

Molecules, morphology, and the monophyly of diapsid reptiles

Michael S. Y. Lee

Department of Zoology, University of Queensland Brisbane, QLD 4072, Australia,

e-mail: mlee@zoology.uq.edu.au

Keywords: Reptilia, testudines, cladistics, osteology, fossils, molecular phylogenetics

Abstract

The morphological and molecular evidence for higher-level reptile relationships is reassessed. A combined analysis of 176 osteological, 40 soft anatomical, and 2903 (1783 aligned) molecular characters in 28 amniote taxa yields the traditional reptile tree. Synapsids (including mammals) are the sister taxon to all other amniotes, including all extant reptiles. Turtles group with anapsid parareptiles and fall outside a monophyletic Diapsida. Within diapsids, squamates and *Sphenodon* form a monophyletic Lepidosauria, and crocodiles plus birds form a monophyletic Archosauria. This tree is identical to the tree strongly supported by the osteological data alone when fossils are included. In a combined analysis the strong osteological signal linking turtles with anapsids is sufficient to override a soft anatomical signal placing turtles next to a heterodox archosaur-mammal clade, and a weaker molecular signal linking turtles with archosaurs. However, the turtle-archosaur clade cannot be statistically rejected. When fossils are ignored, the signal in the osteological data set disappears and, in a combined analysis of morphology and molecules, the molecular (turtle-archosaur) signal prevails. These results highlight the importance of fossils, not just in osteological studies, but even in "combined" analyses where they cannot be scored for the majority of characters (soft anatomy and molecules). Although the total number of molecular traits (2903) is much greater than the total number of morphological traits (216), when only characters informative at the relevant levels are considered, the two data sets are approximately equal in size. The partition homogeneity test yields unreliable results unless uninformative (invariant and autapomorphic) characters are excluded. Analyses of the mitochondrial data suggest that recent evidence from nuclear genes for a heterodox turtle-crocodile clade (excluding birds) might be an artefact of inadequate sampling of a diverse outgroup (mammals) and thus, problems with rooting the reptile tree.

Contents

Introduction	1
Morphological and molecular data sets	4
Data analyses	6
Phylogenetic results	7
Separate analyses	7
Homogeneity tests	8
Combined analyses	9
Molecules, morphology and the importance of fossils	11
Acknowledgements	16
References	16
Appendix	18

Introduction

Despite over a century of research, the relationships between the major groups of living amniotes (mammals, turtles, squamates, *Sphenodon* and archosaurs) remain contentious. There is now a consensus that all living reptiles form a clade, i.e. that turtles are more closely related to other living reptiles than to mammals. This arrangement has now been supported by compelling evidence from both morphology (Gauthier et al., 1988; Laurin and Reisz, 1995; Lee, 1997; de Braga and Rieppel, 1997) and molecules (Marshall, 1992; Eernisse and Kluge, 1993; Van de Peer et al., 1993; Caspers et al., 1996; Zardoya and Meyer, 1998). However, relationships within the reptile clade remain controversial, due mainly to disagreement over the affinities of turtles with the remaining reptiles, which are all conventionally termed "diapsids" on the basis of their possession of two temporal fenestrae (Williston, 1917; Romer, 1966; Carroll, 1988; Benton, 1996). There

are two major clades of living diapsids, archosaurs (crocodiles and birds) and lepidosaurs (*Sphenodon* and squamates). The monophyly of archosaurs, and of lepidosaurs, is extremely strongly corroborated (e.g. Gauthier, 1984; Benton, 1984, 1985; Evans, 1984, 1988; Gauthier et al., 1988; Kemp, 1988; Rieppel and de Braga, 1996; Dilkes, 1998; but see Hedges and Poling, 1999 and comments below). If these two diapsid clades are provisionally accepted, there are three ways in which turtles might be related to diapsids: they might fall outside a monophyletic Diapsida, or might render diapsids (as currently construed) paraphyletic by being the sister group to either lepidosaurs or archosaurs. These three possibilities are here called the “anapsid”, “lepidosaur” and “archosaur” hypotheses respectively (Fig. 1). Each has some recent character support from either morphology or molecules.

The anapsid hypothesis is the traditional arrangement. It implies that turtles retain primitively unfenestrated (anapsid) skulls, and that the presence of two temporal openings is one of the diagnostic derived characters uniting archosaurs and lepidosaurs as a monophyletic Diapsida (e.g. Williston, 1917; Gaffney, 1980; Reisz, 1981; Gauthier, 1984; Benton, 1985, 1996; Gaffney and Meylan, 1988; Carroll, 1988). Under the most recent interpretations of this hypothesis, the nearest relatives to turtles are primitive anapsid reptiles, in particular, procolophonoids (Laurin and Reisz, 1995) or pareiasaurs (Gregory, 1946; Ivachnenko, 1987; Lee, 1997). The hypothesis of diapsid monophyly has been supported largely by morphological, and in particular skeletal, traits. However, even the skeletal traits are relatively few, and some are also correlated with temporal fenestration (e.g. Reisz, 1981; Gauthier, 1984; Benton, 1985; Evans, 1988; Laurin and Reisz, 1995). Furthermore, soft anatomical traits support a radically different arrangement of amniotes (Gauthier et al., 1988; Gardiner, 1982, 1993; see below). Thus, the morphological evidence for diapsid monophyly is not strong. Nevertheless, this arrangement was widely accepted, and analyses of basal amniote relationships continued to assume diapsids were monophyletic, and thus represented them using only one or two presumably basal forms (e.g. Laurin and Reisz, 1995; Lee, 1995).

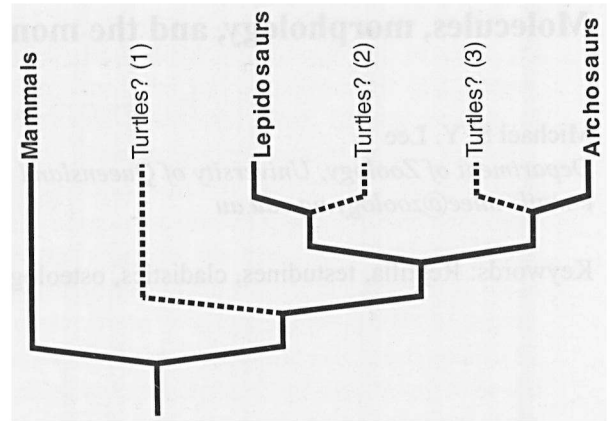


Fig. 1. The three proposed positions for turtles with respect to living diapsids. 1 is the “anapsid” hypothesis, 2 is the “lepidosaur” hypothesis, and 3 is the “archosaur” hypothesis.

The widespread assumption of diapsid monophyly was recently challenged by Rieppel and deBraga (1996; deBraga and Rieppel, 1997), who on the basis of a detailed analysis of osteological traits suggested the second arrangement – that turtles are nested within diapsids as the sister group to extant lepidosaurs. DeBraga and Rieppel’s conclusions were based on the first comprehensive phylogenetic analysis of basal amniotes (including turtles) where diapsids were not assumed to be monophyletic. The results were unexpected: when the Diapsida was split into all its component lineages, these lineages did not form a clade. Rather, turtles nested within the cluster of diapsid lineages, as the sister group to sauropterygians, a marine radiation consisting of placodonts, plesiosaurs, and their relatives. Lepidosaurs were the nearest extant relatives of the sauropterygian-turtle clade. This arrangement implies that turtles have secondarily closed their temporal fenestra. However, a reanalysis of deBraga and Rieppel’s data showed that, even if all their character codings were accepted as correct, support for their preferred phylogeny over the traditional phylogeny was not significant (Wilkinson et al., 1997). Further reanalysis, where some miscoded characters in the data set were rescored, yielded the traditional tree, with turtles grouping with anapsid reptiles and falling outside a monophyletic Diapsida. However, as relatively few characters were changed to cause this topological shift, the revised (traditional) tree was not strongly supported over the heterodox topology

(Lee, 1997). Thus, although deBraga and Rieppel's analysis did not demonstrate that turtles were modified diapsids, it certainly highlighted that the traditional anapsid hypothesis is far weaker than generally assumed, and that the lepidosaur hypothesis deserves serious consideration. A more recent study (Rieppel and Reisz, 1999) has again argued for a turtle-sauroptrygian clade, and discussed a scenario for an aquatic origin of turtles.

It was suggested that molecular sequences would provide independent data to resolve the issue (Wilkinson et al., 1997). However, the molecular data have surprisingly supported neither of the two arrangements proposed on the basis of morphology. Instead, early molecular results seemed most consistent with the third possibility – that turtles are most closely related to extant archosaurs (see Eernisse and Kluge, 1993 for review). A recent study (Zardoya and Meyer, 1998) included sequences from two complete mitochondrial genes (12S rRNA and 16S rRNA, totalling 2903 aligned base pairs) from the two major lineages of turtles (pleurodires and cryptodires), archosaurs (crocodiles and birds), and lepidosaurs (*Sphenodon* and squamates). The optimal trees under parsimony, neighbour-joining, and likelihood analyses grouped turtles as the sister group to archosaurs, but the traditional morphological arrangement of a monophyletic Diapsida was almost equally consistent with the data. The turtle-lepidosaur hypothesis (deBraga and Rieppel, 1997) was the least supported, but even this arrangement could not be statistically rejected (Zardoya and Meyer, 1998). Analysis of complete mitochondrial DNA genomes also grouped turtles and diapsids (represented only by squamates and archosaurs) to the exclusion of mammals (Kumazawa and Nishida, 1999). Another recent study (Hedges and Poling, 1999) proposed additional molecular support for an even more extreme version of the archosaur hypothesis, in which turtles are placed within archosaurs, as the sister group of crocodylians alone (to the exclusion of birds). While both nuclear and mitochondrial genes were evaluated in this study, the heterodox crocodile-turtle clade is supported by the signal found only in most (but not all) of the nuclear data; it was contradicted by sizeable portions of their nuclear data, and all of their mitochondrial data (see discussion). Thus, while

sizeable portions of the mitochondrial and nuclear molecular data support the archosaur hypothesis, the results are not yet conclusive.

Given the disagreement over reptile relationships and the observation that both the morphological and molecular evidence is ambiguous, it was decided to integrate these data into a combined phylogenetic analysis. This study thus revisits an important earlier study of amniote relationships (Eernisse and Kluge, 1993) that included an earlier version of the osteological data (Gauthier et al., 1988) and only the relatively short and poorly sampled gene sequences that were then available. That analysis also only included two "anapsid" groups (turtles and captorhinids), and thus could not test the monophyly of diapsids with respect to other anapsids, or the possible relationships of turtles with other anapsids (e.g. pareiasaurs and procolophonoids) within a combined morphological and molecular framework. Also, that analysis did not include euryapsids, which is a potentially important omission since they are the lineage that apparently "pulls" turtles into diapsids in one study (deBraga and Rieppel, 1997; Rieppel and Reisz, 1999). The present study includes all major anapsid lineages, and euryapsids, and thus can address these questions with a combined morphological and molecular analysis.

Combining data sets can allow secondary phylogenetic signals to emerge, which are not apparent when the data sets are analysed individually (Barrett et al., 1991; Olmstead and Sweere, 1994; Nixon and Carpenter, 1996). Thus, there is the possibility that morphological and molecular data sets, which each give equivocal results, when combined might interact to yield a strong and/or unexpected signal. Also, larger data sets (e.g. molecular sequences) might be expected to swamp smaller data sets (e.g. morphology), since if the strength of the signal (per character) is the same in both data sets, the larger data set would have a larger number of characters supporting its arrangement. However, empirical results suggest that this pattern often does not occur, with a coherent signal in a smaller morphological data set outweighing an ambiguous signal in the larger molecular set (e.g. see Nixon and Carpenter, 1996; Baker et al., 1998). Thus, there is no reason to expect that the emergent signal in

the combined analysis will be that of 2000 aligned nucleotides, rather than that of 200 morphological characters; it could even be some other signal entirely (e.g. turtles as sister group to *Sphenodon* alone). As the results show, the traditional anapsid hypothesis (and thus diapsid monophyly), despite repeated recent challenges, remains the most strongly supported by the combined data sets. The turtle-archosaur clade remains a possibility, but the turtle-lepidosaur clade can be statistically rejected.

Morphological and molecular data sets

The osteological data set of deBraga and Rieppel (1997) was revised and expanded, yielding a total of 176 characters for 28 ingroup taxa. New characters, and modifications to existing character definitions and codings, are listed in the appendix. Most of these modifications have been previously published by Lee (1997a) and subsequently accepted (Rieppel and Reisz, 1999), and are thus not contentious. However, there remains disagreement over some characters, and the reasons for the codings adopted in this study are presented in the appendix, along with references that discuss and/or illustrate these problematic characters. In order to reduce computing time, the two procolophonoids, and the three pareiasaurs, each coded separately in that analysis, are here combined into single terminal taxa (Procolophonoidea and Pareiasauridae respectively). This is not a problematic simplification since in the analyses of the original (deBraga and Rieppel, 1997), and recoded data sets, the procolophonoid lineages formed a clade, as did the pareiasaur lineages. For the osteological data set, the outgroup used to root the tree was coded with the condition found in diadectomorphs, the nearest relatives of amniotes (Gauthier et al., 1988). Ichthyosaurs, an important taxon omitted from the original analysis, were added to the data set by Motani et al. (1998): however, their codings were questionable as they were based on a single (albeit primitive) ichthyosaur, rather than on a range of basal ichthyosaurs. Furthermore, they are being recoded for these traits by other workers (M. Caldwell and P. Spencer, in prep). Hence, they are not included in the formal analysis here. However, exploratory analyses were

performed with ichthyosaurs included and coded according to Motani et al. (1998): their inclusion did not change tree topology – they always emerged as the sister group to euryapsids without changing existing relationships – and had only a minor effect on bootstrap values. Thus, their exclusion because of the problems above does not have a major effect on the results.

The soft anatomical data set was derived from the detailed review in Gauthier et al. (1988). To make the terminal taxa in that analysis comparable to those used in the osteological analysis, crocodiles and birds were combined into the taxon Archosauriformes, and lepidosaurs were split into *Sphenodon* and squamates. The data matrix includes all 40 characters identified in that study that were informative with respect to relationships between mammals, turtles, archosauriforms, *Sphenodon* and squamates (appendix). The outgroup used to root the tree was coded with the condition found in lissamphibians, the nearest extant relatives of amniotes. The soft anatomical characters discussed in Gardiner (1982) and Løvtrup (1985) are reanalysed in Gauthier et al. (1988), and as such have been included in the present study.

The molecular data consists of the complete 12S and 16S rRNA mitochondrial DNA sequences for mammals, turtles, *Sphenodon*, squamates, and archosaurs. Of the total of 2903 sites, 1120 ambiguously aligned sites were excluded, leaving a total of 1783 sites for analysis (Zardoya and Meyer, 1998). Analyses were performed to evaluate if there was any phylogenetic signal in the excluded sites: there was none (see below). Again, in order to make the terminal taxa used in the molecular analysis comparable with those in the other analyses, various species in the molecular matrix had to be grouped into higher taxa. The monotreme, marsupial, and two placental sequences were combined into a single “mammal” sequence (representing part of the osteological taxon Cynodontia), the pleurodire and cryptodire sequences into a single “turtle” sequence (part of the osteological taxon Testudines), and the bird and crocodile sequences into a single “Archosaur” sequence (part of the osteological taxon Archosauriformes). Construction of such “consensus” sequences was done as follows. When the two turtles (a pleurodire and cryptodire) both had the

same nucleotide at a given site (e.g. "A"), the "testudine" sequence was scored with that nucleotide at that site. However, at sites where they differed (e.g. the pleurodire with a "T" and the cryptodire with a "C"), the "testudine" sequence was scored as polymorphic ("TC"). This approach of reconstructing consensus sequences for higher taxa, commonly used in morphological analyses, is now being increasingly used in molecular studies as well (e.g. Zravy et al., 1998), and yields more accurate results than arbitrarily using a single species to represent a large clade in an analysis (Wiens, 1998; Bininda-Emonds et al., 1998). Analyses of the molecular data alone, with the terminal species combined into consensus sequences in this manner, yielded almost identical results (see below) to the previous analysis of Zardoya and Meyer (1998), where the terminal species were coded individually. This suggests that the consensus sequence approach faithfully retained the phylogenetic signal that was present in the original data set.

The present study combines what appear to be recent, comprehensive data sets for each category of traits. Of course, not all characters have been included. There are doubtless many other osteological, soft anatomical and molecular characters that have been discussed in the literature: no study can hope to include every trait in every category. In particular, the soft anatomical traits of Gardiner (1993) and the nuclear DNA data of Hedges and Poling (1999) were not used, for reasons discussed below. Also, the available complete mitochondrial genomes (Kumaza and Nishida, 1999) do not yet include *Sphenodon*, and thus also have not been used in this analysis.

Gardiner (1993) listed additional traits supporting the Haemathermia hypothesis (a bird-mammal clade); however, these newer data require rigorous reassessment. Only 25 of the 97 characters used in Gardiner's (1993) analysis are discussed; the remainder are only briefly tabulated. Many of the inadequately defined characters appear to have been miscoded but a proper assessment of the correct character states for various taxa can only be done after these states are properly defined. The majority of the 25 characters that are discussed have either been inadequately surveyed across amniotes (e.g. "true sleep"), lump together vaguely similar

structures of doubtful homology (e.g. hair and feathers), or are cladistically uninformative, the derived state being uniformly present in only one terminal taxon (e.g. subclavian artery origin, urinary bladder). Gardiner's study no doubt includes some novel and phylogenetically useful traits that have been overlooked by previous workers, but because the state definitions, and distributions, need to be more rigorously assessed, these characters are not included in this study.

The recently discussed nuclear sequences (Hedges and Poling, 1999) have also not been considered here, because the full details of that study could not be published due to space considerations, and the presumed alignments and phylogenetic analyses are currently being reevaluated by several workers. Though the crocodile-turtle clade proposed on the basis of the nuclear data appears to contradict the monophyly of one of the terminal taxa used in this analysis (Archosauriformes), the support for this heterodox clade appears weaker than suggested. The turtle-crocodile-bird trichotomy was interpreted by Hedges and Poling (1999) as resolved in favour of a crocodile-turtle clade. However, four of the nuclear analyses supported a crocodile-bird clade, while two other nuclear analyses supported a turtle-bird clade. These latter clades were also supported by some of their mitochondrial data, which were not included in the "combined" analyses. Thus, a more conservative interpretation of these data would be to treat the turtle-crocodile-bird trichotomy as unresolved, which is consistent with the widespread assumption of monophyly of Archosauriformes. For the purposes of this study, subdivision of Archosauriformes is not necessary because all three data sets considered here individually strongly support the monophyly of Archosauriformes: the osteology (e.g. Gauthier et al., 1988; de Braga and Rieppel, 1997); the soft anatomy (Gauthier et al., 1988), and the mitochondrial DNA (Zardoya and Meyer, 1998). Even if the Archosauriformes was split and coded as several taxa, for the data considered here these taxa would have emerged as a monophyletic group in all the individual and in the combined analyses. Thus, there is no necessity to split up Archosauriformes for this analysis. However, in order to rigorously evaluate the crocodile-turtle hypothesis in a future analysis involving nuclear

genes (which might not support monophyly of Archosauriformes), it will be necessary to subdivide the taxon into its component lineages: Crurotarsi (crocodiles and stem-group relatives), Ornithodira (birds and stem-group relatives), and all more basal archosauriform outgroups to these two clades, such as Proterochampsidae, Euparkeria, Erythrosuchidae, and Proterosuchidae (e.g. Benton and Clark, 1988; Sereno, 1991; Juul, 1994; Witmer, 1997; Gower and Weber, 1998).

Data analyses

The three data sets here were analysed according to maximum parsimony using PAUP* (Swofford, 1999), the only analytic method generally applicable to all three types of traits. (Other analyses applicable to the molecular data, i.e. maximum likelihood and neighbour joining, have already been performed by Zardoya and Meyer (1998), and yielded very similar results to the parsimony analysis). In all analyses, characters were equally weighted and all multistate characters were unordered. The permutation tail probability (PTP) test (Archie, 1989; Faith and Cranston, 1991) was employed to identify phylogenetic structure in the data, but the results were interpreted conservatively (Slowinski and Crother, 1998). Bootstrapping frequencies (Felsenstein, 1985) were based on 500 heuristic searches in PAUP: all cladistically uninformative characters were excluded during resampling since large numbers of these characters (e.g. as in many molecular data sets) have been demonstrated to lower bootstrap values (Carpenter, 1996). In order to standardise results, therefore, it seems best to delete all uninformative characters from bootstrap analysis. The alternative, of including such characters, is difficult as there are potentially an infinite number of "uninformative" characters for each class of data. Bremer support (Bremer, 1988) for each clade was calculated using PAUP commands generated in TreeRot (Sorenson, 1996), which were modified so that each search employed 100 rather than 10 random addition replicates. Initially, the three data sets were analysed separately. The osteological data was analysed two ways: all taxa included, and only extant taxa. The soft anatomi-

cal and molecular analyses each included only extant taxa (since fossils could not be scored for either class of trait).

The partition homogeneity (= incongruence length difference) test (Farris et al., 1994) as implemented in PAUP* was applied to see if any of the data sets were significantly incongruent. This test randomly assigns characters to partitions equal in size to the original partitions, and then compares the between-data set incongruence of the randomised partitions to that in the original partitions. These tests will be inaccurate if invariant characters are unequally distributed in the original data sets (e.g. if they are only present in molecular partitions but not in morphological partitions), as the randomisation procedure will assign them to both resampled data partitions. It has been recommended that such characters be deleted before employing the test (Cunningham, 1997). The same arguments apply to other uninformative (e.g. autapomorphic) characters, which should therefore be deleted as well (Lee, 2000). Accordingly, two sets of partition homogeneity tests were performed, one including all characters, and the other with uninformative (invariant and autapomorphic) characters deleted. Partition homogeneity tests might also be adversely affected if characters with large amounts of missing data are unequally distributed in the original data sets (e.g. DNA and soft anatomical characters can be scored only for living taxa, while osteological characters can potentially be scored for all taxa). In the randomised data sets, traits scorable in living forms only, and in all forms, will again be distributed evenly throughout the data partitions, which will therefore again not be representative of the original partitions. However, unlike the case with uninformative characters, there is no easy way to correct for this bias. Deleting fossil taxa will make traits in each partition scorable in all taxa, but might also significantly change the phylogenetic signal in the osteological data set (Gauthier et al., 1988; Donoghue et al., 1989; Lee, 1998; see results below).

It is arguable whether significant incongruence justifies keeping data sets separate during analyses (e.g. see Bull et al., 1993; Nixon and Carpenter, 1996; Ballard et al., 1998). Thus, regardless of the results of the partition homogeneity tests, combined

analyses were performed, incorporating all 176 osteological, 40 soft anatomical, and 1783 aligned molecular characters. Two such combined analyses were undertaken: one including all taxa (including fossil taxa, which could not be scored for soft anatomical or molecular traits), and one including only taxa with living representatives (i.e. terminals which could be scored for all traits). Comparison of the results of these two analyses would demonstrate whether fossils could have any effect on tree topology in a combined analysis where they could only be scored for 176 characters (at most) out of nearly 2000.

Phylogenetic results

Separate analyses

The analysis of the 176 osteological characters, including extant and fossil taxa, yielded a single most parsimonious tree (length = 771, consistency index = 0.54, retention index = 0.67) consistent with the traditional, anapsid hypothesis (Fig. 2). The PTP test found significant phylogenetic signal in this data set (P= 0.0057). The relationships discovered among the living forms were mammals (turtles(archosaurs+lepidosaurs)). Turtles fell outside diapsid reptiles and grouped strongly with anapsid “parareptiles”; in particular, with pareiasaurs. All diapsid reptiles, including fossil forms, formed a robust clade (Bremer and bootstrap support for all groups shown in Fig. 2).

When the osteological analysis was performed with only living taxa, three most parsimonious trees resulted (L=323, CI=0.84, RI=0.40): turtles either fell outside diapsids, or grouped with lepidosaurs, or grouped with archosaurs (i.e. the anapsid, lepidosaur, and archosaur hypotheses were all equally supported). The strict consensus of the three trees is shown in Figure 3. The PTP test found the data set contained a significant phylogenetic signal (P=0.007), presumably concentrated at the two resolved clades in the strict consensus (reptiles and lepidosaurs). Thus, it appears that the osteological data strongly supports diapsid monophyly (i.e. the anapsid hypothesis) only when fossil taxa are con-

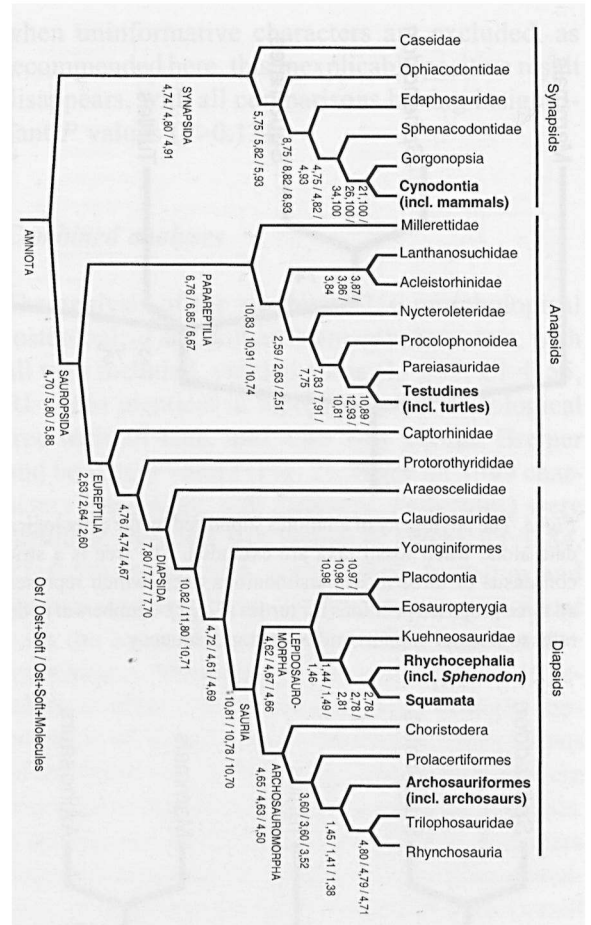


Fig. 2. The phylogeny of amniotes supported by the osteological data alone, by the morphological data (osteology + soft anatomy), and by the entire data set (osteology + soft anatomy + molecules). Fossils are included in all analyses. Living taxa in bold text, fossils in plain text. Turtles are related to anapsids, and diapsids are monophyletic. Bremer and bootstrapping support for each of these analyses in this order is listed at each node.

sidered. When fossils are ignored, the osteology is equally consistent with all three hypotheses.

The soft anatomical characters supported a very heterodox tree (L=63, CI=0.70, RI=0.61) which placed turtles as the sister group of an archosaur-mammal clade (Fig. 4). This arrangement is very similar to that proposed by Gardiner (1982, 1993) and Løvtrup (1985), who both considered many of these characters, but does not appear to have any other modern advocates. It can be concluded that the soft anatomical data by itself does not strongly support any of the three currently accepted positions for turtles. The PTP test again suggested



Fig. 3. The phylogeny of amniotes supported by the osteological data alone when fossil taxa are excluded. The tree is a strict consensus of three most parsimonious trees, which represent all three proposed positions for turtles (Fig. 1). Numbers at nodes refer to Bremer support and bootstrap frequency.

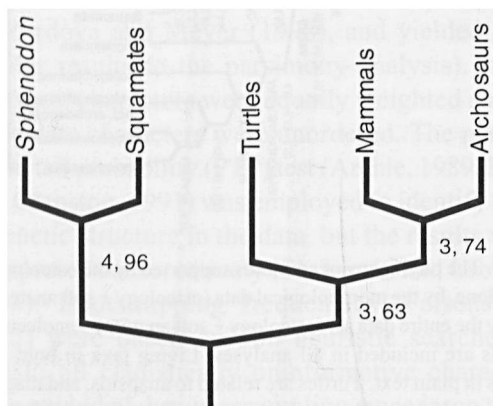


Fig. 4. The phylogeny of amniotes supported by the soft anatomy alone. Numbers at nodes refer to Bremer support and bootstrap frequency.

significant phylogenetic signal in this data set ($P=0.001$).

The analysis of the 1783 unambiguously-aligned nucleotide sites yielded a fully resolved tree ($L=2785$, $CI=0.93$, $RI=0.35$) supporting the Archosauria hypothesis: mammals(lepidosaurs(turtles+archosaurs)). The high CI is an artefact of the inclusion of large numbers of autapomorphic characters; the index falls to 0.80 when such characters are excluded. All clades received moderate to strong support (bootstrap around 70% or more, Bremer >7), including the turtle-archosaur clade (Fig. 5).

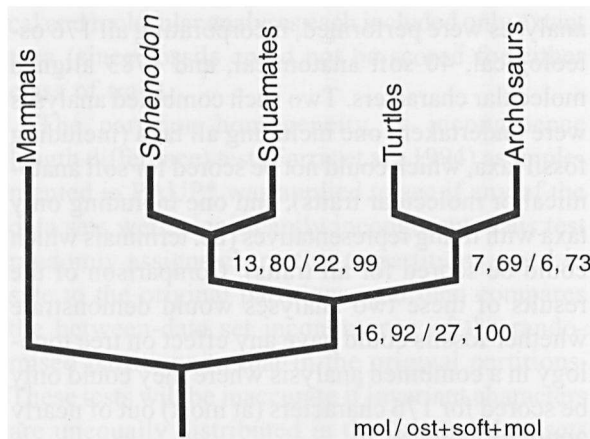


Fig. 5. The phylogeny of amniotes supported by the molecular data alone, and by the entire data set (osteology + soft anatomy + molecules) when fossils are ignored. Numbers at nodes refer to Bremer support and bootstrap frequency for the two respective analyses.

The PTP test indicated significant phylogenetic structure ($P=0.008$).

Analyses including all (well and poorly-aligned) sites yielded an identical tree, and almost identical branch supports, to those including only well-aligned sites. Analyses including only the poorly aligned sites gave a single tree but with extremely poor support on all clades (all had a bootstrap of less than 50% and a Bremer index of 1). Relaxation of parsimony by a single step produces a totally unresolved tree. The poorly-aligned sites also failed the PTP test ($P=0.21$). These results indicate that exclusion of the poorly aligned sites had almost no effect on the phylogeny, and that there is almost no signal in the poorly aligned sites. It is recommended that these tests should be routinely performed on excluded sites assumed to be ambiguously aligned; at the moment few studies do this, and the influence of excluded sites is not explicitly investigated.

Homogeneity tests

The partition homogeneity test was applied to each possible pair of data sets, and to a comparison between the entire morphological (osteological plus soft anatomical) and molecular data (Table 1). For

Table 1. The results of partition homogeneity (incongruence length difference) tests for various comparisons.

Partitions	All Characters	Informative Chars Only
Osteology vs Soft		
Anatomy	$P = 0.95$	$P = 0.95$
Osteology vs Molecules	$P = 0.77$	$P = 0.33$
Soft Anatomy vs		
Molecules	$P = 0.037^*$	$P = 0.13$
All Morphology vs		
Molecules	$P = 0.68$	$P = 0.21$

* Denotes significant at $P = 0.05$.

the soft anatomy versus molecules comparison, only living forms were included since fossil forms could not be scored for either data set. For all other comparisons, all taxa (living and fossil) were included, since the latter could be scored for one of the two data sets. Two analyses were performed, with all characters included, and with uninformative characters excluded. No significant heterogeneity was detected in seven of the eight comparisons: the exception was the comparison involving the soft anatomy and the molecules with all characters included. However, when only informative characters were included the soft anatomy and molecular data sets are not significantly incongruent.

The arguments that uninformative characters should be excluded before applying this test (see Lee 2000) are supported by these results (Table 1). Firstly, excluding uninformative characters substantially changed the resultant P values in all comparisons except for the soft anatomy vs osteology. This is because all comparisons except the latter contained a molecular partition which had many uninformative characters. The soft anatomy vs osteology comparison involved two morphological data sets, both with very few uninformative characters, so exclusion of the latter did not change the test results. Secondly, employing the test with all characters included produces rather counterintuitive results. In particular, the osteology is very congruent with both soft anatomical ($P=0.95$) and molecular ($P=0.77$) data, suggesting that the latter two data sets should also be quite congruent with one another. Yet, the test suggests that the molecular and soft anatomical data are highly incongruent with one another ($P=0.037$). However,

when uninformative characters are excluded, as recommended here, this inexplicable positive result disappears, with all comparisons having insignificant P values ($P>0.13$).

Combined analyses

The analysis of the combined 216 morphological (osteological and soft anatomical) characters, with all taxa included, yielded a tree ($L=837$, $CI=0.55$, $RI=0.66$) identical in topology to the osteological tree with all taxa, and with very similar Bremer and bootstrap values (Fig. 2). When all 1999 characters (osteology, soft anatomy, molecules) were included, the same tree results ($L=3630$, $CI=0.84$, $RI=0.59$), again with very similar bootstrap and Bremer supports (Fig. 2). This tree is consistent with the anapsid hypothesis: turtles(lepidosaurs+archosaurs). Monophyly of extant diapsids, in relation to other extant amniotes, was strongly supported in all these analyses. Moreover, the positions of the fossil taxa in these combined analyses were identical to that found in the osteological analysis. It appears that the fossils and osteological characters largely determined the results of the combined analyses, even though the latter comprised only a small proportion of the total data set (176 out of 1999 characters).

In order to test the lepidosaur and archosaur hypotheses against the osteological, morphological and entire data sets, which all yielded the same phylogeny, the “backbone constraints” function in PAUP was employed to find the best trees consistent with the following relationships between extant taxa:

- A. mammals((lepidosaurs)(turtles+archosaurs)), i.e. the “archosaur” hypothesis.
- B. mammals((lepidosaurs+turtles)(archosaurs)), i.e. the “lepidosaur” hypothesis.

In each of these analyses, all fossil taxa were allowed to “float”, i.e. insert anywhere in the backbone phylogeny of extant forms. The most parsimonious trees consistent with each backbone constraint were found using three data sets: the osteological data alone, the morphological data (osteology + soft anatomy), and for the entire data (osteology + soft anatomy + molecules).

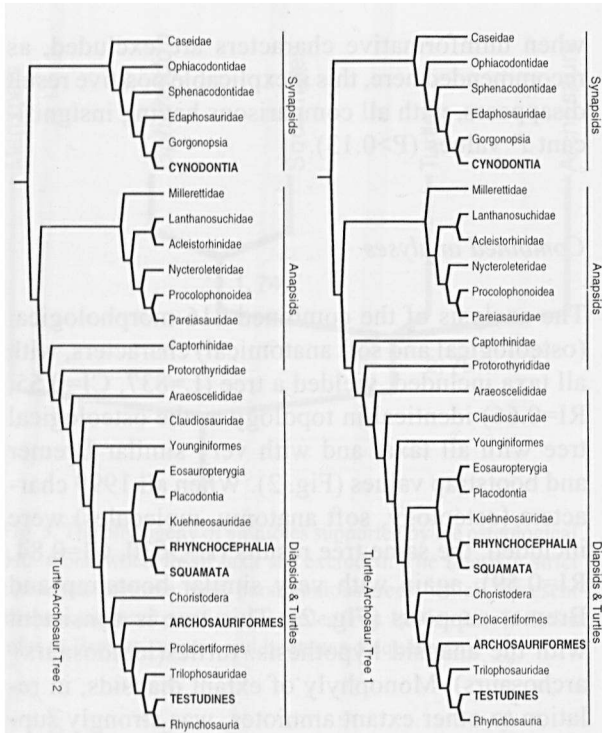


Fig. 6. The two most parsimonious trees consistent with the “backbone” constraint of a turtle-archosaur clade to the exclusion of other living reptiles. These same two trees are found in analyses employing the osteological, morphological (osteology + soft anatomy) and entire (osteology + soft anatomy + molecules) data. Living taxa in bold text, fossils in plain text.

For backbone constraint A (the archosaur hypothesis), the same two most parsimonious trees were found for all three data combinations (Fig. 6). Both trees grouped turtles with rhynchosaurs in the archosauromorph clade. Templeton tests (Templeton, 1983; Larson, 1994) indicate that these trees are

significantly worse than the most parsimonious (traditional) tree for the osteological, and for the combined morphological (osteology + soft anatomy) data (Table 2). However, for the entire (morphological + molecular) data, these trees are not significantly worse than the most parsimonious tree (Table 2).

For backbone constraint B (the lepidosaur hypothesis), the same two most parsimonious trees were found for all three data combinations (Fig. 7). These trees both grouped turtles with sauropterygians in a lepidosauromorph clade, and are very similar to the arrangement proposed by deBraga and Rieppel (1997). Templeton tests (Table 2) indicate that both trees are significantly worse than the most parsimonious (traditional) tree for the osteological, and the combined morphological data. For the entire data set, the difference verges on significant under the two-tailed test ($P=0.059$ and $P=0.064$) and is significant under the more liberal one-tailed test ($P=0.029$ and $P=0.032$). As the Templeton test is rather conservative and often appears to underestimate differences (Larson, 1994), and the one-tailed test is more applicable here as the alternative tree is known to be shorter a priori, these results thus constitute evidence against the lepidosaur hypothesis.

Thus, the osteological and combined morphological data sets both support the traditional, anapsid hypothesis and statistically reject the archosaur and lepidosaur hypotheses. The entire (morphological and molecular) data also supports the anapsid hypothesis over the other two hypotheses, and can statistically reject the lepidosaur, but not the archo-

Table 2. Support for the most parsimonious trees consistent with the turtle-archosaur and turtle-lepidosaur hypotheses compared to the most parsimonious (turtle-anapsid) tree. The trees are illustrated in Figs 2, 6 and 7. The same trees apply to the entire data set, the morphology, or just the osteology. The extra length, and two- and one-tailed Templeton test results, for each tree and each data set are shown.

MPT's consistent with alternative hypotheses	Entire data (Ost+Soft+Mol)	Morphology (Ost+Soft)	Osteology
turtle-archosaur 1 (Fig. 6)	+10, $P = 0.34, 0.17$	+18, $P = 0.029^{**}, 0.015^{**}$	+16, $P = 0.045^{**}, 0.023^{**}$
turtle-archosaur 2 (Fig. 6)	+10, $P = 0.36, 0.17$	+18, $P = 0.029^{**}, 0.015^{**}$	+16, $P = 0.045^{**}, 0.023^{**}$
turtle-lepidosaur 1 (Fig. 7)	+19, $P = 0.059^*, 0.029^{**}$	+20, $P = 0.015^{**}, 0.008^{**}$	+17, $P = 0.035^{**}, 0.018^{**}$
turtle-lepidosaur 2 (Fig. 7)	+19, $P = 0.064^*, 0.032^{**}$	+20, $P = 0.018^{**}, 0.009^{**}$	+17, $P = 0.041^{**}, 0.020^{**}$

* = significant at $P<0.1$, ** = significant at $P<0.05$.

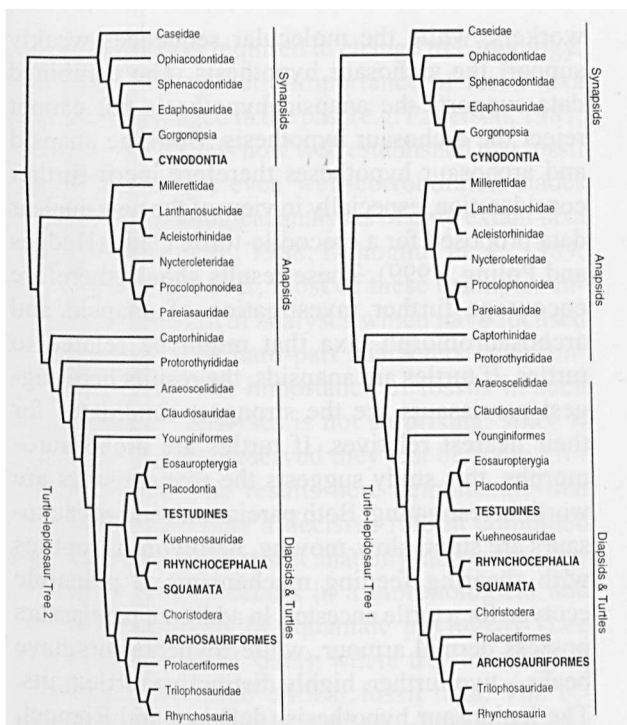


Fig. 7. The two most parsimonious trees consistent with the “backbone” constraint of a turtle-lepidosaur clade to the exclusion of other living reptiles. These same two trees are found in analyses employing the osteological, morphological (osteology + soft anatomy) and entire (osteology + soft anatomy + molecules) data. Living taxa in bold text, fossils in plain text.

saur, hypothesis. In terms of rank order, the anapsid hypothesis is the most supported in all three analyses, followed by the archosaur, with the lepidosaur hypothesis being the least supported (Table 2).

Interestingly, when fossil taxa are excluded, the osteological signal for the anapsid hypothesis disappears, and the osteological data are equally consistent with all three hypotheses (see above results for osteological data, and Fig. 3). As would be expected, therefore, when fossils were excluded, the combined analyses were not greatly influenced by the osteological data. The combined 216 morphological (osteological and soft anatomical) characters yielded a tree (L=389, CI=0.81, RI=0.46) consistent with the anapsid hypothesis but with diapsid monophyly very poorly supported (bootstrap=58%, Bremer=2) (Fig. 8). The entire morphological and molecular data set yielded a tree (L=3176, CI=0.91, RI=0.38) with the same topology as that implied by the molecular data alone, and with bootstrap frequencies hardly changed (Fig. 5). This tree is

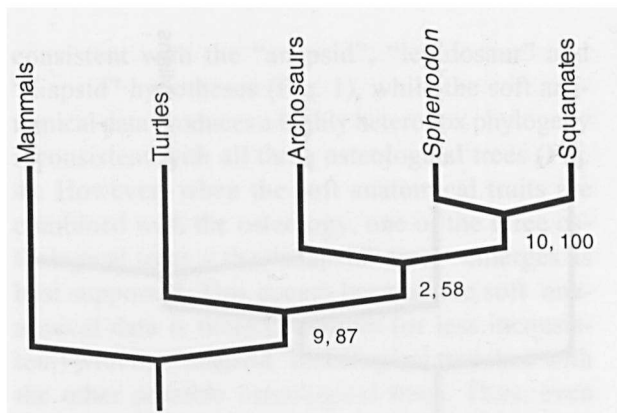


Fig. 8. The phylogeny of amniotes supported by the combined morphological (osteological and soft anatomical) data, when fossils are ignored. Numbers at nodes refer to Bremer support and bootstrap frequency.

consistent with the archosaur hypothesis.

Thus, when all relevant taxa are considered, the traditional anapsid hypothesis is supported in the morphological and combined (morphological + molecular) analyses, mainly due to a strong signal from the osteology that is sufficient to override the weaker signals from the soft anatomy and molecules. However, if fossils are ignored, the osteological signal disappears and the heterodox arrangement supported by the molecular sequences remains uncontradicted, even in the combined analyses that include osteology.

Molecules, morphology and the importance of fossils

The above results suggest that the anapsid hypothesis of turtle affinities – and thus diapsid monophyly – remains not only tenable, but is the most supported arrangement when the morphological, molecular and fossil evidence is considered simultaneously. The results are thus consistent with those of an earlier study of amniote relationships based on morphology and molecules (Eernisse and Kluge, 1993), which also found evidence that the traditional, osteological tree was preserved in a combined analysis. The broad structure of the current tree (Fig. 2) is highly congruent with traditional views, and most widely-recognised groupings are moderately to strongly corroborated. The “pelycosaur” and “therapsid” lineages cluster as successively

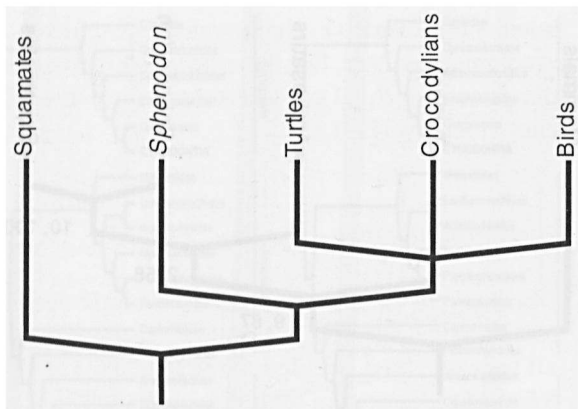


Fig. 9. The “conservative” tree based on the results published in Hedges and Poling (1999).

closer outgroups to mammals as part of a monophyletic Synapsida (e.g. Broom, 1932; Romer, 1966; Kemp, 1982). All other fossil reptiles group with living reptiles, comprising the Sauropsida. Sauropsids consist of two major lineages, parareptiles and eureptiles. The parareptile clade includes, among other things, millerettids, procolophonoids, pareiasaurs and turtles (e.g. Laurin and Reisz, 1995; Lee, 1995). Within parareptiles, the nearest relatives of turtles are pareiasaurs, slow-moving, armoured herbivores (e.g. Gregory, 1946; Lee, 1995). These results thus differ from those of Eernisse and Kluge (1993), which grouped turtles with captorhinids. Here, captorhinids, protorothyridids, and a monophyletic Diapsida group together, to form a clade of Eureptiles (Laurin and Reisz, 1995; Lee, 1997). Within diapsids, araeoscelids, younginiforms and certain other early forms are basal lineages outside the extant crown-clade (Laurin, 1991; deBraga and Rieppel, 1997; Motani et al., 1998). Finally, advanced, crown-clade diapsids (Sauria) consist of two diverse sister groups, lepidosauromorphs and archosauromorphs (Gauthier, 1984; Benton, 1985; Evans, 1988; Gauthier et al., 1988; deBraga and Rieppel, 1997). The marine sauropterygians – placodonts, plesiosaurs and relatives – are part of the Lepidosauromorpha (deBraga and Rieppel, 1997; Motani et al., 1998; Rieppel and Reisz, 1999).

When the data sets are analysed separately, the results are more equivocal. The osteology strongly supports the anapsid hypothesis when all taxa are considered, the soft anatomy gives a heterodox arrangement considered untenable by almost all

workers, while the molecular sequences weakly support the archosaur hypothesis. The combined data supports the anapsid hypothesis but cannot reject the archosaur hypothesis. Both the anapsid and archosaur hypotheses therefore merit further consideration, especially in view of the new nuclear data proposed for a crocodile-turtle clade (Hedges and Poling, 1999). These results should therefore encourage further investigation of anapsid and archosauromorph taxa that might be related to turtles. If turtles are anapsids, the results here suggest pareiasaurs are the strongest candidates for their nearest relatives. If turtles are archosauromorphs, this study suggests the rhynchosaurs are worth investigating. Both pareiasaurs and rhynchosaurs are stout, slow-moving, herbivorous reptiles with shearing feeding mechanisms, a plausible ecotype for a turtle ancestor. In addition, pareiasaurs possess dermal armour, while rhynchosaurs have beaks – two further, highly distinctive turtle traits. The lepidosaur hypothesis (deBraga and Rieppel, 1997) does not receive support from any individual data set (osteology, soft anatomy, mitochondrial DNA, or nuclear DNA), is statistically rejected by the combined analysis, and appears to be the least likely of the three. If turtles are forced to be lepidosauromorphs, they indeed group with the marine sauropterygians, as suggested by Rieppel and colleagues (deBraga and Rieppel, 1997; Rieppel and Reisz, 1999). However, this tree is significantly worse than the most parsimonious tree. There is thus no phylogenetic support for an aquatic origin of turtles. Also, it should be noted that all basal turtles described to date (e.g. *Proganochelys*, *Australochelys*, *Palaeochersis* and *Proterochersis*) appear to be terrestrial, being found in fully or partly terrestrial deposits and exhibiting tortoise-like features such as extremely short, robust digits and domed shells (Rougier et al. 1995).

Previous surveys have demonstrated how morphological traits can drive the results of combined morphological and molecular analyses, even when they are in the minority (Nixon and Wheeler, 1996; Baker et al., 1998). In this analysis, however, it is notable that the traditional tree is not retrieved when morphological traits are added to the molecular data set without increasing taxonomic sampling. Rather, the addition of morphological traits together

with fossil taxa is required to change tree topology. While the phylogenetic importance of fossil taxa has been disparaged in the past (e.g. Patterson, 1981; Gardiner, 1982), it is now well established that fossil taxa can overturn even well-corroborated clades found in morphological analyses of only extant taxa (e.g. Gauthier et al., 1988; Donoghue et al., 1989; Lee, 1998). However, most of these examples involve morphological analyses which have focused on osteological or hard-part characters determinable in fossils. The importance of fossils in such "osteocentric" analyses is not surprising, since if sufficiently well-preserved they can be scored for all characters. The results here demonstrate that fossils can also play a decisive role in combined analyses of hard parts, soft anatomy, and molecules. A similar pattern occurs in a morphological and molecular analysis of squamate phylogeny (Lee and Reeder, unpubl. data), where tree topology is determined by three critical fossil taxa. The effects of fossils were not explicitly addressed in the study of amniote relationships by Eernisse and Kluge (1993), but as their combined (morphological + molecular) tree was identical to their morphological tree, and their morphological tree was shown elsewhere to be structured by fossil taxa (Gauthier et al., 1988), the same effects can be inferred. Thus, these studies challenge the tempting assumption that fossils might be unimportant in combined molecular and morphological analyses, because they can be scored for very few characters (only a varying proportion of the "hard part" morphology). A recent analysis of relationships within turtles based on morphology and molecules (Shaffer et al., 1997) showed a much weaker effect of fossils consistent with this assumption: in that combined analysis, addition of fossils did not change topology but reduced support for certain clades, because they contained large amounts of missing data and/or inserted along branches leading to (previously) well-supported clades. The results here demonstrate that fossils are not always so indifferent, but can contribute strong and unique signals in phylogenetic analyses.

Another interesting result concerns the separate and combined analyses of the osteology and soft anatomy, when fossils are omitted. The osteological analysis produces three most-parsimonious trees,

consistent with the "anapsid", "lepidosaur" and "diapsid" hypotheses (Fig. 1), while the soft anatomical data produces a highly heterodox phylogeny inconsistent with all three osteological trees (Fig. 4). However, when the soft anatomical traits are combined with the osteology, one of the three osteological trees – the "anapsid" tree – emerges as best supported. This occurs because the soft anatomical data is more consistent (or less inconsistent) with the "anapsid" osteological tree than with the other possible osteological trees. Thus, even though at first glance the soft anatomical data might seem to contain little information since it implies a rather bizarre tree, it still serves to arbitrate between the three equally-parsimonious osteological hypotheses. Thus, the affinities of turtles are unresolved when the osteology is analysed separately, strangely resolved when the soft anatomy is analysed separately – but resolved (albeit weakly) in favour of the traditional view when the data sets are combined. This is a good example of a secondary phylogenetic signal emerging in a combined analysis, which is not apparent when the data sets are partitioned and analysed separately (Barrett et al., 1991; Nixon and Wheeler, 1996; Wiens and Reeder, 1997).

This study also underscores the observation that, while molecular data sets often contain many more characters than morphological ones, when only characters informative at the relevant levels are considered, the difference can be much less pronounced. This is because molecular studies sequence all sites in a certain region. This increased objectivity, however, means that most characters in aligned sequence data sets will not be cladistically informative. There will often be alignment ambiguities which will exclude certain segments from analysis. Of the aligned sites, many will either be invariant or cladistically uninformative (e.g. each derived condition characterising only a single terminal taxon). Finally, of the remaining sites which are cladistically informative in molecular analyses, some will not be informative at the relevant levels but serve to unite species which are already "known" to be closely related. For instance, the analysis of Zardoya and Meyer (1998) was aimed at testing relationships between mammals, turtles, archosaurs and lepidosaurs. However, their terminal taxa were

species, and thus not all the aligned, cladistically informative characters helped illuminate the relationships between these four groups, i.e. they are not informative at the relevant level. For some informative sites, the derived state united only the two turtles, or the two archosaurs, or the two lepidosaurs, or various mammal clades, and were thus effectively “autapomorphies” of widely-accepted terminal taxa (turtles, archosaurs, lepidosaurs, mammals). In other sites the derived state cuts across the four putatively monophyletic groups of interest: uniting, for instance, a turtle (e.g. the pleurodire) and a mammal (e.g. the monotreme); such characters were also effectively uninformative with respect to the question being investigated (relationships between the four groups), since they are equally consistent with any of the phylogenetic hypotheses being tested. These comments, it should be emphasised, only apply to parsimony analyses. Likelihood and neighbour-joining analyses make use of more or all of the sequence data. In likelihood analyses, for instance, invariant sites, and changes within putatively monophyletic groups, will contribute towards calculations of gamma parameters, invariant sites, and transition-transversion ratios, etc.

The same problems of course also affect morphological data: many morphological characters will be uninformative at the relevant levels, being of dubious homology (= “alignment”), or invariant in all taxa, or autapomorphies of a single terminal taxon of accepted monophyly, or convergently present only in derived members of accepted monophyletic terminal taxa. However such characters are usually deliberately culled from morphological data sets before analysis – whether or not this is advisable (e.g. Yeates, 1995). Thus, unlike in molecular analyses, almost all the characters in the morphological data sets are directly relevant to resolving relationships among the groups of interest. The methodology of morphological systematics is therefore more eclectic and subjective, but it means that the information density of such data sets is higher. In this analysis, although initially the data consisted of 216 morphological and 2903 molecular traits, when only traits cladistically informative at the relevant levels are considered, there are actually 204 morphological traits and 288

molecular traits. Hence, it is not surprising that the strong morphological signal regarding relationships between turtles, lepidosaurs and archosaurs managed to overwhelm the more ambiguous molecular signal addressing these relationships. Not only that, but the addition of the 2903 (1783 aligned) base pairs of molecular data to the morphological data set had only minor effects on the Bremer and bootstrap support for all clades. There is certainly a strong phylogenetic signal in both the morphological and molecular data sets, as shown by the PTP test results. However, while a large portion of the signal in the morphological data contributed to the resolution of relationships between turtles, archosaurs, and lepidosaurs, much of the molecular signal consisted of support for parts of the tree (e.g. monophyly of mammals, therians, lepidosaurs, archosaurs and turtles) not directly relevant to the problem under investigation (turtle-diapsid relationships).

Finally, although an attempt to integrate the nuclear data (Hedges and Poling, 1999) into these analyses is currently premature since these data are currently being reevaluated and supplemented by several groups, it might be worthwhile speculating how the traditional arrangement supported in this study might be reconciled with the heterodox arrangement proposed in that study. As stated in the introduction, a conservative interpretation of their data suggests a turtle-bird-crocodylian trichotomy (Fig. 9). If this interpretation is adopted, there is a single unrooted network for reptiles that is consistent with both the traditional arrangement, and the heterodox phylogeny proposed by Hedges and Poling (Fig. 10, 11). If the tree is rooted at turtles, the traditional topology retrieved in the current study results. If, however, this tree is rooted at squamates, a tree consistent with Hedges and Poling’s (conservative) tree results, with the squamates and *Sphenodon* being basal to a turtle-crocodile-bird clade.

There remains the problem of deciding which rooting position is more likely to be correct. A rooting problem is not as likely to affect the mitochondrial or morphological data sets considered here since the outgroups to reptiles in both these data sets were well-sampled: basal mammals (marsupials and monotremes) as well as eutherians were included in the mitochondrial and soft anatomical

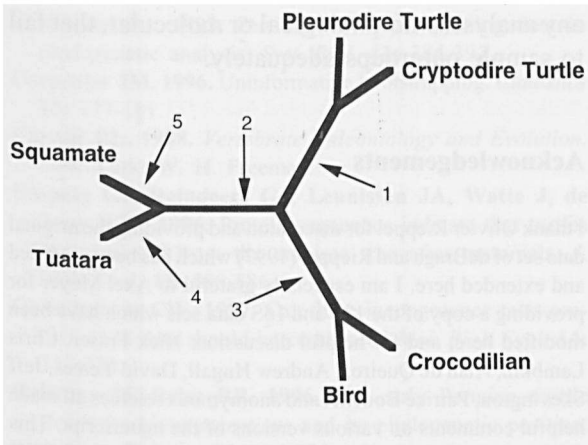


Fig. 10. The unrooted network consistent with all three hypotheses of turtle affinities, and the recent arrangement proposed by Hedges and Poling (1999). Rooting at position 1 yields the traditional “anapsid” hypothesis, position 2 gives the “archosaur” hypothesis, position 3 gives the “lepidosaur” hypothesis, and position 5 yields a tree very similar to that of Hedges and Poling (see Fig. 9).

data sets, along with the next extant outgroup (amphibians). The osteological data set used all mammals as well as fossil synapsids (pelycosaur and therapsids), and again employed the next nearest scorable outgroup (diadectomorphs). Thus, the rooting near turtles implied by the combined signal from these data seems reasonably robust, given the exhaustive characterisation of the outgroups. However, a rooting problem is more likely to affect the nuclear data, since only two derived eutherian mammals (rodent and primate) were used and more basal eutherians (e.g. edentates), marsupials and monotremes were not sampled. This observation is particularly interesting since analyses of the 12S and 16S sequences using different samples of mammals to root a “reptile” tree gave highly variable results. Under parsimony, for instance, four different rooting positions were found, three of these corresponding to the “anapsid”, “archosaur”, and “lepidosaur” hy-

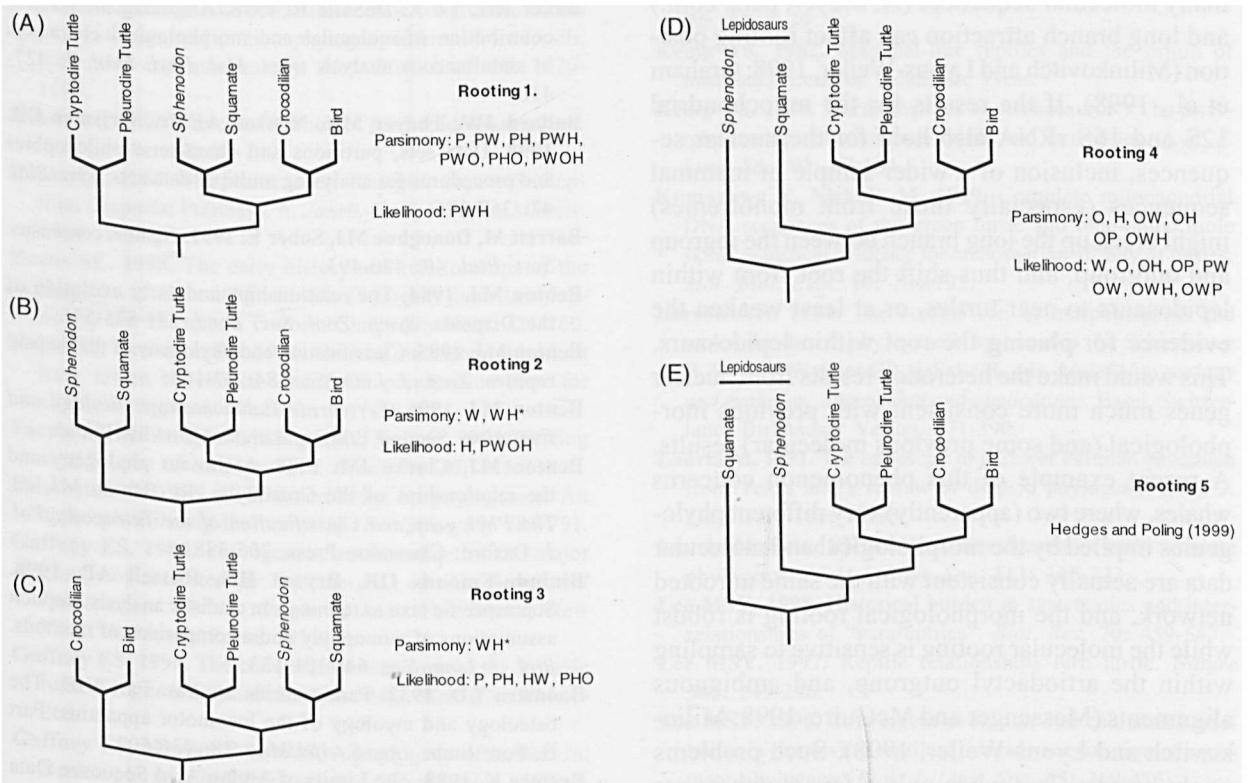


Fig. 11. Five different rooted trees consistent with the unrooted network in Fig. 10, and the combinations of mammalian outgroups that yield each particular rooting. The results of a likelihood and parsimony analyses are listed. Note how outgroup sampling can change the rooting position and thus, the preferred tree. P – Platypus, O – Opossum, W – Whale, H – Human. OP means the tree is rooted with only the opossum and platypus, W means the tree is rooted only with the whale, etc. Note that two equally-parsimonious trees were found when the tree was rooted with the human and the whale (*).

potheses, and the other being very similar to the tree of Hedges and Poling, with lepidosaurs (*Sphenodon* and squamates) being paraphyletic with respect to a turtle-archosaur clade (Fig. 10, 11E). Inclusion of all mammal outgroups, and most subsets of the outgroups that included the monotreme (platypus), resulted in a rooting near turtles (position 1) consistent with the “anapsid” hypothesis and diapsid monophyly (Fig. 11A). However, if the monotreme sequence was excluded, the tree tended to be rooted within lepidosaurs, resulting in a paraphyletic Lepidosauria and a turtle-archosaur clade (Fig. 11B,D). This is very similar to the tree proposed in Hedges and Poling (1999). It is perhaps significant that Hedges and Poling’s study used only two eutherians (rodent and primate) to root their tree and did not consider any monotreme or marsupial sequences. This might predispose it to root within lepidosaurs, especially since lepidosaurs appear to be at the end of long branches for many molecular sequences (A. Meyer, pers. com.) and long branch attraction can affect rooting position (Milinkovitch and Lyons-Weiler, 1998; Graham et al., 1998). If the results for the mitochondrial 12S and 16S rRNA also hold for the nuclear sequences, inclusion of a wider sample of mammal sequences (especially those from monotremes) might break up the long branch between the ingroup and outgroup, and thus shift the root from within lepidosaurs to near turtles, or at least weaken the evidence for placing the root within lepidosaurs. This would make the heterodox results from nuclear genes much more consistent with previous morphological (and some previous molecular) results. A recent example of this phenomenon concerns whales, where two (apparently) very different phylogenies implied by the morphological and molecular data are actually consistent with the same unrooted network, and the morphological rooting is robust while the molecular rooting is sensitive to sampling within the artiodactyl outgroup, and ambiguous alignments (Messenger and McGuire, 1998; Milinkovitch and Lyons-Weiler, 1998). Such problems might be prevalent when only a few members are used to represent a very diverse outgroup (mammals, artiodactyls). This is not to imply that rooting problems only affect molecular analyses, and do not affect morphological ones. Rather, they will affect

any analyses, morphological or molecular, that fail to sample outgroups adequately.

Acknowledgements

I thank Olivier Rieppel for discussion and providing the original data set of deBraga and Rieppel (1997) which has been modified and extended here. I am especially grateful to Axel Meyer for providing a copy of the 12S and 16S data sets which have been modified here, and for helpful discussion. Nick Fraser, Chris Lambkin, Alan de Queiroz, Andrew Hugall, David Yeates, Jeff Skevington, Patrice Bouchet and anonymous referees all made helpful comments on various versions of the manuscript. This work is supported by an Australian Research Council QEII Research Fellowship and Large Research grant.

References

- Archie JW. 1989. A randomization test for phylogenetic information in systematic data. *Syst. Zool.* **38**: 219-252.
- Baker RH, Yu X, DeSalle R. 1998. Assessing the relative contribution of molecular and morphological characters in simultaneous analysis trees. *Mol. Phyl. Evol.* **9**: 427-436.
- Ballard JW, Thayer MK, Newton AF Jr., Grismer ER. 1998. Data sets, partitions and characters: philosophies and procedures for analysing multiple data sets. *Syst. Biol.* **47**: 367-396.
- Barrett M, Donoghue MJ, Sober E. 1991. Against consensus. *Syst. Biol.* **40**: 486-493.
- Benton MJ. 1984. The relationships and early evolution of the Diapsida. *Symp. Zool. Soc. Lond.* **52**: 575-596.
- Benton MJ. 1985. Classification and phylogeny of the diapsid reptiles. *Zool. J. Linn. Soc.* **84**: 97-164.
- Benton MJ. 1996. *Vertebrate Palaeontology: Biology and Evolution*. Second Edition. London: Unwin Hyman.
- Benton MJ, Clarke JM. 1988. Archosaur phylogeny and the relationships of the Crocodylia. In: Benton MJ, ed. *The Phylogeny and Classification of the Tetrapods. Vol. 1*. Oxford: Clarendon Press, 295-338.
- Bininda-Emonds OR, Bryant HN, Russell AP. 1998. Supraspecific taxa as terminals in cladistic analysis: implicit assumptions of monophyly and a comparison of methods. *Biol. J. Linn. Soc.* **64**: 101-133.
- Boonstra LD. 1932. Pareiasaurian studies. Part VIII. The osteology and myology of the locomotor apparatus. Part B. Fore limb. *Ann. S. Afr. Mus.* **28**: 436-503.
- Bremer K. 1988. The Limits of Amino Acid Sequence Data in Angiosperm Phylogenetic Reconstruction. *Evolution* **42**: 795-803.
- Broom R. 1932. *The Mammal-like Reptiles of South Africa and the Origin of Reptiles*. London: F. and G. Witherby.
- Bull JJ, Huelsenbeck JP, Cunningham CW, Swofford DL,

- Waddell PJ. 1993. Partitioning and combining data in phylogenetic analysis. *Syst. Biol.* **42**: 384-397.
- Carpenter JM. 1996. Uninformative bootstrapping. *Cladistics* **12**: 177-181.
- Carroll RL. 1988. *Vertebrate Paleontology and Evolution*. New York: W. H. Freeman.
- Caspers GJ, Reindeers GJ, Leunissen JA, Watte J, de Jong WW. 1996. Protein sequences indicate that turtles branched off from the amniote tree after mammals. *J. Mol. Evol.* **42**: 580-586.
- Cunningham CW. 1997. Can three incongruence tests predict when data should be combined? *Mol. Biol. Evol.* **14**: 733-740.
- DeBraga M, Reisz RR. 1996. The early Permian reptile *Acleisthorhinus pteroticus* and its phylogenetic position. *J. Vert. Paleont.* **16**: 384-395.
- DeBraga M, Rieppel O. 1997. Reptile phylogeny and the affinities of turtles. *Zool. J. Linn. Soc.* **120**: 281-354.
- Dilkes DW. 1998. The early Triassic rhychosaur *Mesosuchus browni* and the interrelationships of basal archosauriform reptiles. *Phil. Trans. R. Soc. Lond. B* **353**: 501-541.
- Donoghue MJ, Doyle JA, Gauthier J, Kluge AG, Rowe T. 1989. The importance of fossils in phylogeny reconstruction. *Ann. Rev. Ecol. Syst.* **20**: 431-460.
- Ernise DJ, Kluge AG. 1993. Taxonomic congruence versus total evidence, and amniote phylogeny inferred from fossils, molecules and morphology. *Mol. Biol. Evol.* **10**: 1170-1195.
- Evans SE. 1984. The classification of the Lepidosauria. *Zool. J. Linn. Soc.* **82**: 87-100.
- Evans SE. 1987. The braincase of *Youngina capensis* (Reptilia: Diapsida; Permian). *N. Jahrb. Geol. Paläont. Monats.* **1984**: 193-203.
- Evans SE. 1988. The early history and relationships of the Diapsida. In: Benton MJ, ed. *The Phylogeny and Classification of the Tetrapods*. Oxford: Clarendon Press, 221-260.
- Faith DP, Cranston PS. 1991. Could a cladogram this short have arisen by chance alone? On permutation tests for cladistic structure. *Cladistics* **7**: 1-28.
- Farris JS, Källersjö M, Kluge AG, Bult C. 1994. Testing significance of incongruence. *Cladistics* **10**: 315-319.
- Felsenstein J. 1985. Confidence limits on phylogenies: An approach utilizing the bootstrap. *Evolution* **39**: 783-791.
- Gaffney ES. 1980. Phylogenetic relationships of the major groups of amniotes. In: Panchen AL, ed. *The Terrestrial Environment and the Origin of Land Vertebrates*. New York: Academic Press, 593-610.
- Gaffney ES. 1990. The comparative osteology of the Triassic turtle *Proganochelys*. *Bull. Amer. Mus. Nat. Hist.* **194**: 1-263.
- Gaffney ES, Meylan PA. 1988. A phylogeny of turtles. In: Benton MJ, ed. *The Phylogeny and Classification of Tetrapods*. Oxford: Clarendon Press, 157-219.
- Gardiner BG. 1982. Tetrapod classification. *Zool. J. Linn. Soc.* **74**: 207-232.
- Gardiner BG. 1993. Haemothermia: warm-blooded amniotes. *Cladistics* **9**: 369-395.
- Gauthier, J. 1984. A cladistic analysis of the higher systematic categories of the Diapsida. Ph.D dissertation, University of California, Berkeley.
- Gauthier J, Kluge AG, Rowe T. 1988. Amniote phylogeny and the importance of fossils. *Cladistics* **4**: 105 – 209.
- Gow CE. 1972. The osteology and relationships of the Millerettidae (Reptilia: Cotylosauria). *J. Zool.* **167**: 219-264.
- Gow CE. 1975. The morphology and relationships of *Youngina capensis* Broom and *Prolacerta broomi* Parrington. *Palaeont. Afr.* **18**: 89-131.
- Gower DJ, Weber E. 1998. The braincase of *Euparkeria*, and the evolutionary relationships of birds and crocodylians. *Biol. Rev.* **73**: 367-412.
- Graham SW, Kohn JR, Morton BR, Eckenwalder JE, Barrett SCH. 1999. Phylogenetic congruence and discordance among one morphological and three molecular data sets from Pontederiaceae. *Syst. Biol.* **47**: 545-567.
- Gregory WK. 1946. Pareiasaurs versus placodonts as near ancestors to the turtles. *Bull. Amer. Mus. Nat. Hist.* **86**: 276-323.
- Hedges SB, Poling LL. 1999. A molecular phylogeny of reptiles. *Science* **283**: 998-1001.
- Ivachnenko MF. 1987. [Permian parareptiles of the USSR]. *Trudy Pal. Inst. Acad. Nauk SSSR* **223**: 3-160. In Russian.
- Juul L. 1994. The phylogeny of basal archosaurs. *Palaeont. Afr.* **31**: 1-39.
- Kemp TS. 1982. Mammal-like reptiles and the origin of mammals. London: Academic Press.
- Kemp TS. 1988. Haemothermia or Archosauria?: The interrelationships of mammals, birds and crocodiles. *Zool. J. Linn. Soc.* **92**: 67-104.
- Kumazawa Y, Nishida M. 1999. Complete mitochondrial DNA sequences of the green turtle and blue-tailed mole skink: statistical evidence for archosaurian affinity of turtles. *Mol. Biol. Evol.* **16**: 784-792.
- Larson AR. 1994. The comparison of morphological and molecular data in phylogenetic systematics. In: Schierwater B, Steit B, Wagner GP, deSalle R, eds. *Molecular ecology and evolution: approaches and applications*. Basel, Switzerland: Birkhäuser Verlag, 371-390.
- Laurin M. 1991. The osteology of a Lower Permian eosuchian from Texas and a review of diapsid phylogeny. *Zool. J. Linn. Soc.* **101**: 59-95.
- Laurin M, Reisz RR. 1995. A reevaluation of early amniote phylogeny. *Zool. J. Linn. Soc.* **113**: 165-223.
- Lee MSY. 1995. Historical burden in systematics and interrelationships of "parareptiles". *Biol. Rev.* **70**: 459-547.
- Lee MSY. 1997. Reptile relationships turn turtle. *Nature* **389**: 245-246.
- Lee MSY. 1998. Convergent evolution and character correlation in burrowing reptiles: towards a resolution of squamate phylogeny. *Biol. J. Linn. Soc.* **65**: 369-453.
- Lee MSY. 2000. Uninformative characters and apparent conflict between morphology and molecules. *Mol. Biol. Evol.* in press.
- Løvtrup S. 1985. On the classification of the taxon Tetrapoda. *Syst. Zool.* **34**: 463-470.
- Marshall CR. 1992. Substitution bias, weighted parsimony,

- and amniote phylogeny as inferred from 18S rRNA sequences. *Mol. Biol. Evol.* **9**: 370-373.
- Messenger SL, McGuire JA. 1998. Morphology, molecules and the phylogenetics of cetaceans. *Syst. Biol.* **47**: 90-124.
- Milinkovitch MC, Lyons-Weiler J. 1998. Finding optimal ingroup topologies and convexities when the choice of outgroups is not obvious. *Mol. Phyl. Evol.* **9**: 348-357.
- Motani R, Minoura N, Ando T. 1998. Ichthyosaurian relationships illuminated by new primitive skeletons from Japan. *Nature* **393**: 255-257.
- Nixon KC, Carpenter JM. 1996. On simultaneous analysis. *Cladistics* **12**: 221-241.
- Olmstead RG, Sweere JA. 1999. Combining data in phylogenetic systematics: an empirical approach using three molecular data sets in the Solanaceae. *Syst. Biol.* **43**: 467-481.
- Patterson C. 1981. Significance of fossils in determining evolutionary relationships. *Ann. Rev. Ecol. Syst.* **12**: 195-223.
- Reisz RR. 1981. A diapsid reptile from the Pennsylvanian of Kansas. *Spec. Publ. Mus. Nat. Hist., Univ. Kansas* **7**: 1-74.
- Rieppel O, deBraga M. 1996. Turtles as diapsid reptiles. *Nature* **384**: 453-455.
- Rieppel O, Reisz RR. 1999. The origin and early evolution of turtles. *Ann. Rev. Ecol. Syst.* **20**: 1022.
- Romer AS. 1966. *Vertebrate Paleontology*. Third Edition. Chicago: University of Chicago Press.
- Romer AS, Price LI. 1940. Review of the Pelycosauria. *Geol. Soc. Amer. Spec. Pap.* **28**: 1-538.
- Rougier GW, de la Fuente MS, Arcucci AB. 1995. Late Triassic turtles from South America. *Science* **268**: 855-858.
- Sereno PC. 1991. Basal archosaurs: phylogenetic relationships and functional implications. *Soc. Vert. Pal. Memoir* **2**: 1-53.
- Shaffer HB, Meylan P, McKnight ML. 1997. Tests of turtle phylogeny: molecular, morphological, and paleontological approaches. *Syst. Biol.* **46**: 235-268.
- Slowinski JB, Crother BI. 1998. Is the PTP test useful? *Cladistics* **14**: 297-302.
- Sorenson MR. 1996. *TreeRot*. Computer Program for Macintosh. Distributed by the author, University of Michigan, Ann Arbor.
- Swofford DL. 1999. *PAUP**. *Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Version 4. Sunderland MA: Sinauer Associates.
- Templeton AR. 1983. Phylogenetic inference from restriction site endonuclease cleavage site maps with particular reference to the humans and apes. *Evolution* **37**: 221-244.
- Van de Peer Y, Neefs JM, de Rijk P, De Wachter R. 1993. Reconstructing evolution from eukaryotic small-ribosomal-subunit RNA sequences: calibration of the molecular clock. *J. Mol. Evol.* **37**: 221-232.
- Wiens JJ. 1998. The accuracy of methods for coding and sampling higher-level taxa for phylogenetic analysis: a simulation study. *Syst. Biol.* **47**: 397-413.
- Wiens JJ, Reeder TW. 1997. Phylogeny of the spiny lizards (*Sceloporus*) based on molecular and morphological evidence. *Herp. Monog.* **11**: 1-101.
- Wilkinson M, Thorley J, Benton MJ. 1997. Uncertain turtle relationships. *Nature* **387**: 466.
- Williston SW. 1917. The phylogeny and classification of reptiles. *J. Geol.* **25**: 411-421.
- Witmer LM. 1997. The evolution of the antorbital cavity of archosaurs: a study in soft-tissue reconstruction in the fossil record with an analysis of the function of pneumaticity. *Mem. Soc. Vert. Paleont.* **3**: 1-73.
- Yeates D. 1995. Why remove autapomorphies? *Cladistics* **8**: 387-389.
- Zardoya R, Meyer A. 1998. Complete mitochondrial genome indicates diapsid affinities of turtles. *Proc. Natl. Acad. Sci.* **95**: 14226-14231.
- Zrzavý J, Mihulka S, Kepka P, Bezdek A, Tietz D. 1998. Phylogeny of the Metazoa based on morphological and 18S ribosomal DNA evidence. *Cladistics* **14**: 249-285.

First draft received: 14 November 1999

Appendix

Osteological Characters (1-176)

Characters 1-168 are as described in deBraga and Rieppel (1997) with modifications as discussed by Lee (1997) and Motani et al. (1998) incorporated. These changes have been mostly accepted and thus not contentious (Rieppel and Reisz, 1999, p.3). However, changes regarding eight characters (characters 65, 82, 103, 120, 121, 127, 140 & 152) are questioned by Rieppel and Reisz (1999) and are discussed below. Some additional changes are made in the present analysis and also discussed below. Reasons for the codings adopted in this analysis discussed, and the references cited illustrate the contentious characters and/or describe them fully. Other characters are not discussed in order to save space. Characters 169-176 are new and thus all listed.

Rieppel and Reisz (1999 p.3, paragraph 2) listed, without discussion, additional changes they also made to the matrix of deBraga and Rieppel (1997); these are not adopted here until their supporting reasons are presented.

19. Well-known lathanosuchids have state 0 (Ivachenko, 1987)
32. Both *Acleistorhinus* and lathanosuchids have state 0 (Ivachenko, 1987; Lee, 1995; deBraga and Reisz, 1996).
41. *Acleistorhinus* is coded as unknown following deBraga and Reisz (1996), where the area is noted to be poorly preserved and the relevant suture not discussed or illustrated in the reconstructions.
51. Lower temporal fenestra absent (0); present (1). The complex multistate character 51 of deBraga and Rieppel included three binaries – the presence of the lower temporal fenestra (this character), the ventral opening of the fenestra (character 169 below), and the contribution of the quadrato-

- jugal to the fenestra (correlated with the presence or absence of the quadratojugal). Contrary to de Braga and Rieppel (1997), nycteroleterids lack a lower fenestra (Ivachnenko, 1987).
63. Younginids (Gow, 1975; Evans, 1987), millerettids (Gow, 1972) and all pelycosaurs (Romer and Price, 1940) retain the ventral otic fissure.
 65. Pareiasaurs have basal tubera which are formed by both the basisphenoid and basioccipital, in contrast to the basal tubera of turtles and other taxa which are formed entirely by the basioccipital. All basal tubera are here considered sufficiently similar to be potentially homologous in this study; however, Rieppel and Reisz (1999) suggest that they are different enough to fail the similarity test of homology.
 72. This character is here redefined as basicranial articulation unfused (0) or fused (1), rather than the inferred functional states of palate “kinetic” or “akinetic”. The palate in *Acleistorhinus* and lanthanosuchids is not fused; rather, it is a tight but non-sutural joint. Under the redefined character they now have state 0.
 74. Millerettids, captorhinids, nycteroleterids and procolophonids do not have a suborbital fenestra (only a small foramen, which has been coded as a separate character, 170). See Laurin and Reisz (1995).
 82. The mandibular joint is here considered to be primitively in front of the occiput in turtles, since this is the condition found in all basal turtles such as *Proganochelys*, *Kayentachelys* and *Australochelys* and most pleurodires and cryptodires (e.g. Gaffney, 1990; Rougier et al., 1995). In contrast, Rieppel and Reisz (1999) code this character as polymorphic in turtles.
 97. The condition in lanthanosuchids is known; they have state 0 (Lee 1995).
 103. The narrow ventral margin on the cervicals of pareiasaurs is considered potentially homologous to the ‘keels’ on diapsids; Rieppel and Reisz (1999) consider them ‘ridged but not keeled’, thus arguing that they are not similar enough to warrant conjecture of homology. This again is a subjective disagreement that is difficult to resolve. However, it should be noted that basal turtles exhibit a very similar morphology to that of pareiasaurs, which Rieppel and Reisz accept as potentially homologous to that of diapsids.
 120. Pareiasaurs are here coded with two coracoid ossifications (e.g. Boonstra, 1932, Gregory 1946); Rieppel and Reisz (1999) code them with one but did not cite the source.
 121. Pareiasaurs are here coded with the coracoid foramen within the coracoid, at least when the lateral surface of the scapulocoracoid is considered (e.g. Boonstra, 1932; Gregory, 1946); Rieppel and Reisz (1999) code them with the foramen between the coracoid and scapula without citing a source.
 127. An ectepicondylar foramen is found in most turtles, including all basal forms (e.g. Gaffney, 1990), and thus coded as primitively present. It is absent mainly in highly derived aquatic forms. Rieppel and Reisz (1999) code either state as potentially primitive for turtles without stating reasons.
 128. The condition in lanthanosuchids is known; they have state 0 (Lee, 1995).
 140. Pareiasaurs lack a discrete fourth trochanter, i.e. a discrete flange on the ventral surface of the femur. Rieppel and Reisz (1999) interpret them as having a ‘weak 4th trochanter ... which has shifted to the edge of the femur’, and thus recognise a weaker structure in a different position as potentially homologous. This again is a subjective disagreement that is difficult to resolve. Regardless of the potential homology, pareiasaurs lack a *well-developed* fourth trochanter, hence the coding adopted here.
 152. The first distal tarsal is present in most turtles, including all adequately known basal forms (e.g. Gaffney, 1990); Rieppel and Reisz (1999) code either presence or absence as potentially primitive for turtles without citing reasons.
 169. Lower temporal fenestra closed ventrally (0); open ventrally (1). A subdivision of DR51; not applicable in taxa without a fenestra.
 170. Suborbital foramen. Absent (0), present (1). LR49
 171. Quadrate flange of pterygoid. Long, almost reaches quadrate condyle (0); short, does not approach quadrate condyle (1). From Lee (1995).
 172. Exoccipitals. Without lateral flanges (0); with lateral flanges (1). From Lee (1995).
 173. Anterior end of splénial. Tapered (0); forked (1). From Lee (1995).
 174. Foramen jugular arterius. Small (0); large (1). From Lee (1995).
 175. Epiphysis of atlas vertebra. Present (0); absent (1). From Lee (1995).
 176. Paired anterior and posterior concavities in neural arches. Absent (0); present (1). From Lee (1995).
- Soft anatomical characters (177-216)*
- Alphanumeric codes following descriptions refer to the numbering system of Gauthier et al. (1988).
177. External auditory meatus. Small (0); large (1). G23.
 178. Secondary tympanic membrane and enclosed metotic fissure. Absent (0); present (1). A21.
 179. Extracolumella. Without Huxley’s foramen (0); with Huxley’s foramen (1). A25 modified.
 180. Nasal conchae. Absent (0); present (1). G24a.
 181. External nasal gland. Within nasal capsule (0); outside nasal capsule (1). A6.
 182. Cochlea. Short (0); elongate (1). G27.
 183. Cerebellum. Small or moderately developed (0); enlarged (1). G7 modified.
 184. Olfactory bulbs. Without peduncles (0); with peduncles (1). A14.
 185. Dorsoventricular ridge of telencephalon. Absent (0); present (1). L15.
 186. Neurofilament proteins. Two types (0); three types (1). L9.
 187. Pineal gland. Sensory (0); secretory (1). G13 modified.

