton watsoni Panchen 1975	M
	Proterogyrinus scheelei Romer 1970	M
Gephyrostegidae	**Gephyrostegidae**	
	Bruktererpeton fiebigi Boy and Bandel 1973	P
	Gephyrostegus bohemicus Jaekel 1902	P
Lysorophia	**Cocytinidae**	
	Brachydectes elongatus Wellstead 1991	P
	Brachydectes newberryi Cope 1868	
Microsauria	**Brachystelechidae**	
	Batropetes fritschia (Geinitz and Deichmüller 1882)	R
	Hapsidopareiontidae	
	Hapsidopareion lepton Daly 1973	R
	Saxonerpeton geinitzi (Credner 1890)	R
	Goniorhynchidae	
	Rhynchonkos stovalli (Olson 1970).	R
	Gymnarthridae	
	Cardiocephalus sternbergi Broili 1904a	R
	Euryodus primus Olson 1939	R
	Hyloplesiontidae	
	Hyloplesion longicostatum Fritsch 1876	P
	Microbrachidae	
	Microbrachis pelikani Fritsch 1876	P
	Odonterpetontidae	
	Odonterpeton triangulare Moodie 1909	P
	Ostodolepidae	
	Micraroter erythrogeios Daly 1973	R
	Pelodosotis elongatum Carroll and Gaskill 1978	R
	Pantylidae	
	Pantylus cordatus Cope 1871	R
	Stegotretus agyrus Berman et al. 1988	P-R
	Tuditanidae	
	Asaphestera intermedia (Dawson 1894)	P
	Tuditanus punctulatus Cope 1874	P
Nectridea	**Diplocaulidae**	
	Batrachiderpeton reticulatum (Hancock and Atthey 1869)	R
	Diceratosaurus brevirostris (Cope 1875)	P
	Diplocaulus magnicornis Cope 1882	R
	Diploceraspis burkei Romer 1952	R
	Keraterpeton galvani Wright and Huxley 1866	P
	Scincosauridae	
	Scincosaurus crassus Fritsch 1876	P
	Urocordylidae	

	Ptyonius marshii Cope 1875	R
	Sauropleura Cope 1868	P
	Urocordylus wandesfordii Wright and Huxley 1866	P
Seymouriamorpha	**Discosauriscidae**	
	Ariekanerpeton sigalovi (Ivakhnenko 1981)	R
	Discosauriscus austriacus (Makowsky 1876)	R
	Utegenia shpinari (Kuznetsov and Ivakhnenko 1981)	R
	Kotlassiidae	
	Kotlassia prima Amalitsky 1921	R
	Leptorophidae	
	Leptoropha talonophora (Tchudinov 1955)	R
	Microphon exiguous (Ivakhnenko 1983)	R
	Seymouriidae	
	Seymouria Broili 1904b	R
Temnospondyli	**Amphibamidae**	
	Amphibamus grandiceps Cope 1865	P
	Doleserpeton annectens Bolt 1969	R
	Eoscopus lockardi Daly 1994	P
	Platyrhinops lyelli (Wyman 1858)	P
	Branchiosauridae	
	Apateon pedestris Meyer 1844	R
	Leptorophus tener (Schönfeld 1911)	R
	Schoenfelderpeton prescheri Boy 1986	R
	Cochleosauridae	
	Chenoprosopus lewisi Hook 1993	P-R
	Cochleosaurus florensis Rieppel 1980	P
	Dendrerpetontidae	
	Dendrerpeton acadianum Owen 1853	P
	Dissorophidae	
	Broiliellus brevis Carroll 1964	R
	Ecolsonia cutlerensis Vaughn 1969	R
	Edopidae	
	Edops craigi Romer 1935	R
	Eobrachyopidae	
	Isodectes obtusus (Cope 1868)	R
	Eryopidae	
	Eryops megacephalus Cope 1877	R
	Micromelerpetontidae	
	Micromelerpeton credneri Bulman and Whittard 1926	R
	Trematopidae	
	Phonerpeton pricei (Olson 1941)	R
	Trimerorhachidae	
	Neldasaurus wrightae Chase, 1965	R
	Trimerorhachis cfr. *insignis* Case, 1935	R
Family incertae sedis	*Balanerpeton woodi* Milner and Sequeira 1994	M
	Capetus palustris (Sequeira and Milner 1993)	P
Crown-group Lissamphibia	*Albanerpeton inexpectatum* Estes and Hoffstetter 1976	M
	Eocaecilia micropoda Jenkins and Walsh 1993	M

	Karaurus sharovi Ivakhnenko 1978	M
	Triadobatrachus massinoti (Piveteau 1936)	M
	Valdotriton gracilis Evans and Milner 1996	M
	Notobatrachus degiustoi Reig 1955	M
	Vieraella herbsti (Bàez and Basso 1996)	M
Crown-group Amniota	**Captorhinidae**	
	Captorhinus aguti Cope 1882	R
	“Protorothyrididae”	
	Paleothyris acadiana Carroll 1969a	P
	Araeoscelidia	
	Petrolacosaurus kansensis Lane 1945	P

APPENDIX 2: PAUP* ANALYSES

Maximum parsimony analysis in PAUP* v. 4.0b10 (Swofford 2003) for single and cumulative time slices gave the following results:

Devonian: 1 tree found, 106 steps long, with Consistency Index (CI) = 0.85 (excluding uninformative characters), Retention Index (RI) = 0.82 and Rescaled Consistency Index (RC) = 0.78.

Mississippian: 3 trees found, 316 steps long, with Consistency Index (CI) = 0.52 (excluding uninformative characters), Retention Index (RI) = 0.57 and Rescaled Consistency Index (RC) = 0.35.

Pennsylvanian: 12 trees found, 616 steps long, with Consistency Index (CI) = 0.39 (excluding uninformative characters), Retention Index (RI) = 0.64 and Rescaled Consistency Index (RC) = 0.27.

Permian: 177 trees found, 698 steps long, with Consistency Index (CI) = 0.34 (excluding uninformative characters), Retention Index (RI) = 0.66 and Rescaled Consistency Index (RC) = 0.25.

Mesozoic: 2 trees found, 102 steps long, with Consistency Index (CI) = 0.69 (excluding uninformative characters), Retention Index (RI) = 0.64 and Rescaled Consistency Index (RC) = 0.52.

Devonian+Mississippian: 3 trees found, 457 steps long, with Consistency Index (CI) = 0.48

(excluding uninformative characters), Retention Index (RI) = 0.61 and Rescaled Consistency Index (RC) = 0.32.

Devonian+Mississippian+Pennsylvanian: 192 trees found, 986 steps long, with Consistency Index (CI) = 0.32 (excluding uninformative characters), Retention Index (RI) = 0.64 and Rescaled Consistency Index (RC) = 0.22.

Devonian+Mississippian+Pennsylvanian+Permian: 54 trees found, 1450 steps long, with Consistency Index (CI) = 0.24 (excluding uninformative characters), Retention Index (RI) = 0.67 and Rescaled Consistency Index (RC) = 0.17.

Devonian+Mississippian+Pennsylvanian+Permian+Mesozoic: 324 trees, 1584 steps long, with Consistency Index (CI) = 0.22 (excluding uninformative characters), Retention Index (RI) = 0.67 and Rescaled Consistency Index (RC) = 0.15 (note that this is a re-analysis of the whole Ruta and Coates' (2007) matrix; the same number of MPTs and the same parameters were found).

Table legends

TABLE 1. The five time slices, their ages, and the number of taxa which have their range in those slices. The number of taxa considered in this study is 102: five taxa cross the boundary between two time slices, and have been considered as belonging to both the time slices (ages from Cohen et al. 2013).

TABLE 2. Comparison between manually pruned and computer re-run time slice trees performed using Partition Metric (δ_{PM}) and two Triplets Based Distance Metrics (δ_{TMs}) measures of distance: DCs (do not conflict) and EA (explicitly agree). Results for the randomization tests (100 replications) and for the three different parametric tests are given in p values. D = Devonian, M = Mississippian, P = Pennsylvanian, R = Permian, Z = Mesozoic.

TABLE 3. Distribution of Δ_1 values (i.e., number of shifts per age) across cumulative time slices and results of ANOVA test, and Wilcoxon Two Sample Test's results of pairwise comparisons of Δ_1 values across time slices. In brackets Δ_1 values $\neq 0$. Statistical significant values highlighted in bold. Dashes identify no node of that age present. Devonian time slice excluded since all nodes are of the same age (i.e., Devonian), and no shifts in diversification was found. D = Devonian, M = Mississippian, P = Pennsylvanian, R = Permian.

TABLE 4. Significant ($p \leq 0.05$) and informative ($0.05 < p < 0.1$) p Δ_1 values found in each cumulatively added time slice. Shifts are labeled with a capital letter as in Fig. 1 and in the text. Age of the shifts is also given. Values corrected for multiple testing (FDR, Benjamini and Hochbert 1995) are provided in italics. Dashes identifies no shift present. D = Devonian, M =

Mississippian, P = Pennsylvanian, R = Permian.

TABLE 5. Character compatibility analysis for both single and cumulative time slices described by Incompatibility count.

TABLE 1

Time bin	Age (Ma)	N of taxa
Devonian	419.2 – 358.9	6
Mississippian	358.9 – 323.2	17
Pennsylvanian	323.2 – 298.9	34
Permian	298.9 – 252.2	39
Mesozoic	252.2 – 66.0	7

TABLE 2

Time slices	**N of taxa/time slice**	**δPM**	**δPM normalized**	**1- δPM normalized**	**δPM randomization**	**δTMs**		**δTMs normalized**		**1 – δTMs normalized**		**δTMs randomization**		**Templeton test**	**Winning-sites test**	**Kishino-Hasegawa test**
						DC	EA	DC	EA	DC	EA	DC	EA			
D	6	0	0.00	1.00	0.01	0	0	0.00	0.00	1.00	1.00	0.01	0.01	1	1	1
M	17	18	0.60	0.40	0.01	102	102	0.15	0.15	0.85	0.85	0.01	0.01	<0.05	0.052	<0.05
P	36	17	0.25	0.75	0.01	71	142	0.01	0.02	0.99	0.98	0.01	0.01	<0.05	<0.05	<0.05
R	41	29	0.37	0.63	0.01	1279	3518	0.12	0.33	0.88	0.67	0.01	0.01	0.210	0.101	0.201
Z	7	1	0.10	0.90	0.01	0	8	0.00	0.23	1.00	0.77	0.01	0.01	<0.05	0.125	<0.05
D+M	23	18	0.43	0.57	0.01	106	106	0.06	0.06	0.94	0.94	0.01	0.01	<0.05	0.087	<0.05
D+M+P	59	43	0.37	0.62	0.01	4226	4876	0.13	0.15	0.87	0.85	0.01	0.01	<0.05	<0.05	<0.05
D+M+P+R	95	2	0.01	0.99	0.01	0	0	0.00	0.00	1.00	1.00	0.01	0.01	<0.05	<0.05	<0.05

TABLE 3

Time slices	D+M	D+M+P	D+M+P+R	Whole
Shifts per age				
Devonian	12 (6)	12 (6)	12 (6)	12 (6)
Mississippian	73 (14)	63 (19)	65 (21)	61 (21)
Pennsylvanian	-	79 (0)	75 (7)	76 (7)
Permian	-	-	33 (0)	32 (0)
Mesozoic	-	-	-	12 (0)
ANOVA test				
F statistic	0.0862	8.297	5.969	5.147
p value	0.809	**0.000381**	**0.000671**	**0.000588**
Pairwise comparison				
Devonian - Mississippian	W = 615.0 p = 0.209	W = 483.5 p = 0.686	W = 473.5 p = 0.7249	W = 464.5 p = 0.7546
Devonian - Pennsylvanian		no Pennsylvanian Δ_1 values $\neq 0$	W = 697.0 p = **0.0369**	W = 706.0 p = **0.0359**
Mississippian - Pennsylvanian		no Pennsylvanian Δ_1 values $\neq 0$	W = 4625.0 p = **0.0175**	W = 4680.5 p = **0.0147**

TABLE 4

Shift	Age of the shift	D+M	D+M+P	D+M+P+R	Whole
A	Devonian-Mississippian	**Significant** (0.04) *(0.05)*	**Significant** (0.03) *(0.05)*	**Significant** (0.03) *(0.05)*	**Significant** (0.03) *(0.05)*
B	Mississippian	Informative (0.07) *(0.07)*	Informative (0.07) *(0.07)*	Informative (0.07) *(0.07)*	Informative (0.07) *(0.07)*
D	Mississippian	Informative (0.07) *(0.07)*	Informative (0.06) *(0.07)*	**Significant** (0.05) *(0.06)*	**Significant** (0.04) *(0.05)*
C	Mississippian	-	**Significant** (0.03) *(0.05)*	**Significant** (0.03) *(0.05)*	**Significant** (0.03) *(0.05)*
F	Mississippian	-	**Significant** (0.04) *(0.05)*	**Significant** (0.04) *(0.05)*	**Significant** (0.04) *(0.05)*
G	Mississippian	-	**Significant** (0.03) *(0.05)*	**Significant** (0.04) *(0.05)*	**Significant** (0.04) *(0.05)*
H	Mississippian	-	**Significant** (0.04) *(0.05)*	**Significant** (0.01) *(0.05)*	**Significant** (0.01) *(0.05)*
E	Mississippian	-	-	**Significant** (0.04) *(0.05)*	**Significant** (0.04) *(0.05)*
J	Pennsylvanian	-	-	**Significant** (0.04) *(0.05)*	**Significant** (0.04) *(0.05)*
I	Pennsylvanian	-	-	-	Informative (0.08) *(0.08)*

TABLE 5.

Data set	Incompatibility count
Devonian	35
Mississippian	1296
Pennsylvanian	4744
Permian	5523
Mesozoic	116
D+M	2997
D+M+P	9498
D+M+P+R	13940
Whole	15503

FIGURE CAPTIONS

FIGURE 1. Chronogram of early tetrapod relationships, proposed by Ruta and Coates (2007), showing stratigraphic distribution, regarded as the reference tree for this study. Taxa whose ranges cross the boundary between two time slices are figured twice. In bold, branches along which significant diversification rate shifts supported by p Δ_1 values were found; shifts are also labeled with capital letters (see text for discussion). Grey brackets at the top highlight the three main groups of tetrapods represented in the cladogram.

FIGURE 2. Comparison between the expected distributions of the ERM-TS and ERM-TI models. The continuous line represents the expected values of *Ic* under the ERM-TS model. The dotted lines represents lower and upper bounds (two standard deviations) of the expected value. The short dashed line represents the expected values of *Ic* under the ERM-TI model. The two long dashed lines represent two standard deviations above and below the expected value. Data for the ERM-TS model from Rogers (1994); data for the ERM-TI model thanks to Dr J. E. Tarver.

FIGURE 3. 'Growth' of the Ruta and Coates' (2007) early tetrapod phylogeny as obtained by comparison between cumulatively added re-run time slice trees and manually pruned time slices. Comparisons obtained using the Partition Metric (δ_{PM}) are shown in A; comparisons obtained using the Triplet Based Distance Metrics (δ_{TMs}) are shown in B. D = Devonian, M = Mississippian, P = Pennsylvanian, R = Permian.

FIGURE 4. Single time slices comparison as obtained by comparison between re-run consecutive time slice trees and pruned time slices from Ruta and Coates' (2007) early tetrapod phylogeny.

Comparisons obtained using the Partition Metric (δ_{PM}) are shown in a; comparisons obtained using the Triplet Based Distance Metrics (δ_{TMs}) are ahown in b. D = Devonian, M = Mississippian, P = Pennsylvanian, R = Permian, Z = Mesozoic.

FIGURE 5. Comparison between Devonian + Mississippian (D+M) time slices: pruned (A) and re-run tree (B). Relocated taxa highlighted in bold; note *Lethiscus* in particular (see text for discussion).

FIGURE 6. Comparison between Devonian + Mississippian + Pennsylvanian (D+M+P) time slices: pruned (A) and re-run tree (B). Relocated taxa highlighted in bold; note (Adelospondyli + *Acheronthiscus*) in particular; see also the position of *Capetus* among Temnospondyli (see text for discussion).

FIGURE 7. Comparison between Devonain + Mississippian + Pennsylvanian + Permian (D+M+P+R) time slices: pruned (A) and re-run tree (B). Relocated taxa highlighted in bold. *Eoscopus* and *Platyrhinops* polytomy becomes resolved; they become more closely related to the other Amphibamidae then to the Branchiosauridae (see text for discussion).

FIGURE 8. Comparison between *I*c for cumulatively added time slices and those expected from the ERM-TI model. Continuous line represent values for early tetrapod time slices. Dotted line represent expected values under the null model with dashed lines being two standard deviations from that value. D = Devonian, M = Mississippian, P = Pennsylvanian, R = Permian.

FIGURE 9. Comparison between *I*c for single, consecutive time slices and those expected from the ERM-TS model. Continuous line represent values for early tetrapods time slices. Dashed

lines represent two standard deviations from the value expected under the null model. D = Devonian, M = Mississippian, P = Pennsylvanian, R = Permian, Z = Mesozoic.

FIGURE 10. Comparison of the observed incompatibilities (gray) to the null distributions of character incompatibilities generated after evolving 318 characters along 100 randomly time scaled phylogenies (width of black bars corresponds to the number of simulated trees with that an incompatibility count falling into each bin; all spindles are based on 100 different trees). Each character was constrained to change at least once, and the total number of changes was set to be 1584 steps (the length of the maximum parsimony tree). The null distributions were generated by taking each time period and slicing the tree there to produce extant (A) and fossil+extant (B) trees, then analysing the simulated incompatibilities.

FIGURE 11. Distribution of incompatibilities when each of the possible taxon pairs are removed. Bars in gray represent the overall distribution, while bars in black represent those pairs without *Lethiscus*. The extreme left value (indicated by the arrow) represents the number of incompatibilities when *Triadobatrachus* and *Caerhachis* are removed.

FIGURE 1

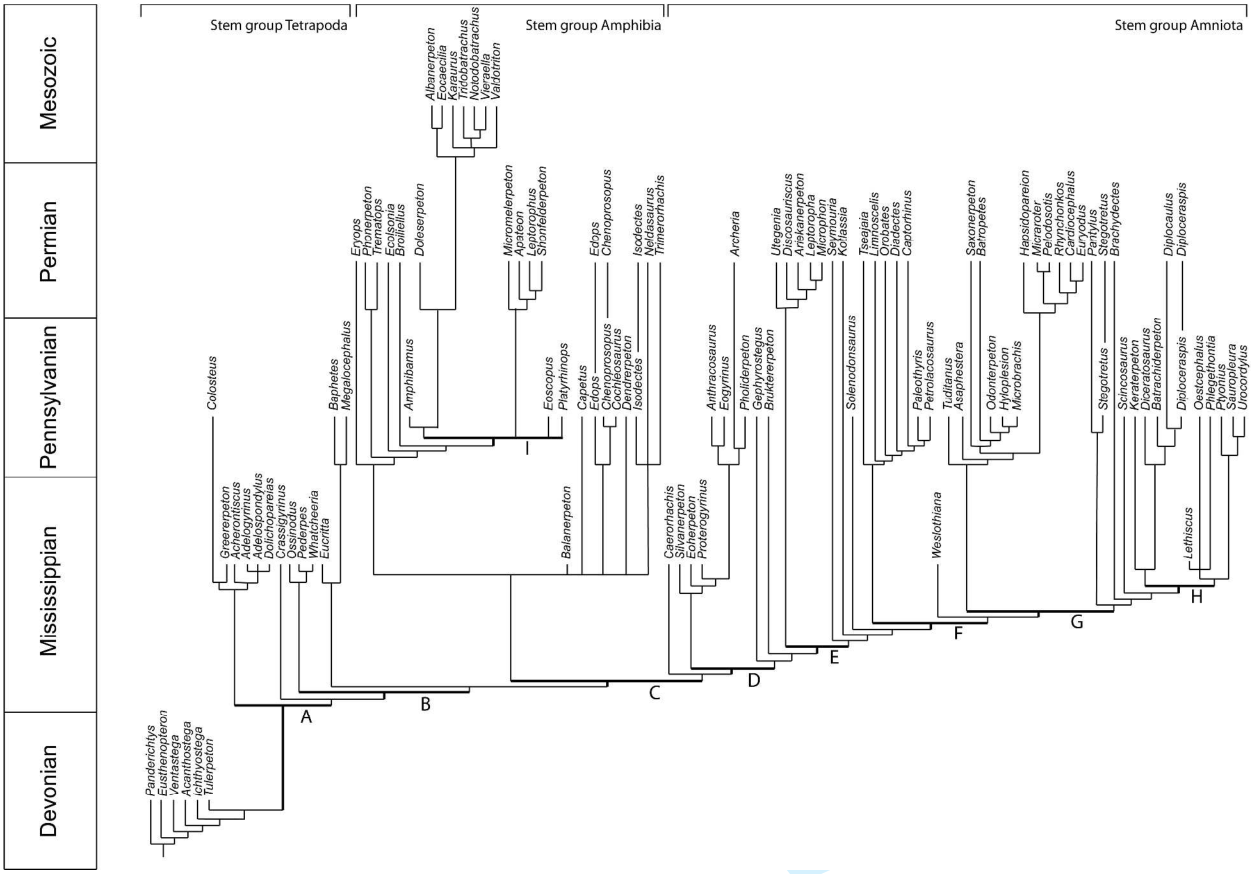

Mesozoic
Permian
Pennsylvanian
Mississippian
Devonian
Stem group Tetrapoda
Stem group Amphibia
Stem group Amniota
Panderichthys
Eusthenopteron
Ventastega
Acanthostega
Ichthyostega
Tulerpeton
Colosteus
Greererpeton
Acherontiscus
Adelogyrinus
Adelospondylus
Dolichopareias
Crassigyrinus
Ossinodus
Pederpes
Whatcheeria
Eucritta
Baphetes
Megalocephalus
Eryops
Phonerpeton
Trematops
Ecolsonia
Broiliellus
Doleserpeton
Amphibamus
Albanerpeton
Eocaecilia
Karaurus
Triadobatrachus
Notobatrachus
Vieraella
Valdotriton
Micromelerpeton
Apateon
Leptorophus
Schoenfelderpeton
Eoscopus
Platyrhinops
Balanerpeton
Capetus
Edops
Chenoprosopus
Cochleosaurus
Dendrerpeton
Isodectes
Neldasaurus
Trimerorhachis
Caerorhachis
Silvanerpeton
Eoherpeton
Proterogyrinus
Anthracosaurus
Eogyrinus
Archeria
Pholiderpeton
Gephyrostegus
Bruktererpeton
Utegenia
Discosauriscus
Ariekanerpeton
Leptoropha
Microphon
Seymouria
Kotlassia
Solenodonsaurus
Tseajaia
Limnoscelis
Orobates
Diadectes
Captorhinus
Paleothyris
Petrolacosaurus
Westlothiana
Tuditanus
Asaphestera
Saxonerpeton
Batropetes
Odonterpeton
Hyloplesion
Microbrachis
Hapsidopareion
Micraroter
Pelodosotis
Rhynchonkos
Cardiocephalus
Euryodus
Pantylus
Stegotretus
Brachydectes
Scincosaurus
Keraterpeton
Diceratosaurus
Batrachiderpeton
Diplocaulus
Diploceraspis
Lethiscus
Oestocephalus
Phlegethontia
Ptyonius
Sauropleura
Urocordylus
A
B
C
D
E
F
G
H
I

FIGURE 2

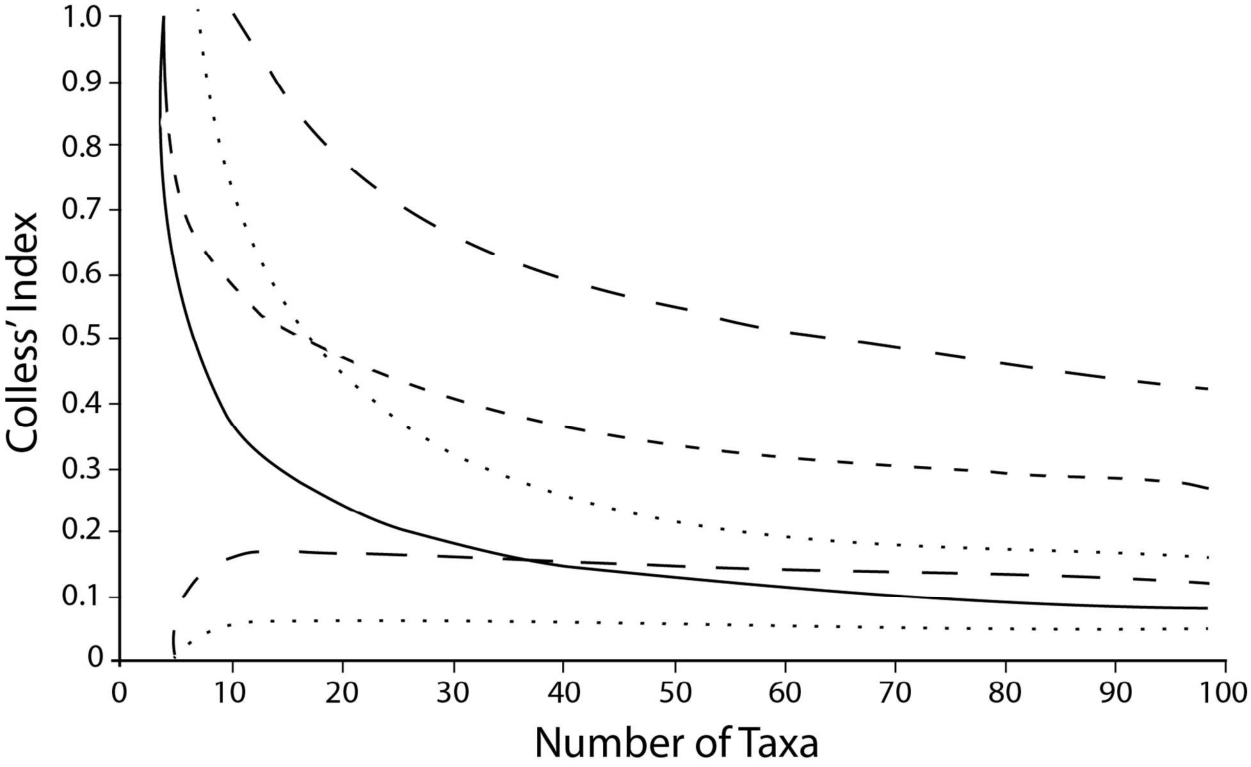

Colless' Index
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
10
20
30
40
50
60
70
80
90
100
Number of Taxa

FIGURE 3

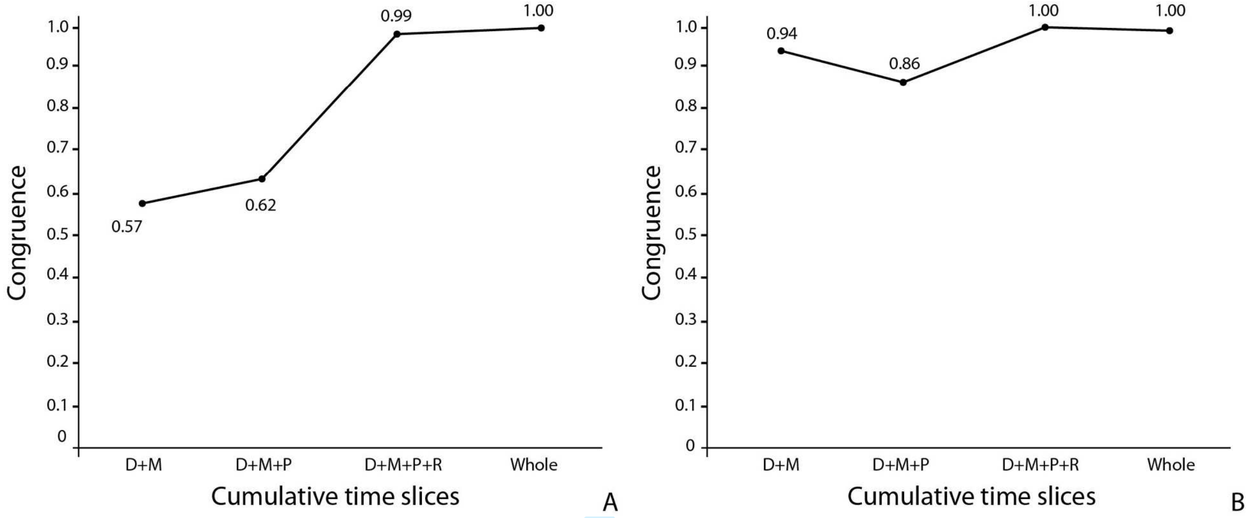

Congruence
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.57
0.62
0.99
1.00
D+M
D+M+P
D+M+P+R
Whole
Cumulative time slices
A
Congruence
0.94
0.86
1.00
1.00
D+M
D+M+P
D+M+P+R
Whole
Cumulative time slices
B

FIGURE 4

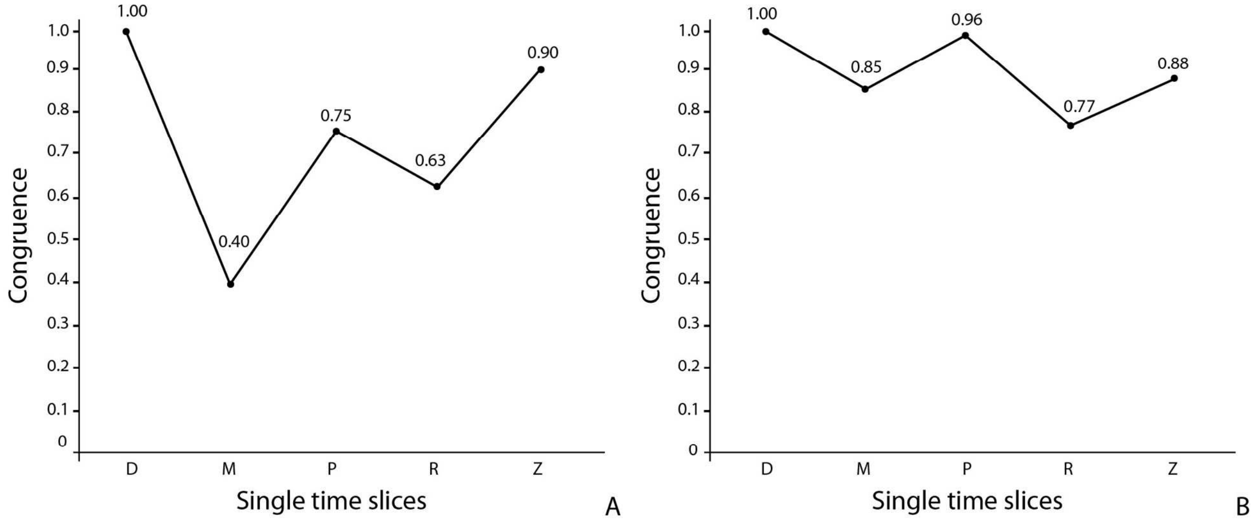

Congruence
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1.00
0.40
0.75
0.63
0.90
D
M
P
R
Z
Single time slices
A
Congruence
1.00
0.85
0.96
0.77
0.88
D
M
P
R
Z
Single time slices
B

FIGURE 5

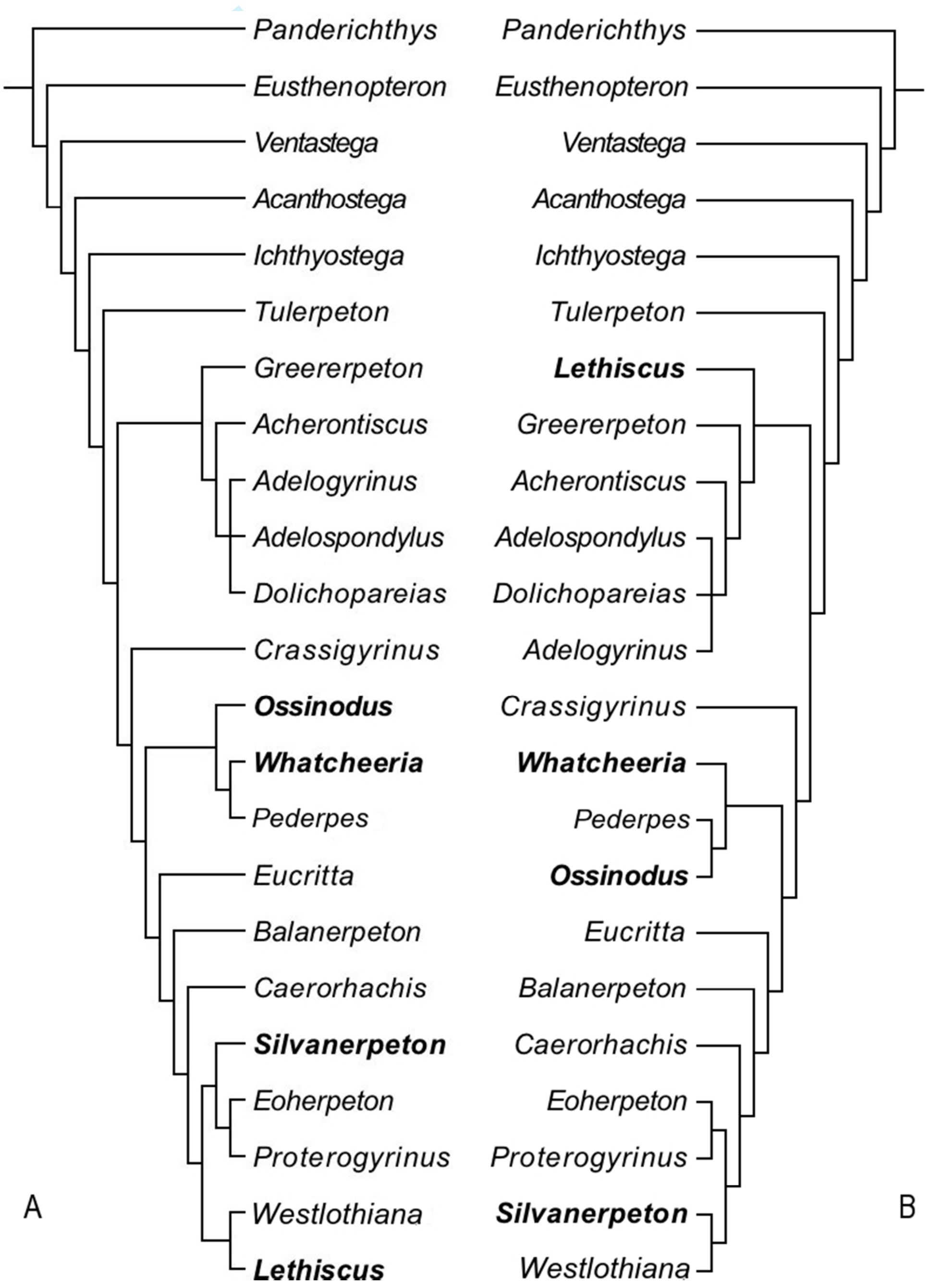

Panderichthys
Eusthenopteron
Ventastega
Acanthostega
Ichthyostega
Tulerpeton
Greererpeton
Acherontiscus
Adelogyrinus
Adelospondylus
Dolichopareias
Crassigyrinus
Ossinodus
Whatcheeria
Pederpes
Eucritta
Balanerpeton
Caerorhachis
Silvanerpeton
Eoherpeton
Proterogyrinus
Westlothiana
Lethiscus
A
Panderichthys
Eusthenopteron
Ventastega
Acanthostega
Ichthyostega
Tulerpeton
Lethiscus
Greererpeton
Acherontiscus
Adelospondylus
Dolichopareias
Adelogyrinus
Crassigyrinus
Whatcheeria
Pederpes
Ossinodus
Eucritta
Balanerpeton
Caerorhachis
Eoherpeton
Proterogyrinus
Silvanerpeton
Westlothiana
B

FIGURE 6

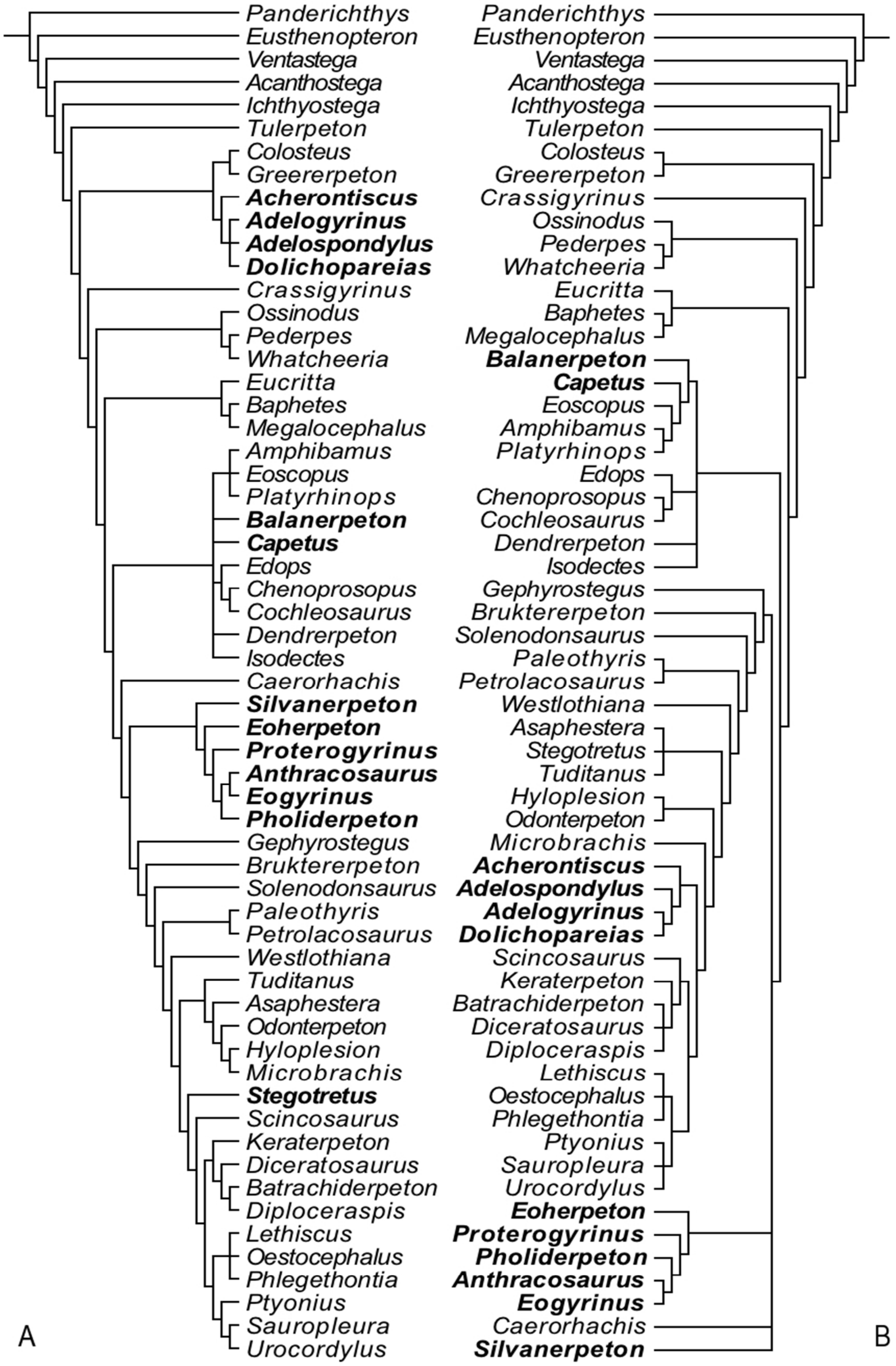

Panderichthys
Eusthenopteron
Ventastega
Acanthostega
Ichthyostega
Tulerpeton
Colosteus
Greererpeton
Acherontiscus
Adelogyrinus
Adelospondylus
Dolichopareias
Crassigyrinus
Ossinodus
Pederpes
Whatcheeria
Eucritta
Baphetes
Megalocephalus
Amphibamus
Eoscopus
Platyrhinops
Balanerpeton
Capetus
Edops
Chenoprosopus
Cochleosaurus
Dendrerpeton
Isodectes
Caerorhachis
Silvanerpeton
Eoherpeton
Proterogyrinus
Anthracosaurus
Eogyrinus
Pholiderpeton
Gephyrostegus
Bruktererpeton
Solenodonsaurus
Paleothyris
Petrolacosaurus
Westlothiana
Tuditanus
Asaphestera
Odonterpeton
Hyloplesion
Microbrachis
Stegotretus
Scincosaurus
Keraterpeton
Diceratosaurus
Batrachiderpeton
Diploceraspis
Lethiscus
Oestocephalus
Phlegethontia
Ptyonius
Sauropleura
Urocordylus
A
Panderichthys
Eusthenopteron
Ventastega
Acanthostega
Ichthyostega
Tulerpeton
Colosteus
Greererpeton
Crassigyrinus
Ossinodus
Pederpes
Whatcheeria
Eucritta
Baphetes
Megalocephalus
Balanerpeton
Capetus
Eoscopus
Amphibamus
Platyrhinops
Edops
Chenoprosopus
Cochleosaurus
Dendrerpeton
Isodectes
Gephyrostegus
Bruktererpeton
Solenodonsaurus
Paleothyris
Petrolacosaurus
Westlothiana
Asaphestera
Stegotretus
Tuditanus
Hyloplesion
Odonterpeton
Microbrachis
Acherontiscus
Adelospondylus
Adelogyrinus
Dolichopareias
Scincosaurus
Keraterpeton
Batrachiderpeton
Diceratosaurus
Diploceraspis
Lethiscus
Oestocephalus
Phlegethontia
Ptyonius
Sauropleura
Urocordylus
Eoherpeton
Proterogyrinus
Pholiderpeton
Anthracosaurus
Eogyrinus
Caerorhachis
Silvanerpeton
B

FIGURE 7

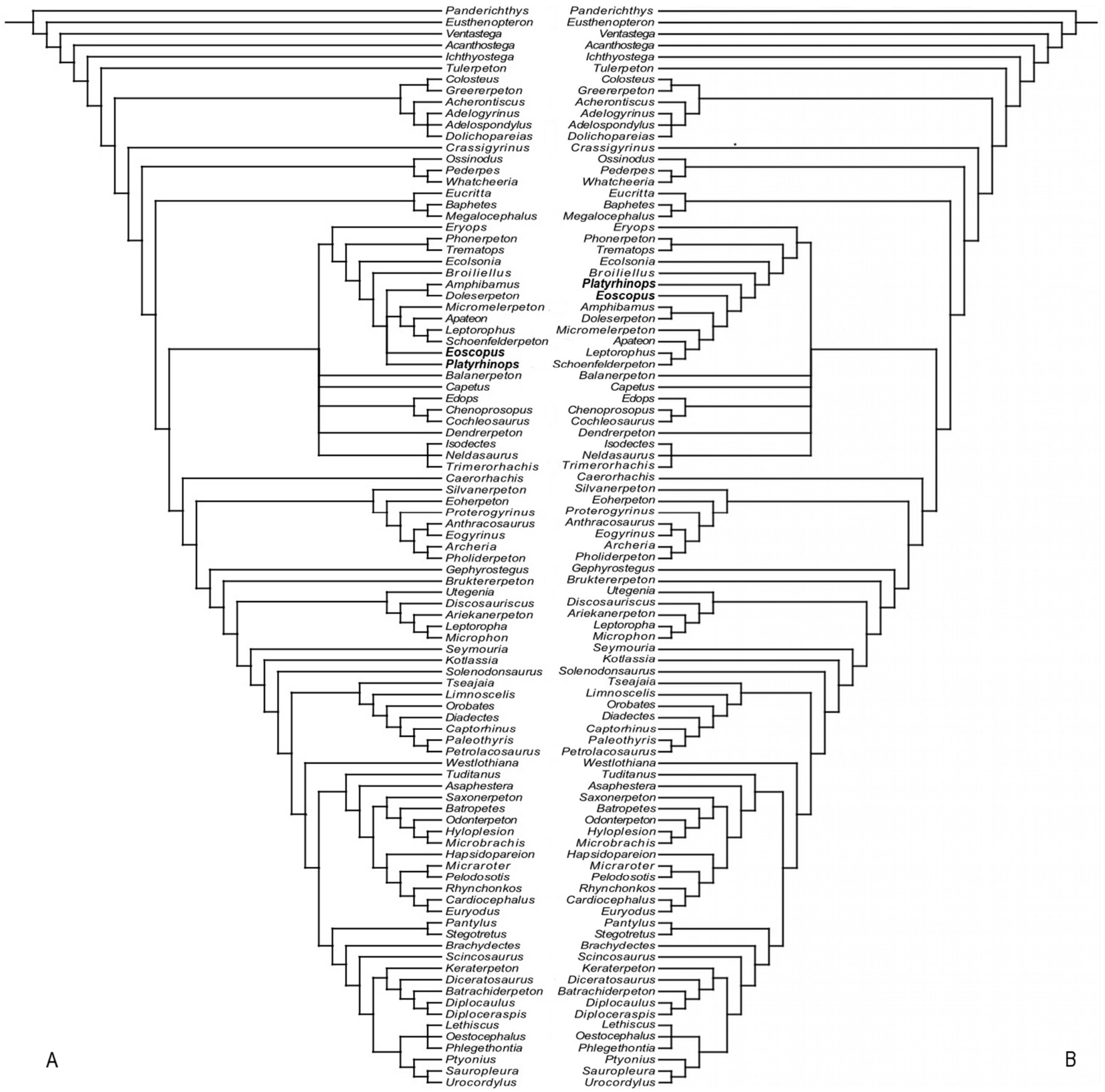

Panderichthys
Eusthenopteron
Ventastega
Acanthostega
Ichthyostega
Tulerpeton
Colosteus
Greererpeton
Acherontiscus
Adelogyrinus
Adelospondylus
Dolichopareias
Crassigyrinus
Ossinodus
Pederpes
Whatcheeria
Eucritta
Baphetes
Megalocephalus
Eryops
Phonerpeton
Trematops
Ecolsonia
Broiliellus
Amphibamus
Doleserpeton
Micromelerpeton
Apateon
Leptorophus
Schoenfelderpeton
Eoscopus
Platyrhinops
Balanerpeton
Capetus
Edops
Chenoprosopus
Cochleosaurus
Dendrerpeton
Isodectes
Neldasaurus
Trimerorhachis
Caerorhachis
Silvanerpeton
Eoherpeton
Proterogyrinus
Anthracosaurus
Eogyrinus
Archeria
Pholiderpeton
Gephyrostegus
Bruktererpeton
Utegenia
Discosauriscus
Ariekanerpeton
Leptoropha
Microphon
Seymouria
Kotlassia
Solenodonsaurus
Tseajaia
Limnoscelis
Orobates
Diadectes
Captorhinus
Paleothyris
Petrolacosaurus
Westlothiana
Tuditanus
Asaphestera
Saxonerpeton
Batropetes
Odonterpeton
Hyloplesion
Microbrachis
Hapsidopareion
Micraroter
Pelodosotis
Rhynchonkos
Cardiocephalus
Euryodus
Pantylus
Stegotretus
Brachydectes
Scincosaurus
Keraterpeton
Diceratosaurus
Batrachiderpeton
Diplocaulus
Diploceraspis
Lethiscus
Oestocephalus
Phlegethontia
Ptyonius
Sauropleura
Urocordylus
Panderichthys
Eusthenopteron
Ventastega
Acanthostega
Ichthyostega
Tulerpeton
Colosteus
Greererpeton
Acherontiscus
Adelogyrinus
Adelospondylus
Dolichopareias
Crassigyrinus
Ossinodus
Pederpes
Whatcheeria
Eucritta
Baphetes
Megalocephalus
Eryops
Phonerpeton
Trematops
Ecolsonia
Broiliellus
Platyrhinops
Eoscopus
Amphibamus
Doleserpeton
Micromelerpeton
Apateon
Leptorophus
Schoenfelderpeton
Balanerpeton
Capetus
Edops
Chenoprosopus
Cochleosaurus
Dendrerpeton
Isodectes
Neldasaurus
Trimerorhachis
Caerorhachis
Silvanerpeton
Eoherpeton
Proterogyrinus
Anthracosaurus
Eogyrinus
Archeria
Pholiderpeton
Gephyrostegus
Bruktererpeton
Utegenia
Discosauriscus
Ariekanerpeton
Leptoropha
Microphon
Seymouria
Kotlassia
Solenodonsaurus
Tseajaia
Limnoscelis
Orobates
Diadectes
Captorhinus
Paleothyris
Petrolacosaurus
Westlothiana
Tuditanus
Asaphestera
Saxonerpeton
Batropetes
Odonterpeton
Hyloplesion
Microbrachis
Hapsidopareion
Micraroter
Pelodosotis
Rhynchonkos
Cardiocephalus
Euryodus
Pantylus
Stegotretus
Brachydectes
Scincosaurus
Keraterpeton
Diceratosaurus
Batrachiderpeton
Diplocaulus
Diploceraspis
Lethiscus
Oestocephalus
Phlegethontia
Ptyonius
Sauropleura
Urocordylus
A
B

FIGURE 8

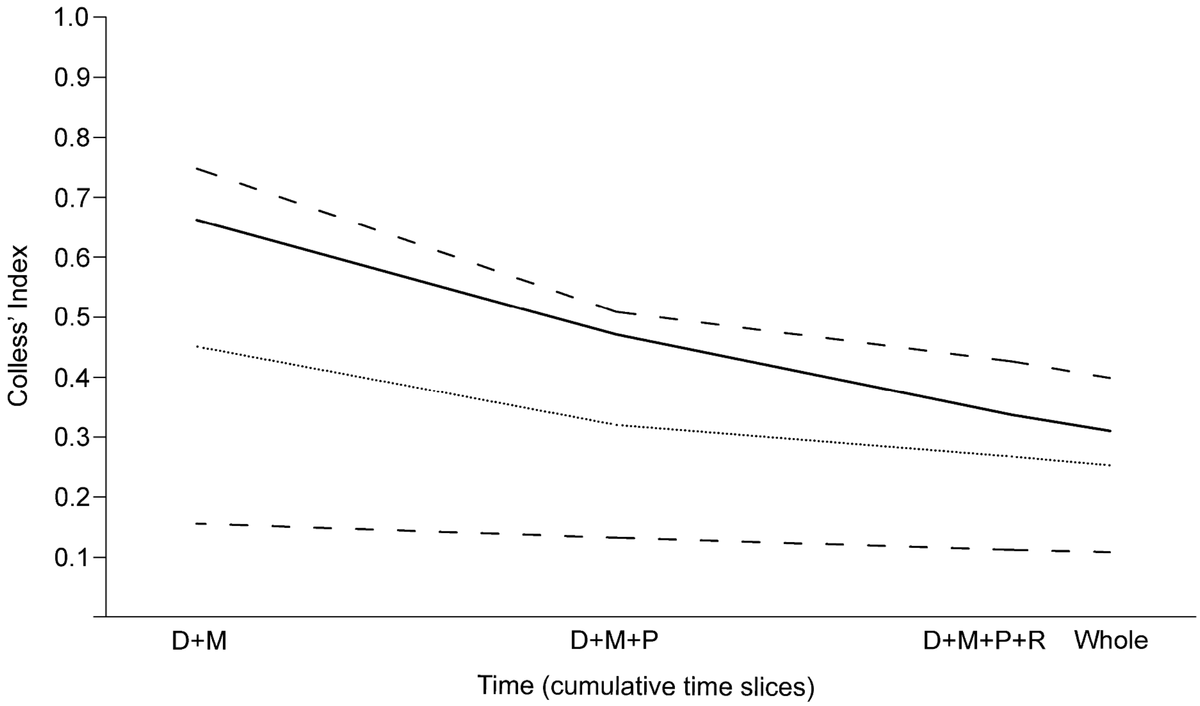

1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
Colless' Index
D+M
D+M+P
D+M+P+R
Whole
Time (cumulative time slices)

FIGURE 9

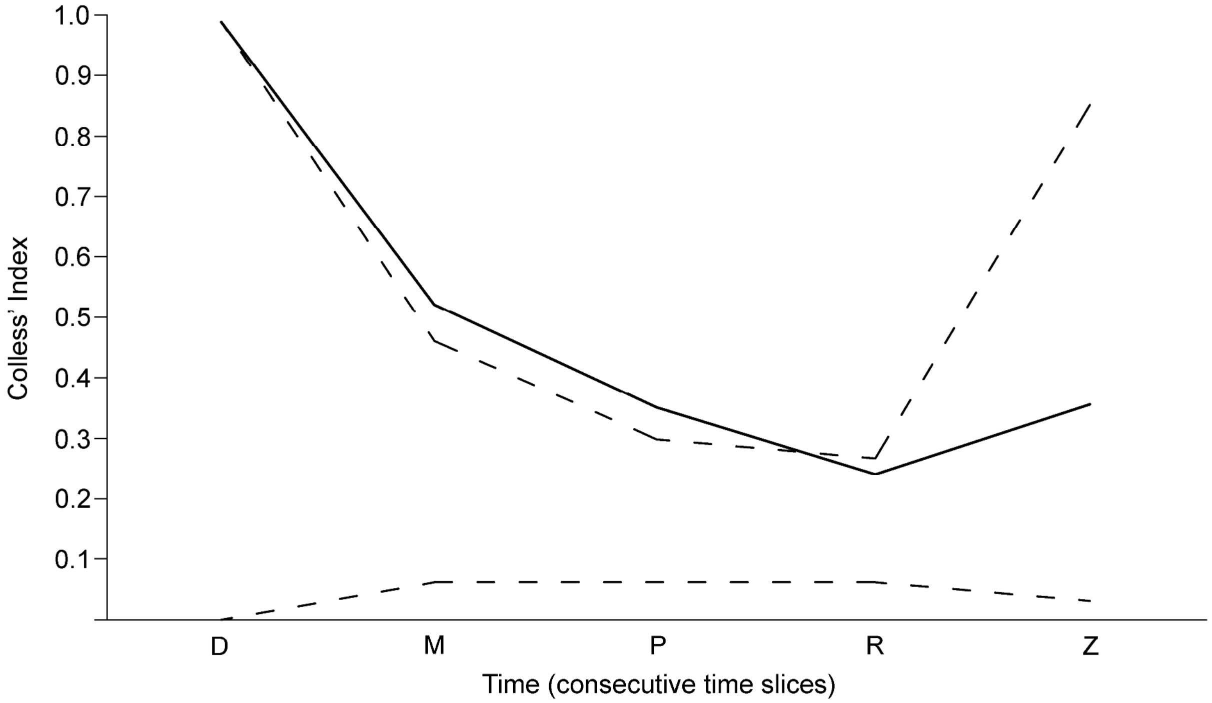

1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
Colless' Index
D
M
P
R
Z
Time (consecutive time slices)

FIGURE 10

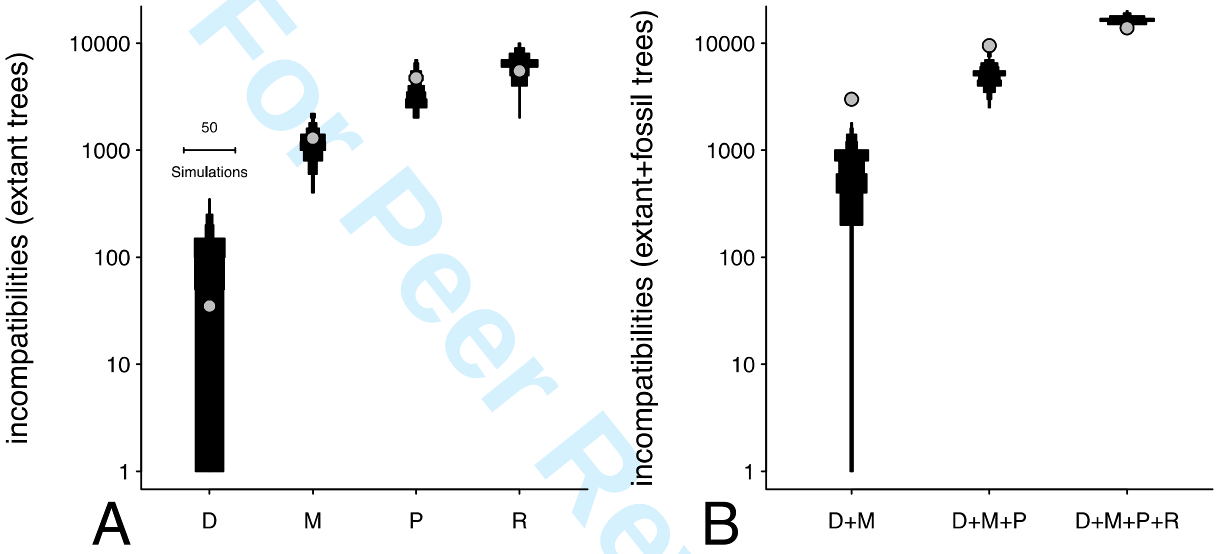
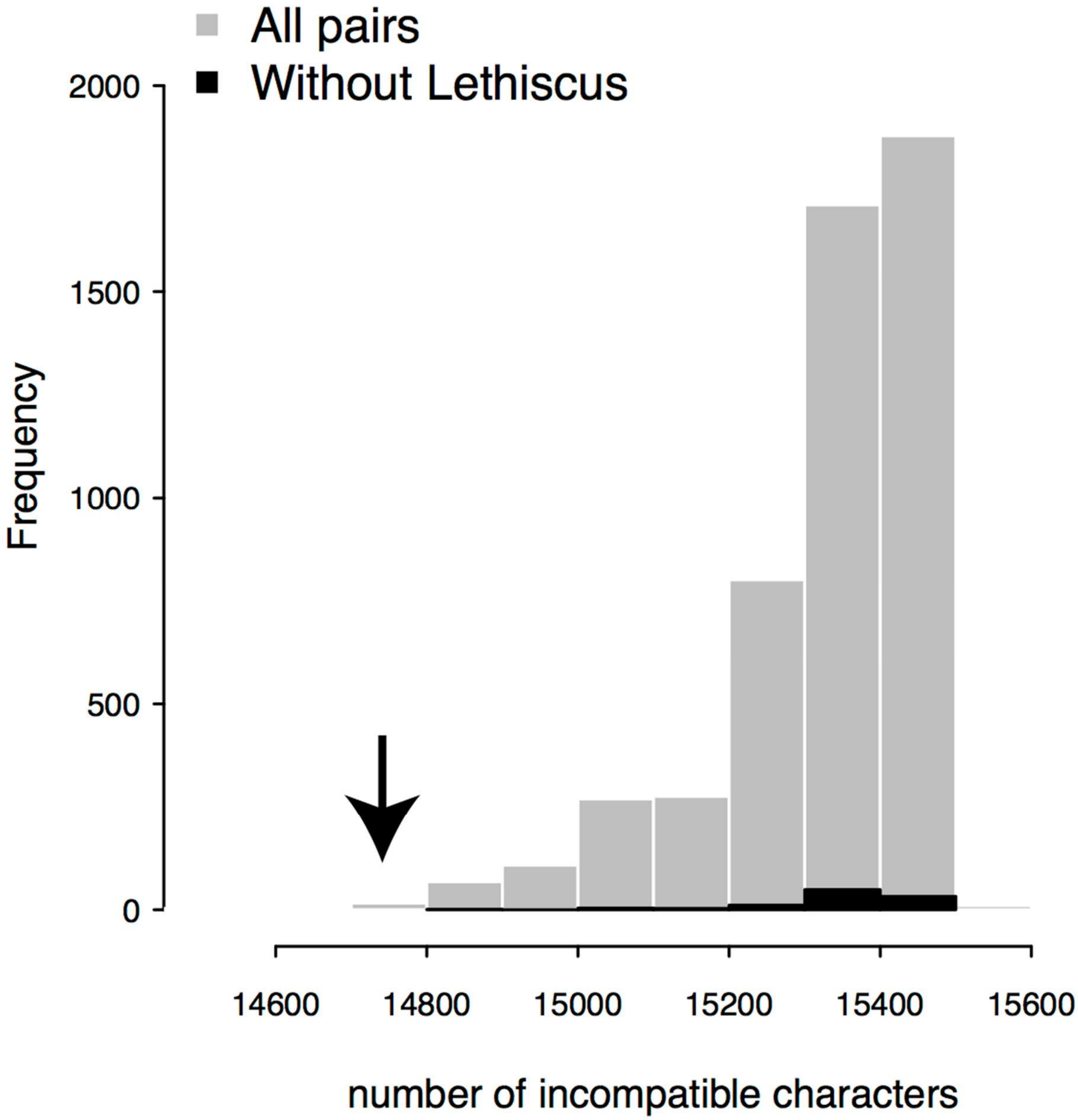

incompatibilities (extant trees)
10000
1000
100
10
1
50
Simulations
A
D
M
P
R
incompatibilities (extant+fossil trees)
10000
1000
100
10
1
B
D+M
D+M+P
D+M+P+R

FIGURE 11



All pairs
Without Lethiscus
2000
1500
1000
500
0
Frequency
14600
14800
15000
15200
15400
15600
number of incompatible characters