Character Construction in Morphological Phylogenetics, and the Affinities of Turtles

Simon R. Harris

A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of PhD in the Faculty of Science, Department of Earth Sciences.

Word Count: 78,043

For my mother and father, who have encouraged me throughout

Abstract

The process of character construction, by which morphological variation is partitioned into discrete characters and character states and the taxa under investigation are scored for those characters, is the most important process in any phylogenetic analysis of morphological data. Currently, the number of cladistic analyses being carried out each year is increasing dramatically, but little work has been done to try to improve the inherently subjective character construction process.

Here it is shown that there is a problem of inconsistency in character construction between workers in different fields of biology and between individual authors. Although the infinite variability of morphological variation makes it impossible to describe a definitive procedure for character construction, guidelines are presented for the best available character construction in specific circumstances. The application of these guidelines would be a start in improving the quality of data subjected to cladistic analysis and increasing the consistency, and therefore repeatability and comparability, of phylogenetic studies.

Analytical methods for identifying conflict in datasets are described, including two new methods that improve upon those previously available. These methods, which are implemented in a new computer program, *Boildown*, are utilised in an attempt to provide new evidence in the debate over the affinities of turtles, which using conventional analysis techniques has become deadlocked. Although this debate seemingly cannot be satisfactorily resolved until new fossil evidence is unearthed, both the analytical methods and a process of improving the character constructions and scorings in previously published analyses of the group produced a diapsid placement for turtles, and this combined evidence is compelling given the lack of other resolution to the debate.

In the light of the proliferation of the field of systematics in zoology, and the vital importance of character construction in that field, the process deserves, and requires far more attention in order to eliminate its image of a black box and to provide advice to those attempting to construct phylogenies. Without more discussion of the problems involved in character construction, morphological phylogenetics may never be thought of as an objective, or scientifically rigorous process.

Acknowledgements

First and foremost I would like to thank my PhD supervisors, Prof. Michael Benton, Dr. Dave Gower, and Dr. Mark Wilkinson for all of their help during the production of this dissertation. I would especially like to thank Mark and Dave for their input, allowing me to raid their reference collections and for making me feel at home at the NHM.

I would also like to thank the many friends who have shared an office with me during my PhD: Max Langer, Duncan Stewart and Kate Harcourt-Brown made me feel comfortable when I nervously moved into their office four years ago. Kate especially made me feel welcome, by giving me a crash course in PAUP (and Pacman). I must also mention the coffee club members, Max Langer (the Brazilian), Davide Pisani (the Italian), Manu Fara (the French), and Adam Yates (the Aussie) for teaching me a lot, although not always about palaeontology. Davide particularly was of immense help with all of the cladistics questions I confronted him, and also a great friend. I would also like to thank Simon Loader in London for keeping me sane during my time in the museum, and to the late Garth Underwood, who is a great loss to the scientific community. Returning to Bristol I found my office inhabited by a new batch of PhD students, who also deserve my gratitude. Keith DeBlanger has made me realise quite how many toys there are to be found on the internet and Stewart Knott has suffered alongside me as we both struggled to write our dissertations.

Finally, I would like to thank two people I've met through palaeontology who have become two of my best friends and have taken the brunt of most of the complaining when it all seemed too much. Erik Tetlie has inspired me against my will to gain an interest in eurypterids, and Rachel Moore, who has suffered the most, and has progressed with me through undergraduate and postgraduate years with surprisingly few arguments (I'll get the milk next time, honest!!!). And finally thanks to the many other friends I've made during my time in Bristol and have not time or space to mention individually.

Author's Declaration

I declare that the work in this dissertation was carried out in accordance with the regulations of the University of Bristol. The work is original, except where indicated by special reference in the text, and no part of the dissertation has been submitted for any other degree. Any views expressed in the dissertation are those of the author.

Signed: ..... Date: .....

Contents

Figures

Chapter 1: Introduction

Chapter 2: A Survey of Character Construction Methods

Chapter 3: Character Construction and the Phylogeny of Aetosaurian Archosaurs (Reptilia, Diapsida)

Chapter 4: Compatibility Methods and Their Uses

Chapter 5: The Boildown Computer Program

Chapter 6: The Stagnation of Phylogenetic Debates and the Affinities of Turtles

Chapter 7: A Closer Look at Turtle Morphology

Tables

Chapter 2: A Survey of Character Construction Methods

Chapter 3: Character Construction and the Phylogeny of Aetosaurian Archosaurs (Reptilia, Diapsida)

Chapter 4: Compatibility Methods and Their Uses

Chapter 6: The Stagnation of Phylogenetic Debates and the Affinities of Turtles

Chapter 7: A Closer Look at Turtle Morphology

Chapter 1: Introduction

1.1 The Origins of Classification

It is human nature to attempt to classify the things around us into groups based on similarity. Between the ages of just one and two years old children first begin to group objects based on simple attributes, such as colour. Such grouping is a necessity for the development of complex language in which commonly understood nouns are vital for communication to be efficient. Organisms therefore became known by common names that often implied relationships or hierarchies based on physical similarities. Over 2,000 years before Darwin (1859) famously published his groundbreaking explanation of the complexity of life, Aristotle differentiated plants from animals, and even separated plants into three groups based on differences in the morphology of their stems. This was the precursor of modern taxonomy, homology assessment and phylogenetics, and the basis for the concept of evolution. However, by the fifteenth and sixteenth centuries Aristotle's classification was becoming seen as inadequate. At this time many more formal classification schemes were being produced. The most famous of these was by the Swedish zoologist, Carl von Linné. Linné also produced a *scala naturae*, or nature's ladder, which placed all organisms in a hierarchy from inanimate objects at the base to humans at the pinnacle. The Linnean system remains the classification system of choice for most taxonomists today. However, with the introduction of the concept of evolution the way relationships between organisms were thought about began to change. Soon after the publication of Darwin's (1859) On the Origin of Species by Means of Natural Selection, relationships began to be represented on tree-shaped diagrams to represent the branching process of evolution (Fig.1.1). However, the relationships presented in these early trees often still suffered from the influence of religious and political beliefs.

1.2 Cladistics

In the last century phylogenetics benefited from a relaxation of the concept of man being at the pinnacle of evolution. Trees began to represent a more scientific view of morphological relationships, which led to more accurate phylogenies. However, these trees would now still be considered less than rigorous, in the sense that they were produced purely on the basis of a personal opinion of the available evidence, and this process was often impossible to replicate. Therefore, Hennig (1966) presented a new method of tree

production that classified organisms based on shared, derived characters. This method, known as cladistics, was extended to allow trees to be produced from observed data using mathematical procedures that were repeatable. Since its introduction the number of phylogenies published has increased dramatically, so that now huge numbers of cladistic analyses are carried out every year.

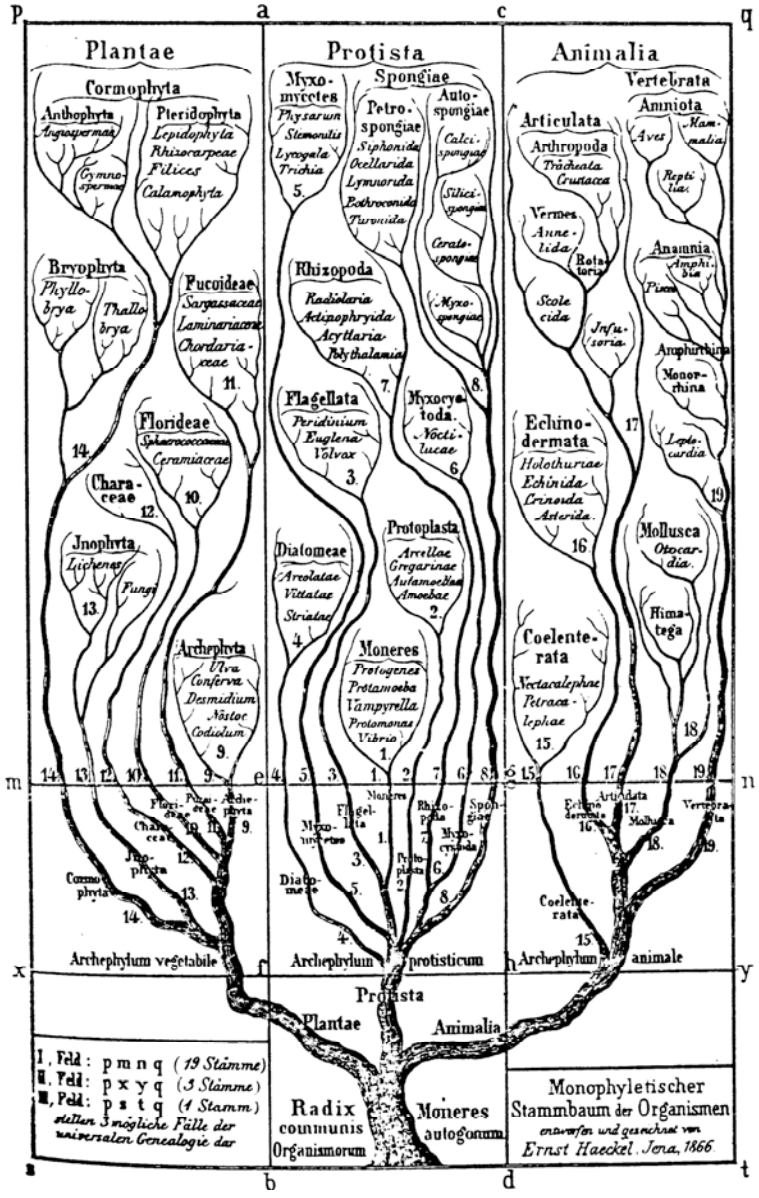


Figure 1.1. An early phylogenetic tree of life drawn by Ernst Haeckel in 1866.

Morphological phylogenetic analyses comprise two important procedures: matrix preparation, where morphological data are represented as a taxon versus character grid, and quantitative data analysis, where relationships between taxa are identified using analytical procedures. The latter of these, the analysis of the data matrix, has received virtually all of the attention in the literature, with the development of many optimality

criteria upon which the best tree is defined (e.g. distance, maximum likelihood, maximum parsimony), techniques for finding the optimal tree (exhaustive, branch and bound and heuristic searches of tree space including nearest neighbour interchange, tree pruning and regrafting and tree bisection and reconnection), methods for quantifying support for clades on trees (e.g. bootstrapping (Felsenstein, 1985), decay analyses (Bremer, 1988; Donoghue *et al*., 1992)), and many powerful software implemetations for carrying out such analyses on large data matrices (e.g. Swofford, 2003). This quantitative stage of the cladistic analysis is, if all input parameters are reported, repeatable, and any results obtained should be reproducible and falsifiable, two necessities for any scientific method if its results are to be testable by other workers.

The initial stage of morphological phylogenetic analysis, in which observed variation is partitioned into characters and character states, is not so clearly objective or obviously repeatable. Representing the plasticity of morphological variation with discrete characters lends itself to criticism as a subjective process that is not reproducible by other workers. Analytical results are simply representations of information derived from the processes of character construction and scoring. Therefore, if the compilation of input data is subjective, then so are the results obtained. In simple terms, garbage in leads to garbage out. For this reason, the construction of the data matrix is undoubtedly the most important part of any phylogenetic analysis.

Phylogeneticists have recognized a number of methodological issues concerning character construction, including the treatment (ordered or unordered) of multistate characters (e.g. Hauser and Presch, 1989; Wilkinson, 1992a), the interpretation of complex structures as complex characters or character complexes (Pleijel, 1995; Wilkinson, 1995a), the treatment of polymorphism (e.g. Wiens, 1995; Kornet and Turner, 1999), and the representation of inapplicability (e.g. Maddison, 1993; Strong and Lipscomb, 1999 and see Chapter 2). Practicing phylogeneticists necessarily confront issues of character construction, and the approaches they adopt have practical consequences for what they can infer using numerical phylogenetic methods. Despite this, there has been surprisingly little discussion of generalities. Chapter 2 of this thesis confirms the results of Hawkins' (2000) survey, which demonstrated the existence of a variety of approaches to character construction, but both studies also found that there was little discussion of why any particular approach was selected. Similarly, Poe and Wiens (2000) found that few workers provided any explicit justification for the approaches they adopted to morphological character selection. The comparison of alternative approaches to character construction,

although important, is still in its infancy and deserves more attention (see also Wiens, 2001; Rieppel and Kearney, 2002), as it has been shown in a number of studies that applying different coding strategies to the same morphological data can lead to the production of very different trees (e.g. Wilkinson 1995a; Harris et al. 2003a).

1.3 Homology and the Character Concept

1.3.1 Homology, What's in a Word?

Probably the most important issue in comparative biology, and therefore in the creation of the data matrix for cladistic analysis, is that of homology assessment. However, the literature on homology is vast and inconsistent, and arguments still rage as to the best definition of 'homology' (see Patterson, 1982 for a discussion of some of the many published definitions). The word homologous, in the non-biological sense, is defined in the Oxford English Dictionary simply as "corresponding", the etymology of the word being from the Greek *homo* meaning 'same' and *logos* meaning 'word'.

Popper (1962: 14) wrote, "definitions do not play any important part in science". Although this extreme view is rejected here, it does highlight an important point. The philosophy of definitions of words is extremely complex, and is not the key in the homology debate. What is important is having a consistent concept of homology. Bock (1973) discussed definitions of words in a simplified way and raised a number of important issues. He believed that what is of paramount importance is that the usage of words, such as homology, is consistent. He suggested that flaws can be found in any operational definition (a clear, concise, detailed definition), and so rejected the notion that only operational definitions are valid in science. Instead he favoured what he called a "theoretical definition", by which the definition of a word is the concept, an idea I agree with. Similarly, Rieppel (1980) realised that the problem is not the definition of the word homology, but rather the operational criteria used to recognise it. For this reason, here homology is discussed as a concept that cannot necessarily be defined by a catchy single sentence as has so often been attempted in the past.

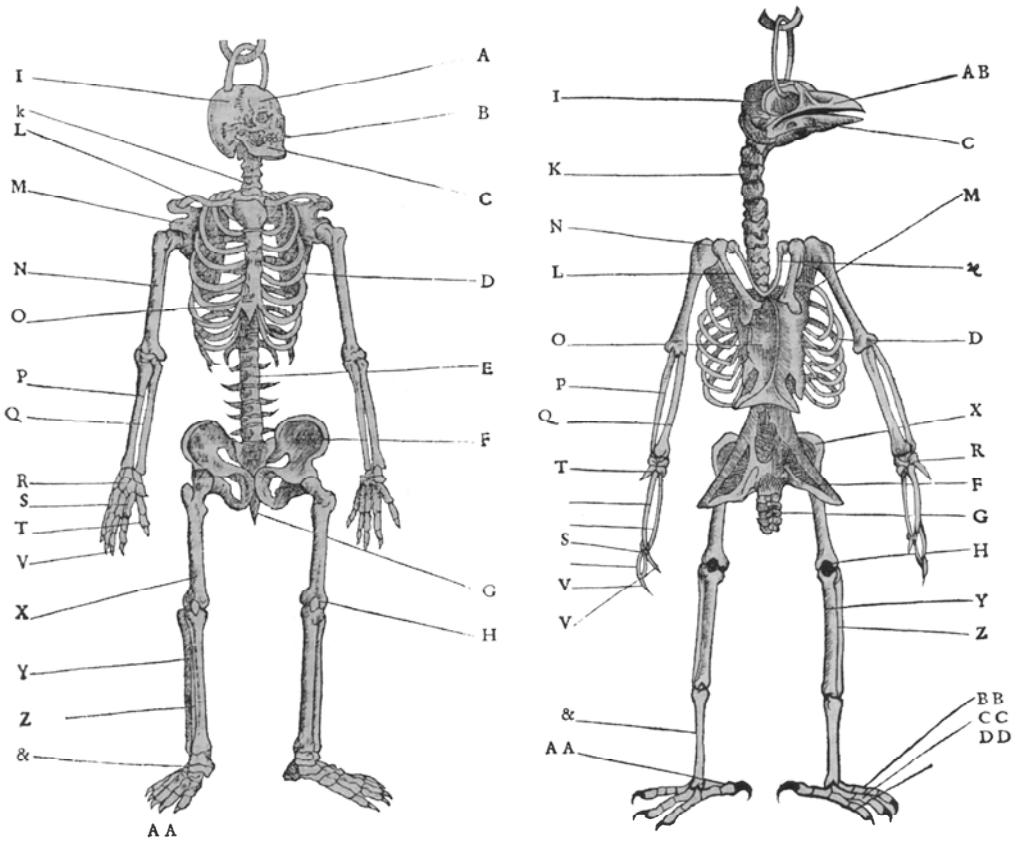


Figure 1.2. Belon's (1555) diagrammatic representation of homologies between the skeletons of humans and birds.

It is impossible to discuss homology without looking at the history of the concept. When early biologists were beginning to group organisms they did so on the basis of observed similarities. By the fourteenth century it had been noticed that even seemingly very different organisms appeared to be built on very similar basic body plan. The most famous example of such thinking can be seen in Belon's (1555) much-reproduced figure illustrating the similarities between the skeletons of birds and man (Fig. 1.2). By the nineteenth century the means by which correspondence of morphological elements in different organisms became a key area of debate. In 1818 Étienne Geoffroy Saint-Hilaire introduced his *théorie des analogues*, which identified what he called analogous parts based on their relative positions (his *principe des connexions*). Saint-Hilaire (1818) did not, however, use the word homology. Richard Owen (1843: 374) did use the word homologue, and famously defined it as "The same organ in different animals under every variety of form and function". Later (1847: 175-6) he split this definition into two, special homology, "the corresponding of a part or organ, determined by its relative position and connections, with a part or organ in a different animal" and general homology, "a relation

in which a part or series of parts stands to the fundamental or general type". He also recognised that the function of structures can cause mistaken homology assessment, and called this analogy. This change of the usage of the word analogy makes sense when its general, non-biological definition is considered. In the Oxford English Dictionary, analogous is defined as "similar in certain attributes, circumstances, relations or uses; having something in parallel". Unfortunately, however, Owen's use of the word analogy was not consistent (Ghiselin, 1976).

It was with the introduction of the theory of evolution that the concept of homology moved on once more. Owen's definitions of homology have often been criticised for including the phrase "the same" organ or part without explaining on what basis sameness should be judged (e.g. Lankester, 1870; Bock, 1963; Ghiselin, 1976). However, as pointed out by Ghiselin (1976), the world in which Owen lived was very different from today. Creation was still accepted in parts of the scientific community, and for this reason Owen maintained that the similarities seen between different organisms were variations of the same part of the idealistic archetype from which a group of organisms was created by "God to deviate" (Van Valen, 1982). Darwin considered homology and the *principe des connexions* as evidence for his theory of evolution (1859), which brought with it a philosophical basis for classification (Bock, 1973) and the concept of homology had to be adapted. Roth (1984: 13) dramatically began her review of definitions of homology with the statement "After 1859 there has been only one definition of homologous that makes biological sense: "A feature [character, structure, and so on] is homologous in two or more taxa if it can be traced back to [derived from] the same [a corresponding] feature in the presumptive common ancestor of these taxa." (Mayr, 1982: 232, brackets are his)". This evolutionary definition of homology is more often, and more simply written as "structural similarity due to common ancestry" (e.g. Boyden, 1973). Roth (1984) tried to add a developmental aspect to this definition by suggesting that a necessary component of homology is the sharing of a common developmental pathway. Unfortunately, however, it is rarely simple to use such developmental pathways to produce hypotheses of homology, especially in the world of palaeontology.

1.3.2 The Bat Wing Bird Wing Problem

An incessantly recurring example in the homology literature is that of the wings of birds and bats, so I will briefly discuss it here. Both of these vertebrates have a forelimb that is adapted for flight, but it is universally accepted that their wings evolved in two

independent events. Most authors simply state that the forelimbs of birds and bats are homologous as arms, but not as wings. This is obviously true. The highly superficial apparent homology between the two types of wing is based solely on function, and any closer examination would show that the wings are not even formed from the same structures, with the bird wing being formed by feathers along a strut-like, largely fused hand and the bat wing by skin stretched between highly elongated fingers. The problem is simply one of names. It seems that the only reason homology is being suggested between the two types of wing is that they are both called wings. By thinking about the meanings of the names forelimb and wing, it becomes clear why such an approach is ill-advised. Forelimb, or arm, in the context of vertebrates, is the name of a structure based on shape (limb = a part or member of the body distinct from the head or trunk according to the Oxford English Dictionary) and position (fore = front). Shared shape and position, as already mentioned, are features of homologous structures. The Oxford English Dictionary definition of wing, on the other hand is "each of the organs of flight of any flying animal, as a bird, bat or insect". The basis of this definition is the function of the wing, a feature not of homology, but of analogy.

1.3.3 Homology ≠ Synapomorphy

Although unhappiness with the exact operational definition of homology continued, the concept seemed to have become relatively consistent with the concept of common ancestry. However, the whole subject was thrown back into disarray with the introduction of phylogenetic systematics in the 1960s (Hennig, 1966), which led to a number of authors (e.g. Wiley, 1975; Eldredge, 1979; Patterson, 1982; see de Pinna, 1991: 370 for a list of others) attempting to equate homology with synapomorphy. Patterson's (1982) view on homology was that it should only be thought about in the phylogenetic framework, because "every worthwhile hypothesis of homology specifies a hierarchy of groups" (Patterson, 1982: 34). He therefore postulated the definition: "Homology is the relation which categorizes monophyletic groups" (1982: 21) i.e. synapomorphy. With this definition in place, Patterson went on to describe a number of tests of homology, which can be split into three subsections:

- *The similarity test*. The initial judgement of similarity between parts is not, in fact, a true test of homology, but rather is the factor that leads to the proposal of potential homology (Patterson, 1982; de Pinna, 1991). However, Patterson (1982) included

similarity as one of his three tests of homology. He said that this similarity could be "topographic, ontogenetic, histological etc.".

- *The conjunction test*. If two structures are identified as possible homologues, then according to the conjunction test, this potential homology is refuted if both structures are present in one organism.
- *The congruence test*. Wiley (1975: 235) postulated "the only valid test of homology…is to hypothesize that the supposed homology is a synapomorphy". Similarly, Patterson's (1982) test of congruence is based on the idea that the best test of a hypothesis of homology is congruence with other hypotheses of homology. He suggested both parsimony (Patterson, 1982) and compatibility (Patterson, 1988) as potential methods by which this congruence can be assessed.

More recently, Patterson's (1982) ideas have been developed further, and the distinction between homology and phylogenetic analysis have been reduced further. De Pinna (1991) introduced the terms primary and secondary homology to describe the stages of identification of potential homologues and the testing of those potential homologues using parsimony, respectively. His primary homology is essentially the process of production of a data matrix for cladistic analysis by identifying, and defining characters, coding character states and scoring taxa for those characters (Hawkins *et al.*, 1997). This process involves the similarity and conjunction tests of Patterson (1982). Secondary homology is then equivalent to synapomorphy on the phylogenetic tree produced by analysis of the data matrix. Primary homologues that are not secondary homologues can then be re-examined to try to determine whether it is the primary homology assessment or the tree that is causing the disparity. This process is known as reciprocal illumination (Hennig, 1966).

Here the equivalence of homology and synapomorphy is not accepted. Roth (1988) amongst others noted that such equivalence restricts the meaning of the term homology. In reply, de Pinna (1991) argued that this restriction is the strength of the argument, as it specifies the meaning of homology, which in the past had been difficult to pin down. I agree with Roth that homology has a broader meaning than synapomorphy. However, it is implied by her argument that synapomorphy is included in homology. I disagree, and suggest that the two are different concepts that only become equivalent in certain circumstances.

To explain my reasoning, I will first define what I understand by some of the relevant terminology.

- Primary homology, under any of its alternative names is simply a hypothesis of homology based on perceived similarity of some kind. As described by de Pinna (1991), it is this process of homology assessment that is used in the production of data matrices upon which systematic analyses can be carried out. Primary homology assessment includes both the similarity and conjunction tests identified by Patterson (1982).
- Patterson's (1982) congruence test was originally a means by which agreement between hypotheses of homology can be assessed. He showed how this can be achieved using parsimony (1982) or compatibility (1988). To Patterson, incongruence between two hypotheses of homology meant that at least one of those hypotheses is incorrect.
- A synapomorphy is a derived character state that unites two or more taxa in a monophyletic group on a phylogeny.
- The term secondary homology is generally used for those primary homologies that unite monophyletic groups in the results of a congruence test using parsimony. As such, secondary homology is synonymous with synapomorphy.

This brings me to my concept of homology. Of the published definitions, that which describes best my personal view is that of Van Valen (1982), who defined homology as "resemblance caused by a continuity of information". Van Valen (1982) recognised that the concept of homology encompasses many related terms, and for this reason his definition was very general. He used a very interesting example to explain his definition: "Whether the pyramids of Mexico are homologous or convergent to those of Egypt depends on whether the basic information was transferred or originated independently". This definition encompasses the traditional evolutionary definition and the developmental definition of Roth (1984). Brigandt (2002) criticised Van Valen's (1982) definition exactly for its generality. He disliked both the use of the phrase "continuity of information", because homologous structures can arise from different genes and/or developmental pathways (e.g. see Wagner and Misof, 1993; Hall, 2003), and the word "correspondence", because it is "exactly the concept that an account of homology has to clarify" (Brigandt, 2002: 394). However, I maintain that even when homologues arise from different genes there must be some sort of conservation of information, and Brigandt's argument is possibly better aimed at Roth's (1984) developmental definition of homology than Van Valen's (1982) continuity of information. Further, the criticism of the circularity of Van Valen's (1982) use of the word correspondence in his definition of homology is mistaken,

since Van Valen did not use this word. Instead he used resemblance, which is more easily understood, and implies the identification of resemblance that must be made in order for a primary homology to be suggested. Furthermore, Brigandt (2002) criticises, but does not propose a viable alternative to, all current concepts of homology. He simply calls for a more precise definition which "has to make reference to the different concrete mechanisms that are at work within cells and organisms" (Brigandt, 2002: 405).

The crux of the difference between homology and synapomorphy is that homology is the truth, whereas synapomorphy is a hypothesis. Systematic analyses produce hypotheses of relationships (phylogenetic trees) from hypotheses of homology (primary homologies) using a hypothesis of evolution (parsimony). The tree produced is not necessarily the true evolutionary history of the group of taxa under study. Therefore, two different analyses of the same set of taxa can produce radically different trees, meaning different derived character states will be synapomorphies. Therefore, only when a synapomorphy unites a monophyletic group that is present in the true evolutionary history of the taxa analysed can it be a true homology. However, even then the two concepts are not necessarily the same. There are occasions, according to my understanding of the concept of homology, that even on the true tree a homology is not a synapomorphy. Patterson (1982) himself admitted this could occur in the case of characters involving secondary loss of a structure. Van Valen (1982) used the example of teeth in birds. Kollar and Fisher (1980) showed that recent birds retain the information required for constructing teeth. It is therefore not unfeasible that teeth could be expressed again in a bird of the future. These teeth would have continuity of information with, share developmental pathways with and be derived from the same structure in a common ancestor as the teeth of Mesozoic birds such as *Archaeopteryx*. Van Valen (1982) called this phenomenon latent homology. The two sets of teeth are homologous even though on the true phylogeny of birds they would not constitute a synapomorphy. Real occurrences of the same phenomenon, as identified in this hypothetical example, have been published. For example, it has been shown that highly complex wings with the same genetic control have appeared many times in stick insects, suggesting that they must be retained from a common ancestor of the group (Whiting *et al.*, 2003). These wings must be homologous, since the only factor affecting their presence or absence is a change in a gene that controls their phenotypic expression. However, on the phylogenetic tree of stick insects they appear to be highly homoplastic and therefore do not form a synapomorphy. Latent homology is a problematic side effect of secondary loss of

complex features, which is one of the major problems in the concept of homology and the process of character construction in systematics. Secondary loss is discussed further below.

Conversely, it is also plausible that a hypothesised homology that forms a synapomorphy on the true phylogeny of a set of taxa may not be truly homologous. This, however, can only occur when a mistake is made when proposing the initial hypothesis of primary homology. De Pinna (1991) made the rather contentious claim that "one might speculate that if bats and birds happened to be sister groups, there would not be fierce contention to considering their wings as homologous, despite all the anatomical differences". Ignoring the fact that virtually no vertebrate morphologist would propose primary homology between these two types of wing (see section 1.3.2 above), de Pinna's (1991) example, which was aimed at showing why homology and synapomorphy can be considered equivalent, is here considered to do the exact opposite. If the wings of birds and bats were proposed as primary homologues, and if, as in de Pinna's hypothetical example, birds and bats were sister-groups on the true phylogeny of vertebrates, this primary homology would be a synapomorphy. However, the two structures would still fail to meet the criteria that define homology. They would not have continuity of information with, share developmental pathways with and be derived from the same structure in a common ancestor, and therefore would not be homologous.

1.3.4 The Opposite of Homology: Analogy, Homoplasy or Nonhomology?

Equally as confusing in the literature as the definition of homology is identifying what is the opposite, or antonym, of homology. Ghiselin (1969; 1976) used the term analogy in this way. However, analogy is used more widely in a sense more similar to Owen's (1843: 374) definition: "a part or organ in one animal which has the same function as another part or organ in a different animal". Ghiselin (1976) criticised Owen's definition, as it suggests that parts can be both homologous and analogous. He provided a new definition of analogy: "entities are analogous when they are nonhomologous corresponding elements of a class of more inclusive entities which share properties suggestive of a genuine homology between those elements" (Ghiselin, 1976: 138). Although Owen's definition is flawed, as shown by Ghiselin, the term analogous does tend to be used in cases where nonhomology is caused by convergence in form due to common function. As discussed above with homology, it is not a short, catchy definition of analogy that is important, but that the concept is used consistently. For that reason, I prefer to define analogy in the way it is more often used, as "nonhomologous corresponding

elements of a class of more inclusive entities which share properties suggestive of a genuine homology between those elements due to common function."

A second term that has been used for the opposite of homology is homoplasy, especially by those who support the idea that synapomorphy is equivalent to homology. However, in general the term homoplasy is always thought of in terms of phylogenetic trees, and for this reason makes a better opposite of synapomorphy than homology (Fitch, 2000). To use it as the opposite of homology would confuse my argument for the inequality of synapomorphy and homology.

Instead of either analogy or homoplasy, I consider that the best term for the opposite of homology is simply nonhomology. However, I use this term in exactly the sense Ghiselin used analogy, with a slight change in the definition to avoid circularity: "corresponding elements of a class of more inclusive entities which are not homologous, but share properties suggestive of a genuine homology between those elements".

1.3.5 Types of Homology

It is well recognised that homology refers to a family of related concepts (e.g. Ghiselin, 1969; 1976; Patterson, 1982; Van Valen, 1982; Roth, 1984). Ghiselin (1976) separated what he called evolutionary homology, which is the general case of homology between two different organisms, and iterative homology. This second type of homology, which refers to homology between repeated structures in one organism, has always been problematic, because it is difficult to reconcile with many definitions of homology. However, there is no such problem when using the definition preferred here, because the resemblances between the repeated structures are caused by a continuity of information, even though they are expressed in a single organism. Owen recognised that serially-related bones in many organisms also exhibit a form of homology, which he termed serial homology. However, as noticed by Ghiselin (1976), this is a misnomer. Although in many cases repeated homologues are in a serial arrangement, there are also many cases, such as mammalian hair and feathers where there is no such serial arrangement. The name iterative homology is more inclusive. Ghiselin's (1976) iterative homology included three subtypes:

- *Serial homology*. Iterative homologues repeated in a linear sequence, such as vertebrae.
- *Antimeric homologues*. Iterative homologues due to bilateral symmetry of the body, such as right and left femora.

- *Sexual homologues*. Corresponding reproductive parts in species in which a single kind of reproductive apparatus is repeated with variations in individuals having different reproductive roles, such as the penis and clitoris.

I also use the term intraorganismal homology (see Harris *et al*., 2003a), for those homologues that are present in a single organism, and therefore do not comply with Patterson's (1982) conjunction test. Intraorganismal homology includes both serial homology and antimeric homology, and is discussed further in Chapter 3, so will not be treated in detail here.

1.3.6 Homology Assessment for Phylogenetic Analyses

Although homology assessment is principally a component of comparative biology, as previously mentioned, it also plays a central role in the creation of data matrices for phylogenetic analysis. However, not all homology is useful for such analyses. Ghiselin's (1976) distinction between evolutionary and iterative homology is essentially a distinction between those types of homology that are useful for phylogenetics, and those that are not. For this reason, Roth (1984) replaced the term evolutionary homology with phylogenetic homology.

It is not completely true to say that intraorganismal (iterative) homology is useless for phylogenetic reconstruction, but it is problematic. Many of these problems are discussed in detail in Chapter 3. Sexual homology is a very good example of a problematic type of homology for phylogeny reconstruction. It is often the case that sexual homologues look more similar between species than sexual dimorphs of the same species. Such characters could lead to reconstruction of a phylogeny with females and males branching at the point that the two types of sexual organs started to diverge.

Supporters of the equivalence of homology and synapomorphy generally do not consider intraorganismal homology as homology (e.g. de Pinna, 1991), or do not mention it at all (e.g. Patterson, 1982). This is because it does not fit with their definition of homology. As de Pinna (1991: 376) states, "serial homology and taxic [essentially meaning interorganismal] homology are fundamentally different from the perspective of systematics, which is concerned with interorganismic hierarchical organization. Serial homology, an intraorganismic kind of order, is therefore outside the realm of most present-day systematics". Essentially, this is an admission that some homology is not equivalent to synapomorphy, because some homology cannot be used for phylogeny reconstruction.

As well as not all homology being useful in character construction, not all characters are based on true homologies. One highly problematic type of character is that coding the gain or loss of an element. Often gains are regarded as some of the strongest evidence for relationships, whilst losses are not so highly regarded (e.g. see Hecht and Edwards, 1977). The problem with these types of character is the idea of homology of lacking an element. It can be said that the element is homologous in those taxa that possess it. The problem is with the taxa that lack the element. Proposing the lack of an element as homologous between two taxa is difficult, as it cannot be said to meet the similarity criterion. However, in many character constructions taxa are united by a character state defined by lack of an element, and such characters can be synapomorphies.

1.3.7 The Character Concept

Like homology, the concept of the phylogenetic character has been the subject of a great deal of debate in the literature, and there have been whole books dedicated to the subject (Wagner, 2001). The concept of the cladistic character is of paramount importance, because characters are the data that determine the results of systematic analyses (Richards, 2003). Hennig (1966), when introducing his idea of phylogenetic systematics, noted the importance of the cladistic character, defining it as "a historical event in the evolution of a feature". However, like homology, as noted by Colless (1985), the term "character" is regularly used in several related ways, and not necessarily in the way Hennig intended (Grant and Kluge, 2004). Unless the meaning of the term is universally understood and used in a uniform way, the results of systematic analyses may not be consistent. The problematic nature of the concept can be seen in the Oxford English dictionary, which has many definitions for the word character. The two most relevant of these are in the literal sense as "a distinctive mark impressed, engraved, or otherwise formed", and in the figurative sense as "a feature, trait, characteristic". Similarly, Colless (1985: 230) listed 18 definitions of character that he found in just 50 papers in the cladistic literature. He clustered these definitions into three categories:

1) The ancient, dictionary meaning, which is an attribute, feature, trait or characteristic. This implies a fact about the subject that can be true or false. Henceforth this will be called a characteristic.
2) A part. This includes physical parts, such as a specific bone or muscle, and also more abstract parts, such as forms of behaviour. Henceforth called a part.

3) A set of mutually exclusive attributes of a part, sometimes known as a variable or character-variable. Each part may have many variables. Henceforth called a variable.

So, using an example of a taxon that has red feathers, the three categories of characters would be: 1) characteristic = has red feathers, 2) part = feathers, 3) variable = feather colour, in this case red, but also any alternatives, such as blue or green etc.

The most commonly used definition of a cladistic character would probably be one similar to that of Platnick (1978: 542), who said "a character comprises of two or more different attributes (character states) found in two or more specimens that, despite their differences, can be considered alternate forms of the same thing (the character)". This is a combination of the last two of Colless' (1985) categories. The variable of a part is the character, and the set of mutually exclusive attributes in that variable form the character states (Colless, 1985; Pimentel and Riggins, 1987).

The main philosophical debate regarding characters concerns what criteria should be used to individuate them. Wiley (1981) suggested that a character is any characteristic that can be communicated by one biologist to another, and Eldredge and Cracraft (1980) similarly suggest it is anything that can be named. More usually workers try to individuate characters based on independence, whether that be logical (*sensu* Wilkinson, 1995a), functional or developmental (Richards, 2003). Although philosophically character individuation is problematic, many of the important issues are covered in homology assessment (see section 1.3.1 onwards above), or by logical comparison of characters to check for non-independence.

Another debate in discussions of the character concept revolves around the distinction of "character" from "character state". A number of authors (e.g. Bock, 1973; Platnick, 1978; Patterson, 1982) have claimed that there is no distinction between these two concepts, and that characters and character states are the same thing just at different hierarchical levels (Eldredge and Cracraft, 1980). De Pinna (1991), on the other hand, attempted to differentiate the two on the basis of independence. He said that characters should be independent of one another, while the character states of a character should be transformations of one another. Although this distinction seems strong, it is also true that many character states of one character can also be a variable part in a second, producing the hierarchical levels alluded to by Eldredge and Cracraft (1980). In a practical sense it is useful to differentiate between the two concepts simply for the sake of describing characters used in a phylogenetic analysis. All characters used in an analysis should be described in a way that is easily understandable and allows others to repeat the process of

matrix construction. Therefore, consistency in character definitions is paramount. A character description should include the part and variable followed by each attribute of that variable. So, the above example should be written as:

Feather colour: red (0), blue (1), green (2)

It would be erroneous to describe it as:

Feathers: red (0), blue (1), green (2)

because there could be many variables of the part "feathers", and in some cases excluding the variable from the character description could lead to confusion. Conversely, Colless (1985) also suggested that character states should not include the name of the part.

1.4 Aims of This Study

This thesis is aimed at looking at the process of data matrix construction for phylogenetic analysis in a practical way rather than by debating the philosophical aspect of the subject, which has been much discussed in the literature. It also attempts to use alternative analytical methods in an attempt to quantitatively identify characters that are highly incongruent with other characters in a matrix on the preface that such characters are more likely to be poor assessments of homology.

Chapter 2 is a literature survey based on similar work by Hawkins (2000). This survey looked at the ways in which different authors and workers in different fields constructed characters for analysis to see if the inconsistency identified by Hawkins (2000) for botanical analyses was also true for zoology and palaeontology.

Chapter 3 takes a closer look at some of the mistakes and problems in character construction and scoring that are frequently encountered during morphological analyses. This is done by looking at the three published analyses of a group of Triassic archosaurs, the aetosaurs. Many coding and scoring problems were identified, especially regarding the treatment of the intraorganismally homologous dermal scutes possessed by these animals. A new matrix was produced in which the highlighted problems were corrected as far as possible, and a new analysis was carried out.

Chapter 4 introduces compatibility methods of analysis that are out of favour in the world of modern systematics, but offer many useful techniques for exploring data and identifying possible nonhomology. Some new methods are also introduced that take compatibility methods further.

Chapter 5 describes the features of "Boildown", a computer program written as part of this study to carry out many of the methods in Chapter 4, including the methods new to this study. Chapter 4 also acts as a manual for the program, which is included on a CD at the back of this thesis.

Chapters 6 and 7 are a case study based on the debate over the origin of the turtles. Both the analytical methods (Chapter 6) introduced in Chapter 4 and resolution of problematic character constructions and scorings (Chapter 7) are employed in an attempt to find evidence for the phylogenetic position of this problematic group. This case study has wider implications for all cases of stagnation of phylogenetic debates.

1.5 Glossary

Many terms relating to data matrix production and phylogenetic analyses are used throughout this thesis, so a glossary of relevant terms is provided here.

ANALOGY – see section 1.3 above

APOMORPHY – a derived character state

AUTAPOMORPHY – an apomorphy that defines a single taxon in a phylogenetic analysis

BINARY CHARACTER – a character that has only two defined states

BRANCH – a line on a cladogram. Internal branches connect two nodes, whilst external branches connect nodes to terminal taxa

BRANCH-AND-BOUND – a method for searching for the most parsimonious trees for a set of characters that produces an exact result. This method is more time efficient than the exhaustive search, and so can be used with data matrices including more taxa. The method first uses a heuristic search to find a relatively short tree. It then sets the length of this tree as its upper bound before searching exhaustively through all other trees. When carrying out this exhaustive search, and partially constructed trees that are as long as or longer than the upper bound are discarded, cutting the search time needed. If a new tree is found during the search that is shorter than the upper bound, the upper bound is changed to the length of this new tree

BRANCH LENGTH – the number of character state changes that occur on a branch

CHARACTER & CHARACTER STATE – see section 1.3 above

CHARACTER CODING – see CODING

CHARACTER CONSTRUCTION – the partitioning of phenotypes into discrete characters, the partitioning of variants into character states and hypothesizing the relations among them (choosing a character type)

CHARACTER SCORING – see SCORING

CLADE – a group of taxa on a tree that includes the common ancestor of that group and all of its descendants. Also called a monophyletic group

CLADISTICS – a method of phylogenetic inference that groups taxa together based on shared derived characters

CLADOGRAM – a branching diagram that shows relative relationships between a number of taxa. Time is not indicated absolutely, but relative time of branching is implied.

CLIQUE – a set of mutually compatible characters

CODING – a part of character construction in which morphological variation is split into discrete states (see Chapter 3)

COMPOSITE CODING – a method of coding complex characters in which variation that can feasibly be split into a number of binary characters are combined into a single multistate character

CONGRUENCE TEST – see section 1.3 above

CONJUNCTION TEST – see section 1.3 above

CONSENSUS (TREE) – one of a number of related methods for combining the information about relationships in a number of trees

CONSISTENCY INDEX (CI) – a measure of the amount of homoplasy a character or set of characters exhibit on a specific tree. It is calculated by dividing the minimum possible number of steps the character can take on any tree by the minimum number of steps it takes on the tree in question

CONSTANT CHARACTER – a character in which all of the taxa scored possess the same state

CONTINUOUS CHARACTER – a character in which the variable is continuous. Due to the restriction of most matrix representation methods, continuous variables are usually divided into discrete character states

CONVERGENCE – a homoplastic similarity between two taxa that has evolved independently on more than one occasion

COST – The number of steps specified for a change between two states of a character on a branch

DECAY INDEX – the number of extra steps (above the TL of the MPT) that must be added before a clade is no longer present in all trees produced by parsimony analysis

DISCRETE CHARACTER – a character in which the states are based on a discrete variable

EXHAUSTIVE SEARCH – a method for searching for the most parsimonious trees for a set of characters that produces an exact result by calculating the length of every possible resolved tree for the taxa included in the analysis

HEURISTIC SEARCH – a method for searching for the most parsimonious trees for a set of characters that does not necessarily return the most parsimonious result. Heuristic methods have the advantage over exact search methods in speed, but sometimes at the cost of accuracy

HOMOLOGY & HOMOPLASY – see section 1.3 above

INGROUP – The group of taxa in a cladistic analysis for which phylogenetic relationships are being asserted

LEAF – a terminal taxon on a tree

LENGTH – see TREE LENGTH

MAJORITY-RULE CONSENSUS – a consensus tree including all clades that are present in at least 50% (or greater) of the members of a set of trees

MONOPHYLY – see CLADE

MOST PARSIMONIOUS TREE (MPT) – The tree or set of trees that have the shortest tree length of any trees for the input data

MULTISTATE CHARACTER – a character that has more than two observed states

NODE – an internal node is a point on a tree joining two branches. An external node is the end of an external branch, sometimes called a leaf

ORDERED CHARACTER – a multistate character for which the evolutionary order of transformations between states has been specified. State changes between states that are not adjacent cost more than between adjacent states

OUTGROUP – a taxon or taxa closely related to the ingroup that is used to root a tree and identify which character states are plesiomorphic

PARAPHYLY – a group that contains a common ancestor plus some, but not all, of its descendents

PARSIMONY – a method that is often used for creating cladograms. This method assumes minimal tree length as being the most likely model of evolution

PERMUTATION TAIL PROBABILITY (PTP) – a method for looking for phylogenetic signal in a data matrix. The length of the actual data is compared with that of a number of random permutations of the same data. The PTP is the proportion of these permuted datasets that have a length equal to or shorter than the actual data

PHYLOGENY – the evolutionary relationships of a group of taxa

PLESIOMORPHY – a primitive character or character state

POLARITY – the discrimination between primitive and derived character states

POLYMORPHISM – a scoring of a taxon as possessing more than one state of a particular character

POLYPHYLY – a group of taxa based on convergences

POLYTOMY – a node at the base of three or more branches

PRIMARY HOMOLOGY – see section 1.3 above

REDUCTIVE CODING – a method of coding complex characters in which variation that is be split into a number of characters, which are usually binary

RESCALED CONSISTENCY INDEX (RC) – the product of the consistency index and the retention index

RETENTION INDEX (RI) – a measure of the amount of similarity exhibited by a character or set of characters on a specific tree that can be considered synapomorphy. It is calculated using the following formula:

$$RI = \frac{(g - s)}{(g - m)}$$

where g is the maximum number of steps the character can take on any tree, s is the minimum number of steps the character can take on the tree in question, and m is the minimum number of steps the character can take on any tree

ROOT – the basal point of a cladogram which is often defined by an outgroup. The outgroup is also sometimes referred to as the root

SCORING – the ascribing of character states to a terminal taxon

SECONDARY HOMOLOGY – see section 1.3 above

SIMILARITY TEST – see section 1.3 above

SISTER-GROUP – two taxa that are more closely related to each other than to other taxa

STEP – a change between two states of a character on a branch

STEPWISE ADDITION – the method by which taxa are added to a tree during searches for the most parsimonious tree

STRICT CONSENSUS – a consensus tree including all clades that are present in all members of a set of trees

SYNAPOMORPHY – an apomorphy that defines a clade

TAXON – a group of organisms

TOPOLOGY – the shape of a cladogram

TREE – a cladogram

TREE BISECTION AND RECONNECTION (TBR) – a method by which the positions of branches on a tree are swapped during searches for more parsimonious trees. TBR cuts off subtrees from the cladogram being analysed, changes the position of the root and then reattaches them at a new position on the rest of the cladogram

TREE LENGTH (TL) – the sum of the minimum number of character state changes for all characters on a tree

UNINFORMATIVE CHARACTER – a character that contains only autapomorphies or is constant. Uninformative characters contain no useful information for phylogeny construction using parsimony

UNORDERED CHARACTER – a multistate character for which the evolutionary order of transformations between states has not been specified. All changes between states have the same cost

WEIGHTING – a procedure by which state changes in different characters are assigned costs

Chapter 2: A Survey of Character Construction Methods

2.1 Introduction

Hawkins *et al*. (1997) demonstrated that there are alternative methods in common usage for coding the same data, based on different interpretations of morphological characters. Hawkins (2000) went on to illustrate, using a survey of 34 botanical cladistic analyses, that different authors working on similar data use different character construction methods. Hawkins (2000) defined a number of methods of character coding that she described as unconventional (see below), and counted the occurrence of these. Her findings agreed with those of Pleijel (1995), who concluded that different character coding techniques are being used without discussion of the reasons why. Furthermore, within a single case study different coding methods are often employed for similar characters, again without discussion. Hawkins (2000) found that all of the case studies in her survey included at least one unconventionally coded character and that 16% of all characters examined were unconventional. She discovered that the most common forms of unconventional coding were those types relating to inapplicable data, which she interpreted as indicating a prevalence of hierarchically related characters. She concluded that inconsistency in character coding is due either to failure of homology theory or simply the subjective manner of matrix construction.

In this chapter, Hawkins' (2000) method is applied to 100 zoological datasets (50 using neontological data and 50 palaeontological), in an attempt to assess whether or not there are differences in coding methods employed by different authors, between neontological and palaeontological datasets and/or between invertebrate and vertebrate datasets. Hawkins (2000) identified two main types of character coding: conventional and unconventional.

2.2 Conventional Coding

Hawkins (2000) followed Maddison (1993) in using hypothetical tail characters to illustrate her character types, a trend continued here. She based her concept of conventional coding on de Platnick's (1979: 542) concept of the character as "two or more different attributes (character states) found in two or more specimens that, despite their differences can be considered alternated forms of the same thing (the character)" An example of conventional coding of a binary character is:

Tail colour: red (0), blue (1).

If additional taxa with tails of colours other than red or blue were included, extra states of the character would be added to represent these additional colours. Any other attributes describing the morphology of the tail would be coded as separate characters providing that they were not logically linked (see Wilkinson, 1995a) to the tail colour or to each other.

2.3 Unconventional Coding Methods

Below are brief descriptions of the types of unconventional coding recognised in this survey. Most of these types were identified and discussed by Hawkins (2000). Extent, repetition, behavioural, developmental and landmark coding methods were added here.

2.3.1 Composite Coding

Composite coding is a method of coding that represents multiple attributes of a complex character. It unites attributes that could conceptually be coded as separate characters, effectively assuming the attributes are linked, biologically or logically (Wilkinson, 1995a), so that separate coding of these attributes would lead to overweighting. Hawkins gives the example of red and blue tails and twisted and straight tails. Whereas conventionally these two attributes (colour and shape) would be coded as separate characters, composite coding produces the following single character in which red tails are always twisted and blue tails always straight:

Tails: red and twisted (0), blue and straight (1).

Such coding is satisfactory if there is reason to believe that the characters are genuinely linked, such as consistent covariation among taxa (see Harris *et al*., 2003a and Chapter 3). However, in such cases authors should specify their reasons for employing a composite coding. A problem arises when a form is identified with, for example, a straight, red tail, which would falsify the coding.

2.3.2 Conjunction Coding

Conjunction coding consists of a number of states describing attributes of a character, and a further state that describes specimens that exhibit more than one of these attributes. For example:

Tail colour: red (0), blue (1), red and blue (2).

Although in this case it is possible to imagine taxa with blue and red colouration on the same tail, usually the existence of an individual exhibiting more than one state of a

character suggests that those states are not homologous, since it is a direct violation of Patterson's (1982) conjunction criterion of primary homology assessment. In some cases, such as heterozygosity, however, it may be that failure of the conjunction test does not mean non-homology.

2.3.3 Inapplicable Data Coding

A major problem in character construction revolves around characters that describe attributes that are applicable only to a subset of taxa in an analysis. Such characters often occur due to the hierarchical nature of organisms, for example where some taxa lack a structure that exhibits phylogenetically informative variability. It is desirable for the character construction used to include both the distinction between taxa with and without the structure, and any information presented by attributes of the structure in taxa that possess it. Detailed discussion of the best available treatment for inapplicable characters has been played out in the literature (see discussion) and, unfortunately, it appears that all current methods for coding such characters have flaws. Hawkins included inapplicable data as one of the unconventional coding types in her survey, and she noted two types. The first is a reductive method in which one character in the data codes for the presence or absence of a structure, while further characters code for attributes of this structure. For example:

Tail: absent (0), present (1).

Tail colour: red (0), blue (1).

The character relating to tail colour is inapplicable to taxa that have no tail and therefore, these taxa are scored as unknown (?) for the tail colour character. Hawkins named this type 'inapplicable missing'.

The second type, named 'inapplicable multistate' coding by Hawkins, employs a single composite character in which absence of the structure is coded as one state, and attributes of the structure in taxa possessing it as further states. For example:

Tail: absent (0), present, red (1), present, blue (2).

Hawkins also noted that inapplicable data characters need not always contain a state 'absent'. She gave the following example of what she named 'cryptic' inapplicable data:

Tail feathers: one (0), two (1).

Tail feathers: united (0), separated (2).

The second of these characters would be inapplicable to taxa with one tail feather.

2.3.4 Logically Related Coding

Two characters are completely logically related if the knowledge that a taxon possesses one state for one character imposes restrictions on the possible states it could possess for the second (Wilkinson, 1995a). Characters can also be incompletely logically related if in a taxon the possession of a certain state for one character increases the probability that the taxon possesses a certain state for the second character. Such logical linkage of characters can lead to the overweighting of some attributes. For example, the character:

Wingspan: small (0), large (1).

is logically related to a second character:

Tail length: less than or approximately equal to wingspan (0), greater than wingspan (1).

In this example, if it is assumed that the length of the tail is independent of the wingspan, a taxon coded as having a small wingspan is more likely to be coded as state 1 for the tail length character than a taxon with a large wingspan. This is because a taxon with a large wingspan would have to have a relatively longer tail for the tail to be longer than the wingspan. These two characters are therefore incompletely logically related.

2.3.5 Nominal Variable Coding

Nominal variable (Pimentel and Riggins, 1987), or presence/absence (Pleijel, 1995), coding is a method by which each attribute of a character is constructed as a binary character with the states being the absence or presence of the attribute. For example, if all taxa in the analysis possessed either red or blue tails, a nominal variable coding of tail colour would be as follows:

Red tail colour: present (0), absent (1).

Blue tail colour: present (0), absent (1).

Any additional taxa possessing tails of other colours would necessitate the addition of further characters to code these colours. As discussed by Hawkins, nominal variable coding (NVC) is problematic when thought of in terms of homology assessment. In the above example, the first character in isolation simply tells us that taxa with red tails possess the same tail colour, but does not tell us what the complement relation of red tails is. This character alone is therefore an unspecified homologue character (see below). It is only when two or more nominal variable characters describing the same attribute are present that the homologue of red is revealed. NVC can have analytical consequences. Pairs of homologous nominal variable characters, such as those in the example above are

completely logically linked. So, if a taxon is scored as present for the red tail character, then it must be scored as absent for the blue tail character and vice versa. Both characters show covarying complementary character state distributions. This distribution is the same as the single character that would be created by a conventional coding of the same data. Therefore, providing all characters in a matrix are binary, a nominal variably coded matrix should give identical results in parsimony analysis to conventional coding, although the number of characters and the tree length would be twice as large, because there was two identical characters for each character in the conventional matrix. Nominal variable characters should generally not be used in the same matrix as conventional characters, because a binary character coded using NVC carries twice the weight of the conventionally coded version. Further problems arise if NVC is applied to characters that would contain more than two states if coded conventionally. In NVC, the relative weight of the character states decreases as more states are added. For example, a three-state character has a minimum length of two steps when coded conventionally, which is twice that of a binary character. A three-state nominal variable character would have a minimum length of three steps, which is only 1.5 times the weight of a nominal variably coded binary character. In general, the minimum number of steps of a conventionally coded character is equal to the number of states-1, whereas in NVC it is simply equal to the number of states. This means that the relative weight assigned to characters containing different numbers of states is not the same in the two coding methods. Therefore, a parsimony analysis of a conventionally coded matrix containing multistate and binary characters would not necessarily give the same result as an analysis of a nominal variable version of the same characters.

2.3.6 Positional Coding

Positional coding describes characters that can be present in more than one topological position in different taxa. For example:

Red colour: on wing (0), on tail (1).

Positional coding is problematic in primary homology assessment, as it appears to contravene Patterson's (1982) similarity criterion. This criterion states that homologues must show compositional, ontogenetic and topological correspondence. Obviously, the states of characters coding the different position of a homologous structure do not share topological correspondence. However, this is not necessarily a problem. There are two types of positional coding, (1) those coding alternate positions of a single structure in different taxa, and (2) those coding the presence of iteratively homologous structures on

different parts of a given individual. The first of these is rarely problematic, because the position of the structure is usually constant relative to other structures in the organism even though its absolute position differs between taxa. This is still topological correspondence. If, however, the position of the structure varies relative to other structures, then the topological element of the similarity criterion has been not been met, and the similarity upon which homology is being hypothesised must be of another kind. In the second case the similarity criterion is more difficult to apply. Iterative homology is discussed in depth in Chapter 3.

Hawkins divided positional characters into nominal and composite versions, but here all types of positional coding were counted together and 'nominal' positional characters were also counted as nominal variable characters.

2.3.7 Ratio Coding

Ratio coding is simply the use of ratios to compare attributes of two structures, or two attributes of a single structure (single-structure ratio coding). The latter is generally used as a measure of the shape of a structure. An example of a ratio coded character is:

Tail length: less than or approximately equal to wingspan (0), greater than wingspan (1).

Often ratio coding is applied in an attempt to discretize states in continuous characters rather than simply using an imprecise character, such as tail length, with the states long or short. One obvious problem with ratio coding is that it is a form of composite coding. Either of the two attributes could vary to produce the same result. So, the relative length of the tail to the wingspan could be altered by changes in the length of the tail or the wing. Therefore, taxa may be coded as possessing the same state for the character when the processes producing the similarity are not homologous. This is an extra form of homoplasy that is not suffered by most characters.

Ratio coding also suffers from the problems of describing continuous data, such as where to establish the division between states (e.g. Archie, 1985; Stevens, 1991; Garland Jr. *et al*., 1992; Thiele, 1993; Rae, 1998; Wiens, 2001). Preferably, state delimitation should be identified on the basis of a multi-modal distribution within the variable. However, often, continuous characters form uniform or unimodal distributions, so this is not possible. Therefore, in most cases, the states of ratio coded characters are arbitrary. If ratio characters are still included in these circumstances, authors should discuss their reasons for choices of state delimitation.

2.3.8 Unifying Coding

Unifying coding involves coding two states, which could be coded separately, as a single state in a character. For example:

Tail colour: red or blue (0), purple (1).

Lumping character states together in this way is problematic for homology assessment, because it appears to contravene the compositional correspondence necessitated by the similarity criterion (Patterson, 1982). The coding suggests that red and blue tails are evolutionarily equivalent or that red and blue colourations of the tail are more homologous to each other than either is to purple. If a unifying coding is employed, an evolutionary hypothesis must be proposed to this effect, and should be discussed in the character description. If no justifiable evolutionary hypothesis is proposed, using such a coding method can provide support for grouping taxa that possess different states for the attribute. In the example above, taxa with red and blue tails would be grouped together and contrasted with purple-tailed forms.

2.3.9 Unspecified Homologue Coding

Unspecified homologue coding involves the production of a character that specifies one or more of its states, but includes an alternative to these that is unspecified. It is similar to including only one of a set of nominal variable characters. An unspecified homologue coding for tail colour might be as follows:

Tail colour: red (0), otherwise (1).

This character would not be altered by the addition of taxa with tails that are of a colour other than red or blue, leading to the lumping together of all taxa without red tails.

Hawkins further divided unspecified homologue characters into those including a state ‘otherwise’ or ‘not as (0)’, and those she described as cryptic, in which the states are presence/absence or yes/no, and the attribute is described in the character definition, as in nominal variable coding. For example:

Red tail colour: present (0), absent (1).

In this study, the distinction between types of unspecified homologue codings was considered unnecessary and all forms were counted together.

A number of additional types of unconventional character coding were added during the course of this survey, when characters were encountered that, although not conventionally coded, did not fit well into any of the unconventional coding categories described by Hawkins.

2.3.10 Extent Coding

Extent coding is similar to both composite and ratio coding. It describes those characters that express the size or shape of a structure by its extent relative to another structure. An example of extent coding might be:

Maxilla: extends past anterior margin of orbit (0), does not reach orbit (1).

A problem with this type of coding is that in some cases two taxa may be scored as possessing the same state when it has arisen in a different way. Like ratio coding, this is an unusual form of homoplasy caused by the character construction method that means that the same state can be acheived either by a change in the length of the maxilla or in the position of the orbit. It is plausible that either the maxilla could elongate or move backwards to pass the anterior orbital margin, or the orbit could move forward so that its anterior margin passes the posterior border of the maxilla. Extent coding is similar to both the ratio and positional coding types defined by Hawkins (see above). However, the prevalence of extent coded characters in the papers studied in the survey led to it being allocated its own category.

2.3.11 Repetition Coding

Repetition coding describes instances where the same character, presumably inadvertently, is entered twice in the same data matrix, although the wording may not necessarily be the same. Any such occurrences of repetition are almost certainly simple oversights by the author rather than coding strategies, but they still have an effect on analytical results, so were recorded in the survey.

A second form of repetition coding is the repetition of a character state in two or more characters, for example:

Tail: absent (0), present (1).

Tail: red (0), blue (1), absent (2).

or

Tail colour: red (0), blue (1).

Tail colour: red (0), purple (1).

The first example involves inapplicable data, but also includes presence/absence as an individual character and uses a multistate character including absence to code attributes of the tail. This means that absence of a tail is afforded twice the weight of other characters. The second example also repeats the state ‘red’, but here this state has two different

homologues in different characters. Such codings are incomprehensible and indefensible, but have been identified in at least one of the studies examined in this survey. Repetition coding is obviously a form of logical linkage, but was counted separately in this survey.

2.3.12 Behavioural, Developmental and Landmark Coding

A number of the neontological studies examined in this survey incorporated behavioural or developmental characters in their data. Similarly, one neontological study used characters and states based on morphometric landmarks. These characters were placed into their own groups to highlight the application of these alternative types of character data or construction. However, if more of these types of characters had been present in the surveyed papers these characters could have been treated differently.

2.3.13 Mixed Coding

Character codings can fit into more than one of the non-conventional groups outlined above. In such cases, these characters are included in the count of each separate type of inapplicable coding that they possess. So, if a character is inapplicable (multistate) and is based on ratios, then it will count as both of these types, so the total number of both inapplicable multistate and ratio characters in that dataset will be incremented by one.

2.4 Methods

The character sets examined in this survey were taken from the papers used by Harcourt-Brown *et al*. (2001) in their study of tree balance. Harcourt-Brown *et al*. carried out a survey of 50 neontological analyses and 50 palaeontological analyses in an attempt to see if and why palaeontological trees tend to be more imbalanced than neontological ones. They chose papers for their survey by examining selected journals in reverse chronological order to identify analyses that met certain criteria (see Harcourt-Brown *et al*., 2001 for complete sampling strategy). In the current survey, the same 100 papers were examined (for references see Appendix 1). However, three of the neontological studies (Eernisse *et al*., 1992; Vane-Wright *et al*., 1992; Wheeler *et al*., 1993) used by Harcourt-Brown *et al*. (2001) had to be excluded from this survey, because they lacked descriptions of characters and character states, and one further reference (Mikkelsen, 1996) could not be obtained. Therefore, only 46 neontological studies were included in the survey. A total of 6360 characters were examined (3289 from neontological studies, and 3071 from palaeontological studies). For each paper, all character codings were scrutinised and any

characters adjudged unconventional were placed into one or more of the coding categories described above. For each paper, the number of characters falling into each coding category was then counted. Because characters can fall into more than one category, the total number of unconventional character codings found in a dataset can exceed the total number of characters in that dataset. However, the number of characters falling into each individual category of unconventional coding must be fewer than or equal to the total number of characters in the dataset.

Comparisons were made between coding methods employed by authors in the fields of neontology and palaeontology and between the 73 vertebrate and 23 invertebrate studies included in the survey. Mann-Whitney U-tests were applied individually to each coding type to assess whether any differences identified were statistically significant. The first step in a Mann-Whitney U-tests is to rank (in an increasing or decreasing order of magnitude) all of the papers together based on the number of characters constructed with the particular coding type being tested. Ties are each assigned the mean of the ranks that they jointly occupy. Next, the data are split into the two groups of interest (for example neontological data versus palaeontological data) and the sum of the ranks of each group calculated. The Mann-Whitney test is a non-parametric test of the null hypothesis that the two samples come from identical populations. This is equivalent to the hypothesis that there is no difference in the distribution of ranks between the two groups. i.e. that the average ranks of the palaeontological and neontological data are not significantly different. If one or other of the types of data has a significantly higher average rank, then the coding type under investigation is significantly more prevalent in that type of data.

2.5 Results

Results of the survey are shown in Appendix 2, summarised in table 2.1 and graphically illustrated in figures 2.1 and 2.2. All papers included in the survey contained unconventionally coded characters, which made up 29% of the 6356 characters examined.

Coding Type	Neontological		Palaeontological		Total	
	Matrices (%)	Characters	Matrices (%)	Characters	Matrices (%)	Characters
Total Number of Characters		3285		3071		6356
Behavioural*	17.4	25	-	-	8.7	25
Composite	52.2	78	66.0	84	59.1	162
Conjunction	13.0	7	2.0	2	7.5	9
Developmental*	2.2	1	-	-	1.1	1
Extent	56.5	74	60.0	72	58.3	146
Inapplicable Data (missing)	63.0	90	40.0	69	51.5	159
Inapplicable Data (multistate)	89.1	291	68.0	226	78.6	517
Landmark	2.2	9	0	0	1.1	9
Logically Related	19.6	20	24.0	17	21.8	37
Nominal Variable	2.2	16	4.0	6	3.1	22
Positional	73.9	139	70.0	135	72.0	274
Ratio	65.2	111	88.0	211	76.6	322
Repetition	4.3	3	0.0	0	2.2	3
Single Structure Ratio	32.6	23	52.0	56	42.3	79
Unifying	63.0	133	74.0	119	68.5	252
Unspecified Homologue	15.2	51	14.0	37	14.6	88
Total Unconventional		**1071**		**1034**		**2105**

Table 2.1. Occurrences of unconventional characters in the survey of 96 cladistic analyses (46 neontological, 50 palaeontological). Occurrences quantified as the percentage of matrices containing at least one character of a coding type, and the total number of each type of character coding found in neontological studies, palaeontological studies and in total. The total number of characters examined is also reported. Asterisks represent those characters that were only found in neontological datasets, as they are unlikely to occur present in palaeontological data.

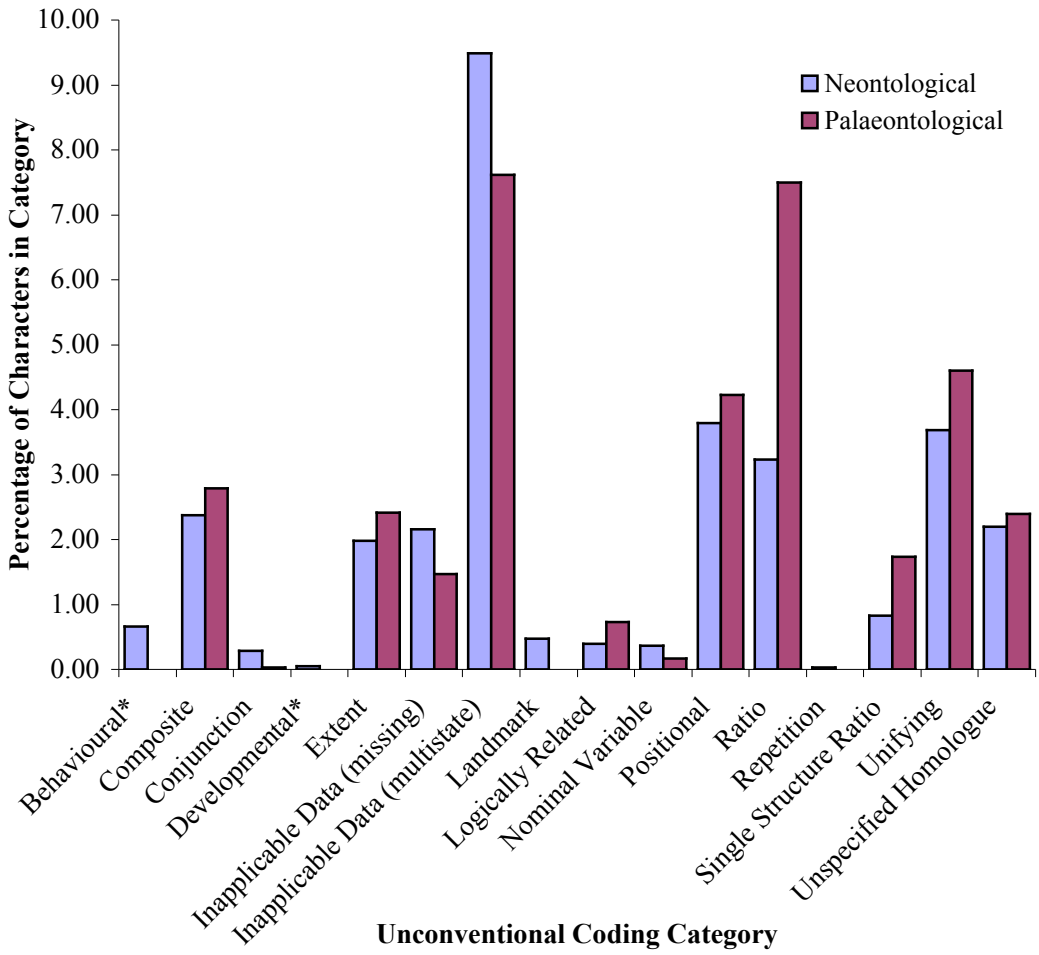


Figure 2.1. Histogram showing the average percentage of characters per study of each coding type. Red bars represent neontological studies and blue bars represent palaeontological studies. Asterisks represent character types that are unlikely to be present in palaeontological datasets.

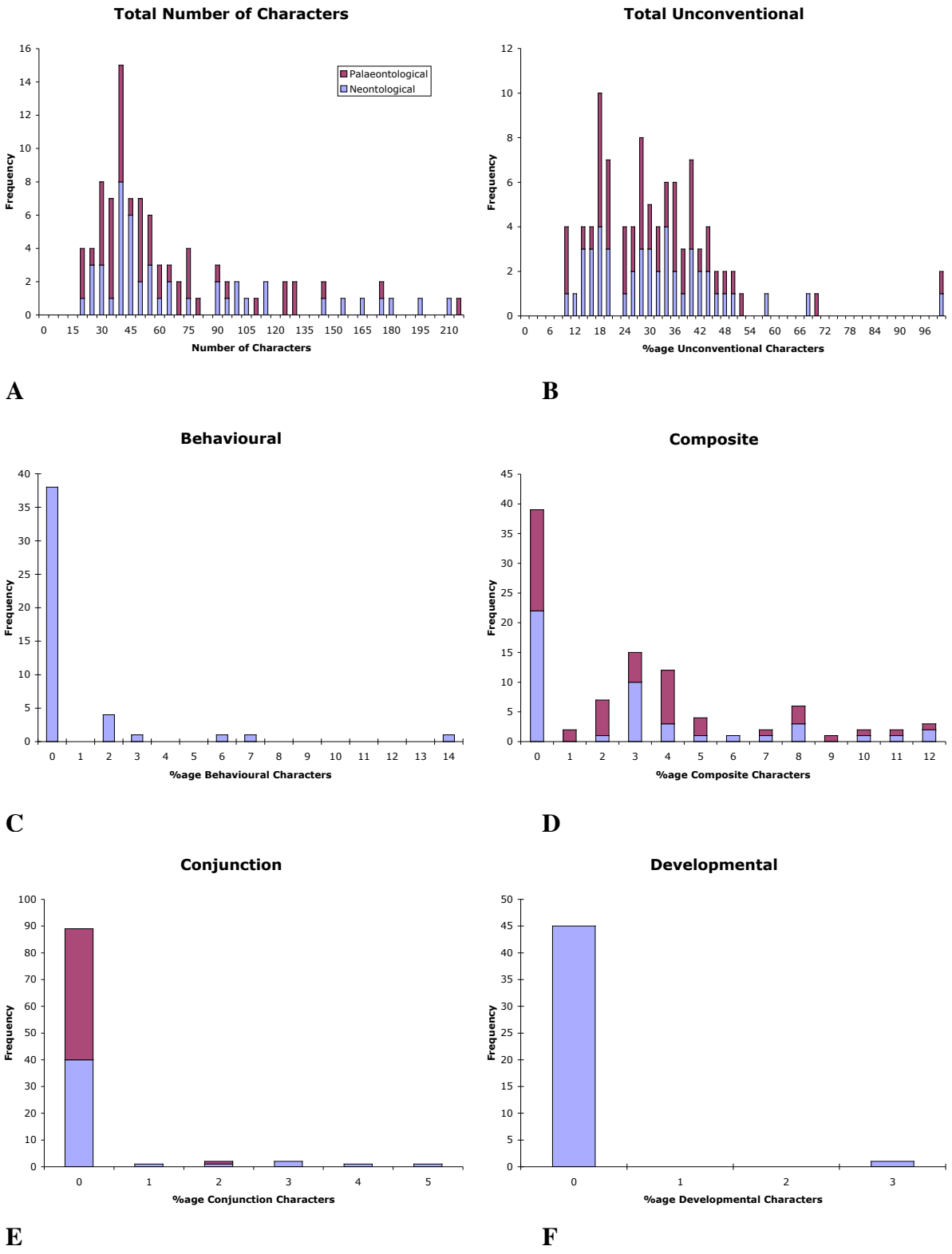
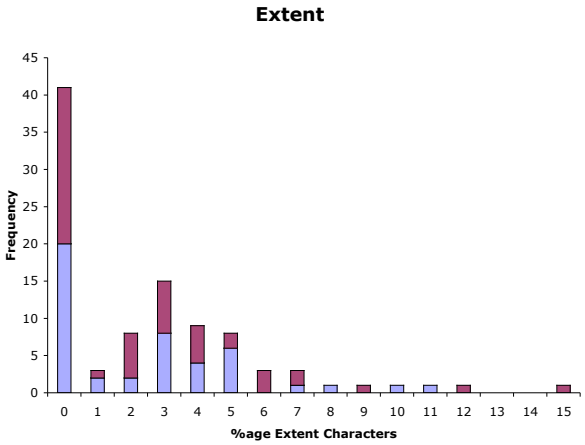
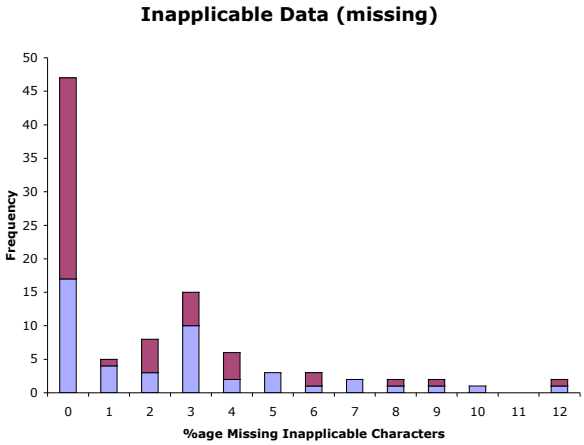


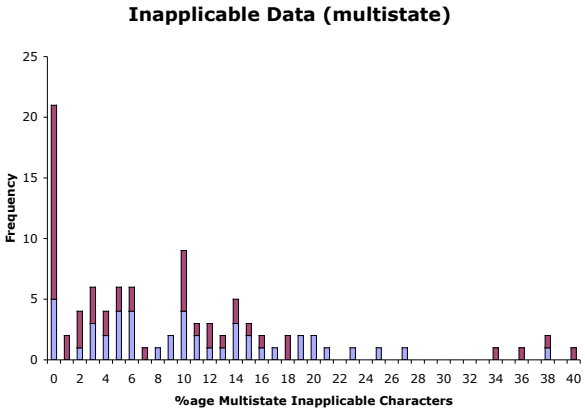
Figure 2.2. Frequency histograms showing the occurrence of each type of unconventional coding. (A) Number of characters present in studies. (B) Proportion of characters in studies that were unconventional. (C to R) Proportion of characters in studies that were of each type of unconventional coding.


Extent
Frequency
%age Extent Characters

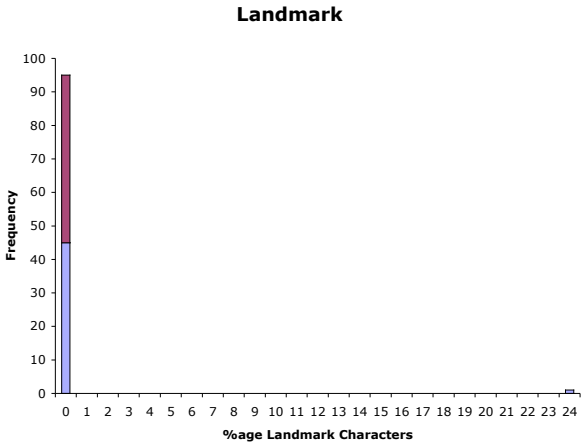
G


Inapplicable Data (missing)
Frequency
%age Missing Inapplicable Characters

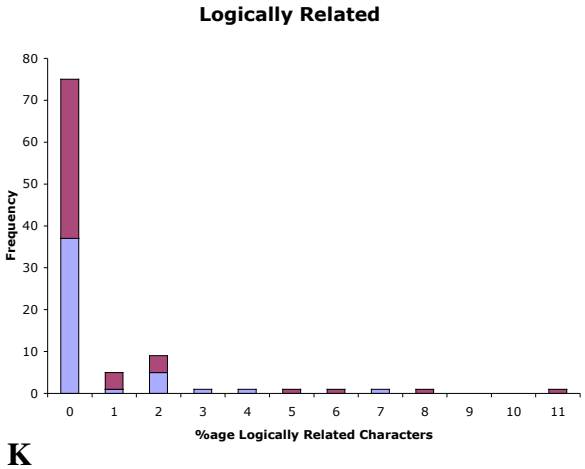
H


Inapplicable Data (multistate)
Frequency
%age Multistate Inapplicable Characters

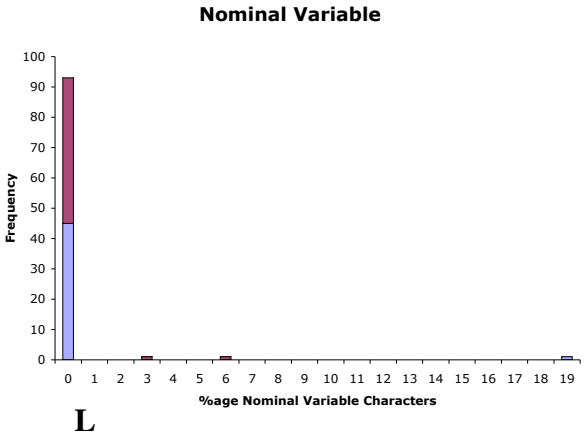
I


Landmark
Frequency
%age Landmark Characters

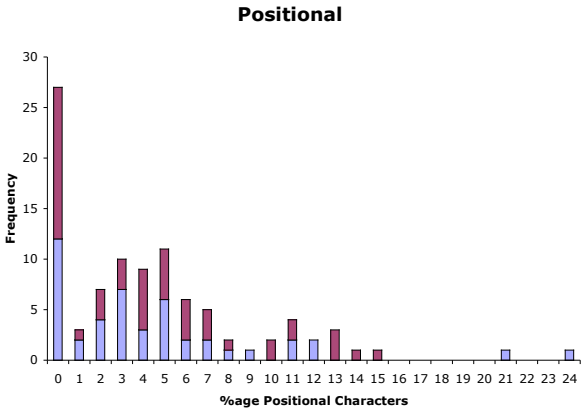
J


Logically Related
Frequency
%age Logically Related Characters

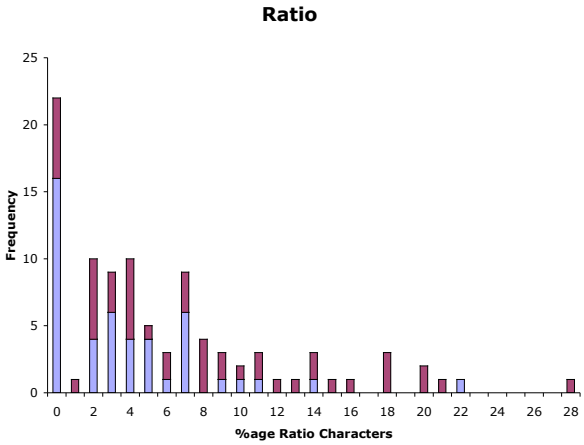
K


Nominal Variable
Frequency
%age Nominal Variable Characters

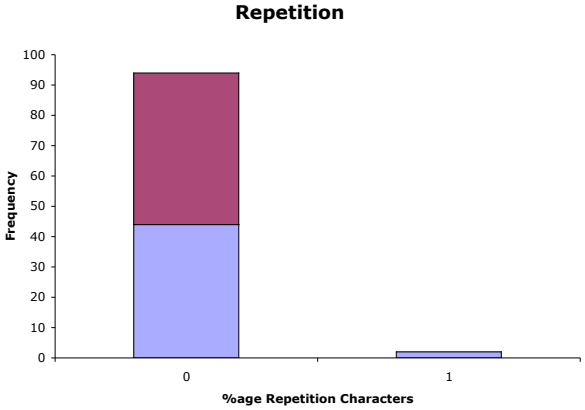
L


Positional
Frequency
%age Positional Characters

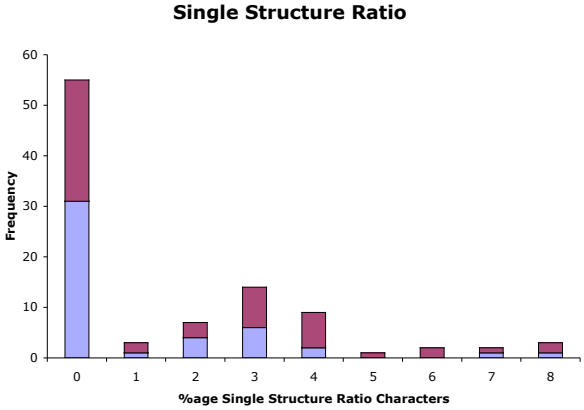
M


Ratio
Frequency
%age Ratio Characters

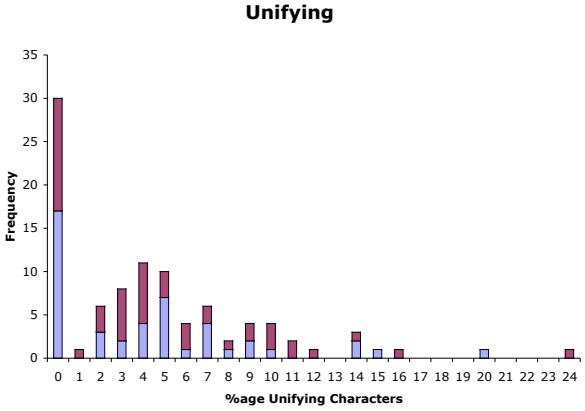
N


Repetition
Frequency
%age Repetition Characters

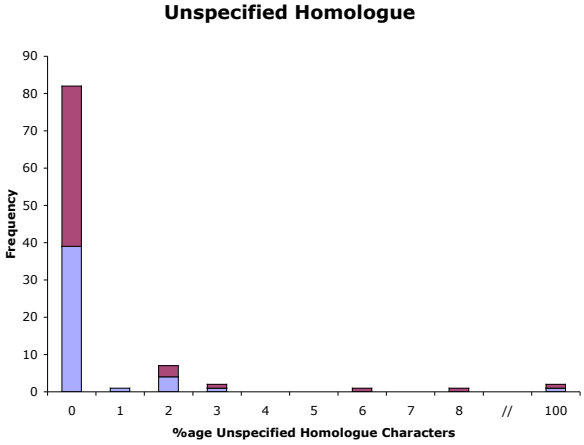
O


Single Structure Ratio
Frequency
%age Single Structure Ratio Characters

P


Unifying
Frequency
%age Unifying Characters

Q


Unspecified Homologue
Frequency
%age Unspecified Homologue Characters

R

2.5.1 Neontological Versus Palaeontological Analyses

It is clear that in neontological and palaeontological papers the relative proportions of different types of unconventionally coded characters are generally similar (see Fig. 2.1). However, there are some differences. Mann-Whitney U-tests (see Table 2.2 for results for all coding types) show that significantly more ratio (p=0.0006) and single structure ratio (p=0.0222) characters are utilised in palaeontological analyses than those including only neontological data. Significantly more inapplicable data (ID multistate p=0.0076, ID missing p=0.0271) character codings were used in neontological analyses. When vertebrate studies alone were subjected to Mann-Whitney U tests, in order to remove any effect of differences caused by the large number of invertebrate studies in the neontological surveys relative to the palaeontological surveys, the significant results for ratio, single structure ratio and inapplicable data (multistate) characters are still recovered.

Coding Type	**Neo. vs palaeo.**		**Invert. vs vert.**	
Behavioural	-	-	0.3539	0.6551
Composite	0.3077	0.2089	0.2529	0.3169
Conjunction	0.8194	0.8280	0.6982	0.7188
Extent	0.5059	0.4883	**0.0021***	**0.0070***
Inapplicable Data (missing)	**0.9589***	**0.9729***	0.8996	**0.9684***
Inapplicable Data (multistate)	**0.9945***	**0.9924***	0.9414	**0.9930***
Logically Related	0.3569	0.3542	0.2086	0.2111
Nominal Variable	0.4417	0.4417	0.5495	0.5495
Positional	0.5641	0.3064	**0.0237***	**0.0239***
Ratio	**0.0032***	**0.0006***	**0.0019***	**0.0063***
Repetition	0.6431	0.6431	0.4217	0.4217
Single Structure Ratio	**0.0269***	**0.0222***	0.1156	0.1407
Unifying	0.5190	0.2453	0.1485	0.0982
Unspecified Homologue	0.5234	0.5249	0.9307	0.9307
Total Unconventional	0.6832	0.7335	0.0796	0.6922

Table 2.2. Results of Mann Whitney U-Tests of occurrences of each type of unconventional character in neontological vs. palaeontological papers and in invertebrate vs. vertebrate papers. For each, values in the left column represent tests of numbers of occurrences and numbers in the right column represent tests of proportion of occurrence. Significant results of U-tests at the 5% level are show in bold and marked with an asterisk.

2.5.2 Invertebrate Versus Vertebrate Analyses

Mann-Whitney U-tests (Table 2.2) indicate that positional (p=0.0239), extent (p=0.007) and ratio (p=0.0063) (but not single structure ratio (p=0.1407)) characters are significantly more common in studies of vertebrates, whereas invertebrate analyses tend to

employ more inapplicable data characters (ID multistate p=0.003, ID missing p=0.0316). Mann-Whitney U tests were also carried out on the neontological results alone in order to correct for any effect on the results caused by the fact that almost all of the invertebrate studies were neontological. In these tests, the only significant difference was that vertebrate studies used significantly more extent characters than invertebrate studies (p=0.0097). However, testing the neontological studies alone reduces the number of data points in the Mann-Whitney test, which reduces the degrees of freedom and makes it less likely that significance will be achieved.

2.6 Discussion

The most striking result of this study is the abundance of unconventional coding methods used in published analyses. All studies were found to contain unconventional codings, and 29% of all characters examined were unconventional, even higher than the 16% found by Hawkins (2000) in her similar survey of botanical papers. This large difference is difficult to explain because a number of factors may influence the results. First, the results may indeed be a true representation of the differences in character coding methods used by authors in the fields of botany and zoology. However, it is probable that at least some of the difference can be explained by the subjective nature of identifying coding categories. Although the current study was based on the descriptions of unconventional coding types published by Hawkins (2000), those descriptions are often open to alternative interpretations, and decisions as to whether or not characters fall into certain categories are therefore themselves subjective. Some coding types are far easier to identify in the literature than others, because their definitions are more clear-cut. Multistate inapplicable data coding, for example, is relatively easily recognised, because generally the character state descriptions include one state with a single attribute (e.g. tail absent) and further states with more than one attribute (e.g. tail present and red, tail present and blue). The ease of recognition of this category may make the results more comparable between studies than some other categories whose identification is more difficult. Recognition of logically linked characters, for example, often requires detailed knowledge of the morphology of the taxa being analysed, so characters of this type may be more commonly missed. Similarly, nominal variable coding can be difficult to identify, because characters only qualify as nominal variables if at least two unspecified homologue characters can together be conceptualised as a conventional character. This requires identification of

homologous states in separate characters, which can be problematic, especially if the two nominal variable characters are not sequential in a large published character list. Another possible explanation for part of the difference in the proportion of unconventional characters between the two studies is that in the current study five new types were added to the list of unconventional coding strategies (behavioural, developmental, extent, landmark and repetition). The addition of these new coding types means that some characters considered conventional in the study by Hawkins would, in the new study, be considered unconventional by falling into these categories. Together these factors mean it may not be wise to compare the results of surveys such as these, which are carried out by different authors, even though the descriptions of character coding categories used are generally the same. The results of each survey individually, however, do provide interesting data. Because the current survey was carried out by a single person, the risk of inconsistencies in the identification of coding categories is lowered.

Most of the results of this survey were markedly different from those found by Hawkins (2000). For example, Hawkins (2000) found unspecified homologue characters to be the second most common coding type (after inapplicable data characters), whereas in this survey such characters were relatively rare (only 88 (<2%) occurrences in 15% of analyses). Conversely, Hawkins found very few positional characters in her survey, but these characters were one of the most common types identified in the current survey. Again, the reasons for these differences between the surveys could be due to differences between the methods of character construction used by zoologists and botanists, the subjective nature of placing characters into categories, or actual differences between the organisms under study. For example, it may be that plants are more modular in their morphology than animals, which may alter the types of characters constructions used.

The two studies did agree in both finding inapplicable data coding to be the most common form of unconventional coding, and that the majority of inapplicable characters were coded using the multistate method, against common advice and generally without discussion as to why they chose that method. A similar lack of discussion about character inclusion was noted by Poe and Wiens (2000), suggesting that it is a major obstacle to understanding the criteria used in constructing many published phylogenetic analyses.

Both studies also highlighted the problem of inconsistency in coding within and between studies. It is clear from the results presented here (see Appendix 2) that there is a great deal of variation in the coding methods used by different authors. For example, the proportion of unconventionally coded characters in single studies ranged from 3 out of 35

characters (8.6% unconventional) in the analysis of Gower and Sennikov (1997, reference F27, Appendix 1) to 54 out of 56 characters (96.4% unconventional) in the study of Alroy (1995, reference F32, Appendix 1). Similarly, the types of unconventional characters used by different authors also varied widely (see Fig. 2.2). For example, of the 88 unspecified homologue characters coded in all analyses (Fig. 2.2r), 44 were coded by Guisu and Winterbottom (1993, reference E48) and 27 by Hooks (1998, reference F22) with only 17 in all other analyses combined. Similarly, inapplicable data (multistate) characters (Fig. 2.2i) made up between zero percent (in 21 studies) and forty percent (Steppan and Pardinas, 1998, reference F15, Appendix 1) of the total number of characters in different papers. Such variation suggests that different authors are constructing characters based on morphological variation in different ways. More worryingly, many authors do not seem to be consistently coding similar types of characters within single analyses. 42% of analyses in the survey included both methods of coding inapplicable data. This type of inconsistency appeared more common in neontological analyses, where 56% of studies included both types, compared with only 28% of palaeontological analyses. Similarly, three studies included both nominal variable characters and conventionally coded characters. The problems associated with this are discussed in the coding type definitions above. Such a lack of consistency in coding similar characters was also identified by both Pleijel (1995) and Hawkins (2000), and demonstrates unequivocally that character construction is not being performed explicitly enough in many phylogenetic analyses.

2.6.1 Palaeontology Versus Neontology

Another inconsistency that is apparent from the survey is that workers in the different disciplines of the biological sciences do not employ the alternative coding methods equally. It was shown that neontologists use significantly more inapplicable data characters than palaeontologists, but this abundance of inapplicable data characters in neontological datasets is difficult to interpret. One possible explanation may be related to the incomplete nature of a great deal of palaeontological material. Hawkins hypothesised that the large number of inapplicable characters present in analyses represented the high numbers of hierarchically related characters that must be dealt with in systematic analyses. The relative completeness of neontological data means that there are more potential characters available to the neontological phylogeneticist than palaeontological workers. Although the survey found that the total number of characters used in the neontological analyses is not significantly greater than the number used in palaeontological studies, the

presence of significantly more multistate inapplicable data characters indicates that neontologists are recognising more attributes on structures, but are combining hierarchically-related attributes into single characters. As discussed above, this leads to loss of potential characters and phylogenetic information. The dearth of available characters in palaeontological data may make phylogeneticists loath to lose any potential phylogenetic information, and so workers in this field may prefer not to combine different attributes of structures into single characters. However, this explanation would not account for the significant difference in the number of inapplicable 'missing' characters, which do not lead to loss of information in the same way. An alternative explanation for the larger number of inapplicable data characters in neontological analyses may be another consequence of incomplete preservation of fossil material. Since most inapplicable data characters are the result of hierarchically related structures, incomplete preservation may restrict the possible number of such characters. As preservation of fossils becomes poorer, the first characters to be lost are likely to be those describing smaller features, such as processes on bones. Such small characters are often hierarchically-related to larger, more general characters, such as the presence or absence of the bone. Therefore, the loss of these characters will lead to less hierarchical-relatedness in the data. For example, in neontological analyses one character may be the presence or absence of the temporal bone, and further characters may describe attributes of the shape or structure of the temporal bone. In palaeontological data, however, preservation may make some of these characters impossible to recognise. The most likely of the characters to be recognisable is the presence or absence of the temporal (the character highest in the hierarchy).

Palaeontological studies employed more ratio characters than neontological studies. This is again probably a consequence of palaeontological specimens being less numerous and well preserved than can be expected in neontological studies. Palaeontologists have to create data matrices based almost entirely on hard-tissues that have a high preservational potential. One type of character that is easy to identify in hard-tissue structures (especially vertebrate bones) is the relative size of different elements or of different measurements from a single element. Although such characters are also widely used in neontological studies, the wealth of other possible characters (such as soft tissues and colour to name just two), mean that ratio characters, with their associated problems, can be replaced by less controversial characters.

Significantly more extent characters were used in invertebrate than vertebrate analyses. The reasons for this are uncertain.

2.6.2 Inapplicable Data Characters

Looking at the results of the current survey in more detail it is clear that some types of unconventional character coding are employed far more frequently than others. Two of the most common coding types concern the coding of inapplicable data. Multistate inapplicable data characters are the most common, being present in 79% of the studies with 517 occurrences in total, with missing inapplicable coding present in 52% of studies with 159 occurrences.

As discussed by Hawkins (2000), inapplicable data are prevalent in cladistic analyses due to the hierarchical nature of the structure of organisms, and this is one of the most problematic aspects of character construction for phylogenetic analyses (e.g. Maddison, 1993; Hawkins *et al*., 1997; Lee and Bryant, 1999; Strong and Lipscomb, 1999). It is widely accepted that there is currently no flawless method for coding inapplicable data, and that both missing and multistate coding methods present different problems in phylogenetic analysis. Maddison (1993) advocated the use of a multistate method. He provided a simple example that focussed on a hypothetical phylogenetic tree of 14 taxa produced by analysis of a large number of characters. The tree (See Fig. 2.3a) is fully resolved except for one clade on the left, and all relationships are supported by a large number of characters. Maddison (1993) envisaged a situation in which organisms in two clades on the tree possessed tails (those at the far left and right of the tree), with tailless taxa basal to both. Both groups of tailed forms contain taxa with blue tails and taxa with red tails. Maddison (1993) argued that two alternative resolutions (Fig. 2.3b, c) of the left-hand clade should be equally parsimonious, because taxa ancestral to this group are tailless and therefore the primitive tail colour is unknown. He demonstrated, however, that if the missing method of coding inapplicable data was employed (so that the presence or absence of tails was coded as one character and tail colour coded as a second with those taxa lacking tails scored as unknown (?)), the resolution of the group on the right of the tree influenced the most parsimonious resolution of the group on the left (See Fig. 2.3b, c), so that a tree with blue tails basal to red in the left-hand clade is one step shorter than with red basal to blue. This is simply an artefact of scoring taxa lacking tails as unknown. The only evidence on the polarity of the tail colour character comes from the resolved clade of tailed taxa on the right. The basal-most taxon in this clade has a blue tail, so the most parsimonious conclusion is that blue tails are primitive and red derived. Effectively, parsimony assumes that had the tailless taxa had tails they would be blue.

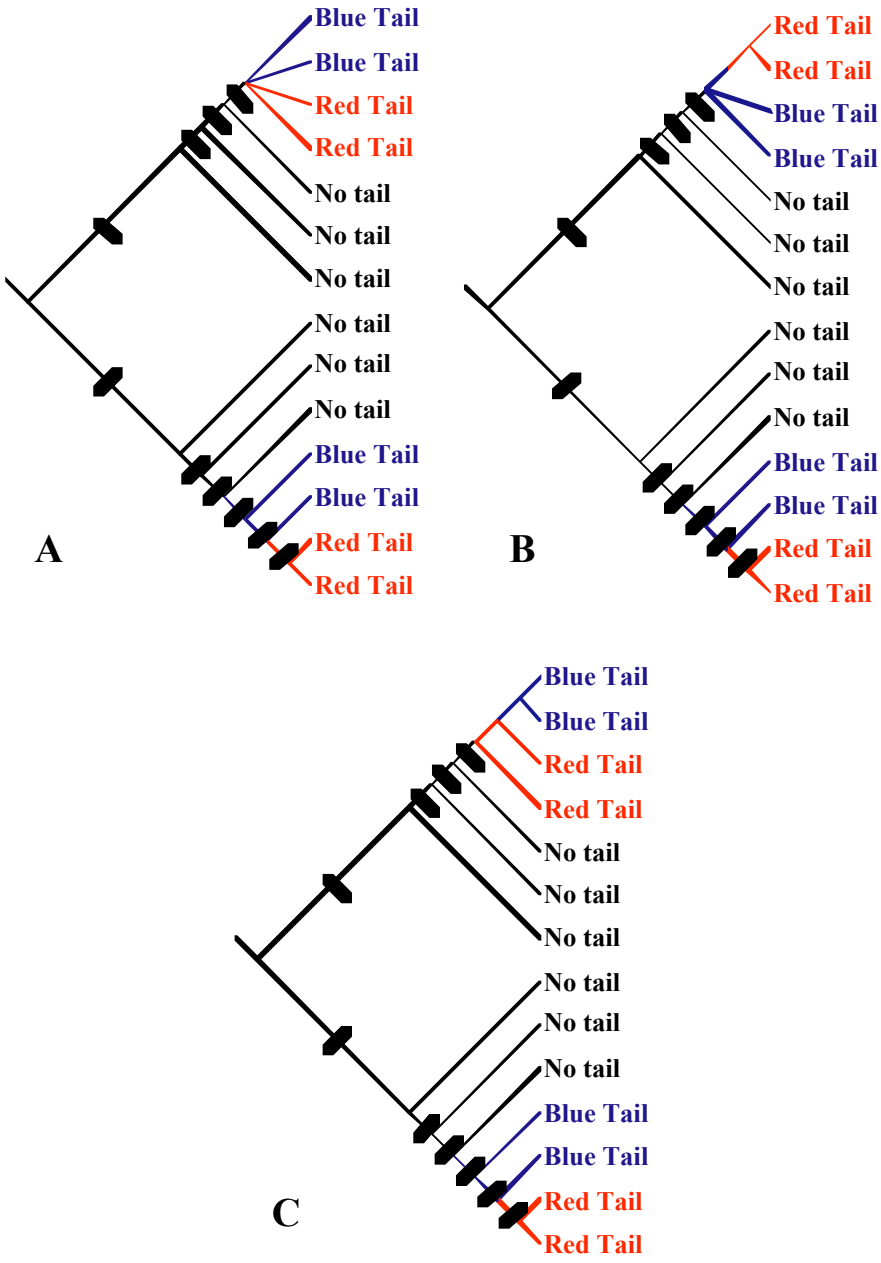


Figure 2.3. Hypothetical phylogenetic tree modified from Maddison's (1993) example. Some taxa lack tails, whilst some have blue tails and some red tails. Bars cutting branches represent nodes supported by a large number of characters. (A) The clade of tailed taxa on the left is unresolved. This would be the expected resolution given the available data. (B, C) Two possible resolutions of the left-hand clade of tailed taxa. Resolution B has a length of 3 steps with the multistate coding method, and 2 steps with the missing method. Tree C has a length of 3 steps with either method.

Therefore, when resolving the clade on the left of the tree it appears more parsimonious for blue-tailed taxa to be primitive to red-tailed taxa. This, of course, would not be a problem if the evolution of tails always went through a blue stage before changing to red. Maddison (1993) concluded that it is preferable to code a single multistate character for tail attributes with three states: no tails, blue tails and red tails. It is true that by coding tail variation in

this way the problem he identified is removed. The tailless taxa between the two groups of tailed taxa are coded with their own state, so the basal blue tails on the right cannot influence taxa on the left and either resolution of the left-hand tailed clade is equally parsimonious (Fig. 2.3b, c).

Maddison's (1993) assessment has been criticised by Hawkins *et al*. (1997) and Lee and Bryant (1999), who argued that, although fusing the presence and colour of tails into one character accounts for their logical hierarchical linkage, such coding is equivocal regarding transformational independence, and leads to a loss of phylogenetic information. Combining the two characters means that there is no character state that is a synapomorphy of 'presence of tails'. It is effectively assumed that no tails, red tails and blue tails are homologous states of the character 'tail anatomy', and that any transformation between these states is equally likely. A transformation series from blue tails to no tails to red tails would be as parsimonious as to change from no tails to blue tails to red tails. This is difficult to justify. Patterson (1982) showed that absence of a structure has no logical relationships with any character variables and has a complement relationship only with presence of the structure. If a multistate coding of Maddison's character is analysed unordered (as such characters usually are), it simply splits the taxa into three groups, those with no tails, those with red tails and those with blue tails, and ignores the hierarchical nature of the data. i.e. that taxa firstly either possess or lack tails, and secondly, that the tails of those that possess them can be red or blue. In other words, multistate coding does not preserve the hierarchical nature of hierarchical data, and leads to the loss of potentially useful phylogenetic information identified during morphology assessment. Strong and Lipscomb (1999) also criticised Maddison's preference for multistate coding. They showed that Maddison's example was misleading, because multistate coding only recovers the correct tree when absence is primitive and there are no secondary losses. They claimed, however, that the method is not implemented satisfactorily in most computer programs. They advocate the use of NONA (Goloboff, 1998), because it does not allow the recovery of semi-strictly supported trees. Other parsimony programs, such as PAUP (Swofford, 1999) and Hennig86 (Farris, 1988) often recover more MPTs than NONA, because they allow semi-strictly supported trees, which can lead to inapplicable data determining the placement of taxa it should not affect in some of the MPTs. However, as long as a strict consensus of resulting trees is calculated, the results returned by all of the programs are the same.

Further problems arise with multistate inapplicable data coding if more than one attributes of tails is identified as being phylogenetically informative. Using the inapplicable missing method, further characters can simply be added for each attribute, with taxa lacking tails again scored as unknown for the new characters. Using multistate coding, as highlighted by Maddison (1993) himself, extra states must be added to the single multistate tail character, which can lead to the production of extremely complex characters. Again these characters lose potentially useful information about which states are homologous to one another.

Maddison's (1993) example is also problematic when thought of in terms of more broad homology assessment. He states that the topology of the tree shown is supported by a large number of characters. In the tree, two clades that are widely separated possess tails, so that the character tails present/absent fails Patterson's (1982) test of secondary homology and it seems likely that tails evolved twice independently. So, the two groups may possess tails, but a large number of characters suggest that these tails are not homologous and hence the problem with the coding of attributes of tails may be immaterial.

In the survey carried out here it was found that inapplicable data characters were prevalent, especially in neontological studies. This suggests that the inapplicable data problem is common and widespread and, therefore, that far more effort should be invested into attempting to solve it. Here it was also found that the multistate method for coding such characters was employed more frequently than the missing method, which suggests that in many cases, such data may not be being treated in the best way currently available.

The second most common type of unconventional coding found in the survey was ratio coding, which occurred 322 times in 77% of the matrices. Additionally, 79 single structure ratio characters were identified, and 145 extent characters. Both of these types are very similar to ratio characters, making such coding rather commonplace in the survey. Other common coding types were positional (275 occurrences), unifying (252), and composite (162) coding. Each of these coding types has associated problems, which are discussed in their descriptions above. All other coding methods together occurred only 194 times.

Conclusions

Hawkins (2000) concluded that her finding of 16% unconventional characters showed that "either guidelines are ignored or they are not interpreted in the same way" (Hawkins 2000:34). Such a conclusion is generally corroborated here, especially because of the large differences in coding practice observed between different authors, but I disagree that guidelines are being ignored, since there appear to be few definitive guidelines to follow. However, one problem with Hawkins' conclusion is that it assumes that the use of unconventional characters is in itself more problematic than conventional coding. In fact, if unconventionally coded characters were being used in the literature in a uniform way, so that all authors were using similar types of coding for similar types of morphological variation, it would not be a cause for less practical concern, because at least the subjective differences would be removed. Problems are likely to arise if, as found here, there is a great deal of variation in the types of coding used by different individual authors, or by workers in different fields of biology, without explanation of their reasons for employing such strategies. For example, how can two or more competing phylogenetic hypotheses be evaluated against each other if they are based on unjustified, unconventionally coded characters? Worse still, it has also been demonstrated here that many authors are not using consistent coding practices in a single analysis. Hawkins (2000) also includes all of the unconventional coding types together as problematic, which ignores the fact that some types are more worrying than others. Some unconventional coding types can be preferable to conventional coding in certain circumstances, and in some cases, conventional characters cannot be applied. For example, the plethora of inapplicable data characters found in both surveys highlights, as stated by Hawkins (2000), the great deal of hierarchically-related variation present in organisms. There is no way to code such variation using conventional characters other than by simply excluding those characters that are linked. The preponderance of hierarchical relatedness between characters means that such an approach would lead to the loss of a large amount of potentially phylogenetically informative data. The only alternative is to implement one of the unconventional inapplicable data coding methods.

Here it is considered that the main problem in character construction is not that guidelines are ignored or misinterpreted, but that there are very few unquestionable guidelines. Character construction is a subjective process, and it is extremely difficult to draw up general rules, although some authors have made an attempt (e.g. Maddison, 1993;

Pleijel, 1995; Wilkinson, 1995a; Harris *et al.*, 2003a). Far more work in this area is needed, and a unified theory of character construction striven for if the process is to be improved. However, another problem is that where guidelines have been suggested, such as the use of nominal variable coding (e.g. Pleijel, 1995) methods, or methods for the treatment of inapplicable data (e.g. Maddison, 1993), there is often so much criticism or disagreement that it is not surprising that no one method is consistently applied. Far more work needs to be carried out to attempt to solve some of the major problems in character construction. The treatment of inapplicable data, for instance, can have major effects on the outcome of analyses. In this survey, it was shown that such characters are extremely common, and that treatment was not consistent within, let alone between, studies. Without a solution to the debate over the treatment inapplicable data, this problem will persist.

There are some simple guidelines that should always be followed when constructing characters that would improve the output of phylogenetic analyses:

- Character construction is the most important stage in recovery of a phylogenetic hypothesis, and for that reason assessment of primary homology should be carried out fastidiously.
- Authors should discuss reasons for employing a certain coding strategy. For example, when using inapplicable data coding, authors should explain why they chose to use multistate or missing coding.
- Some types of unconventional character should certainly be avoided, including repetition, logical linkage and nominal variable coding.
- Authors should discuss their choices of state delimitations whenever they are not clearly defined in the character descriptions.
- Importantly, all workers performing cladistic analyses should strive for consistency within their data, so that similar morphological variations are all coded in the same way. If no best treatment for a character type is known, authors could compare results of analyses using alternative coding types for those characters. If the result is consistent whichever method is used, then it could be concluded that that phylogenetic hypothesis is well supported, whereas if differences occur, further investigation will be necessary.
- Finally, where discussion of the best method for coding a certain type of character has been published (for example, when coding inapplicable data or data relating to continuous variables), authors should consult the literature before attempting to code

their data. If all analyses were carried out following these simple guidelines, it would be possible to place more faith in published phylogenies.

- Until more is known, perhaps authors should be extra cautious about their interpretation of their trees. Confidence should only be placed in clades that are very well supported.

Chapter 3: Character Construction and the Phylogeny of Aetosaurian Archosaurs (Reptilia, Diapsida)

3.1 Introduction

Many phylogeneticists seemingly have their own intuitive approach to character construction rather than making explicit choices among the available alternatives of which they may be only dimly aware (see Chapter 2). The practicing phylogeneticist is most likely to be keenly aware of alternative approaches upon discovering that they would (or do) do things differently to other workers. Such a discovery was the stimulus to Wilkinson's (1995a) discussion of reductive and composite coding approaches to the treatment of complexity. A parallel discovery made during investigations of aetosaurian phylogeny prompts us to highlight and discuss here alternative approaches to the construction of characters from anatomical systems comprising multiple parts that are themselves homologous *within* organisms.

This chapter is based on a paper (Harris *et al*., 2003a) written during my PhD studies, which was written in conjunction with Dr. Mark Wilkinson and Dr. David Gower of the Natural History Museum, London. The review of aetosaur phylogenetic analyses was carried out by myself, while the discussion of character construction, especially that of intraorganismal homology, was produced by a combination of myself and my co-authors.

As noted by Ghiselin (1976: 134) "It is a brute fact of nature that lots of organisms are built up of repeated units having similar, if not identical, arrangements of their components." Owen (1843) coined the term *serial* homology for corresponding anatomical units in different segments within organisms, such as vertebrae or the humerus and femur, in contrast to *special* homology, which pertains to correspondences among organisms, including those of different species. Ghiselin (1976) took serial homology to apply only to features that occur in a linear spatial arrangement within an organism, and he noted the existence of many other kinds of correspondences within organisms. For example, his *antimeric* homology pertains to the correspondence between bilaterally paired structures. Whereas the interpretation of special homology appears to have, for the most part, become evolutionary (see Chapter 1.3), intraorganismal homology remains a poorly understood but seemingly fundamental aspect of organismal organization (Ghiselin, 1976).

Wilkinson (1995a) distinguished between two approaches that have been used to construct characters from interorganismal variation in complex features, ones made of

multiple parts. In the more composite approach, the complex feature is taken as the character and each variant is a different character state. With more reductive coding, separate characters are used to describe variations in the different parts of the complex. In practice there is a continuum of approaches that are more or less composite or reductive. Which approach is adopted can impact upon both what relationships are taken to be supported by the underlying variation and the weight ascribed to that evidence (Wilkinson, 1995a). Intraorganismal homologues are a special case of a complex feature, in which complexity is built upon some degree of repetition. Here we investigate relatively composite and reductive alternative approaches to character construction applied to interorganismal variation in systems of intraorganismal morphological homologues, and we discuss the relative merits of these alternatives in this specific context.

3.1.1 Aetosaurs

Aetosaurians are extinct Triassic suchian archosaurs, the closest living relatives of which are crocodilians (e.g. Gower and Wilkinson, 1996). Their distinctive morphology includes bony dermal armor composed of discrete osteoderms or scutes (Fig. 3.1) and a specialized dentition indicating that they may have been the earliest radiation of herbivorous/omnivorous archosaurs (e.g. Walker, 1961; Parrish, 1994; Small, 2002). The systematics of aetosaurians is of special interest for several reasons. First, their fossil remains have been used as biochron indicators and interest has recently developed in their biogeography and biostratigraphy (e.g. Parrish, 1994; Heckert and Lucas, 1999; 2000). Second, there is lack of agreement concerning their relationships to other major clades of suchian archosaurs (Gower and Wilkinson, 1996; Gower and Walker, 2002). Finally, they have recently been suggested as relevant to the controversy over the phylogenetic affinities of turtles (Hedges and Poling, 1999, and see Chapter 7).

Aetosaurian phylogeny has been addressed in three recently published phylogenetic analyses by Parrish (1994), Heckert *et al.* (1996), and Heckert and Lucas (1999). We provide a critical review of these studies and use this to develop alternative reductive and composite combined data matrices based on these studies. These alternatives differ only in the treatment of intraorganismal homologues. Analyses of these data are used to investigate both aetosaurian phylogeny and the practical impact of alternative approaches to character construction.

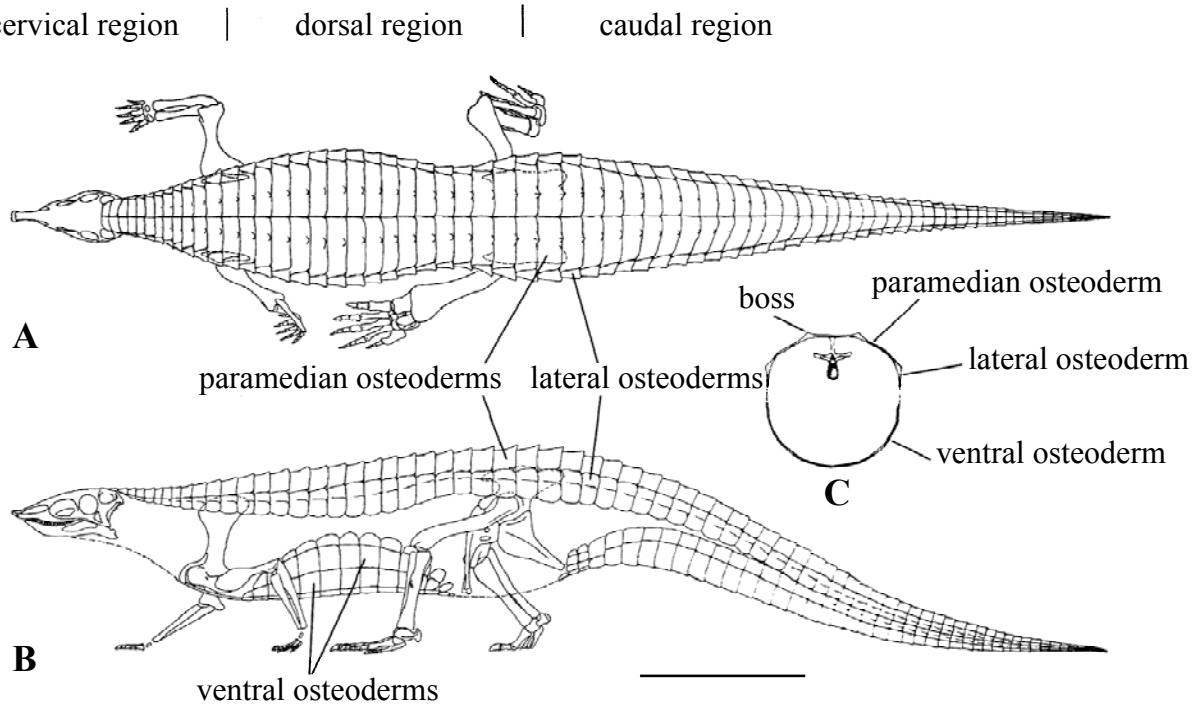


Figure 3.1. Skeletal reconstructions of the Triassic aetosaurian archosaur *Stagonolepis robertsoni* Agassiz, showing the disposition of the dermal ossifications or osteoderms. (A) Dorsal view. (B) Lateral view. (C) Transverse section at mid-body. Scale bar = 0.4 m. Modified from Walker (1961: fig. 23) and reproduced with permission from The Royal Society.

3.2 Materials and Methods

Published data matrices from each of the three previous studies (Parrish, 1994; Heckert *et al*., 1996; Heckert and Lucas, 1999) and any revisions thereof were investigated with parsimony analysis. Combined data matrices incorporating revised characters from all previous studies were developed with either reductive or composite representations of variation in intraorganismal homologues. These were used to investigate the impact of the alternative approaches to character construction in quantitative phylogenetic analyses. Unless stated otherwise, all analyses were performed using PAUP 4.0b10a (Swofford, 2003). Characters were weighted equally and searches for most parsimonious trees (MPTs) were exact (branch and bound). Tree length (TL) and consistency index (CI) were recorded for each MPT. Multiple MPTs were summarized with the strict reduced consensus (SRC) method (Wilkinson and Thorley, 2003) as implemented in RadCon (Thorley and Page, 2000). This method identifies all cladistic relationships that are common to the MPTs and are thus unambiguously supported by the parsimonious interpretation of the data (Wilkinson, 1994a). It may produce multiple consensus trees, together termed a profile. If

the "strict" consensus (Sokal and Rohlf, 1981), referred to here as strict component consensus (Wilkinson, 1994a; Wilkinson and Thorley, 2001a) is informative it will be a member of the SRC profile. RadCon was used to determine Thorley *et al.*'s (1998) cladistic information content (CIC) and Wilkinson and Thorley's (2001b) consensus efficiency (CE). As its name suggests, CIC is a measure of the information content of trees (including consensus trees) based on the number of permitted resolutions of the trees, and CE quantifies how well a consensus represents the set of trees it stands for, scaled between zero (minimal efficiency) and one (maximal efficiency).

Null hypotheses that data are no more structured than expected by chance were tested by randomization using two distinct measures of data quality: parsimony tree lengths (Archie, 1989; Faith and Cranston, 1991) and the number of pairwise hierarchical nestings of characters (Alroy, 1994). These yield matrix parsimony (MP) and matrix nesting (MN) permutation tail probabilities (PTPs) respectively. All randomization tests used 1000 trials giving minimum possible PTPs of 0.001. The distribution of missing data in data matrices is typically non-random and ideally should be held constant during random permutation of the data. This is not possible with PAUP's implementation of the MPPTP, but was applied in our determinations of MNPTPs, using PICA 4.0 (Wilkinson, 2001a). Bootstrapping (Felsenstein, 1985) and decay analysis (Bremer, 1988; Donoghue *et al*., 1992) were used to quantify support for relationships (splits). Bootstrap proportions were based on 1000 replicates and are reported for clades. Decay indices were determined through constrained analyses and are reported for clades and for less inclusive relationships (partial splits) recovered by the SRC method. The latter were determined using backbone constraints (Wilkinson, 1997a). Scope for safe taxonomic reduction, the elimination of taxa that have no effect upon inferred phylogeny (Wilkinson, 1995b), was determined using TAXEQ3 (Wilkinson, 2001b).

We recognize the process that culminates in the recording of a datum in a matrix to be comprised of at least two parts – construction and scoring. Scoring is the ascribing of state(s) to a particular terminal. Construction (also formulation) is more complex, involving the partitioning of phenotypes into discrete characters, the partitioning of variants into character states, and hypothesizing the relations among them (i.e. choosing a character type). Scoring, as understood here, is sometimes termed coding by other authors (e.g. Yeates, 1995), but this term has also been used to describe some aspects of character construction, e.g. additive binary (Farris *et al*., 1970), composite, and reductive coding

(Wilkinson, 1995a). We understand coding to be part of character construction. The dermal ossifications that form the armor of aetosaurians are variably termed osteoderms and scutes throughout the literature, and we use the former term here.

						PTP	
DATA MATRIX	T	C	N	TL	CI	Parsimony	Nesting
P94 full	10	15	3	16	0.938	0.001*	0.001*
P94 ingroup	8	15	-	-	-	0.277	0.005*
P94 ingroup minus *Aetosaurus*	7	15	-	-	-	0.473	0.332
rP94 full	10	15	2	15	1.000	0.001*	0.001*
rP94 ingroup	8	15	-	-	-	0.076	0.002*
rP94 ingroup minus *Aetosaurus*	7	15	-	-	-	0.109	0.118
H96	10	23	2	29	0.793	0.001*	0.001*
rH96	10	22	2	26	0.846	0.001*	0.001*
H99	60	14	10	90	0.656	0.001*	0.001*
rH99	60	14	1	86	0.674	0.001*	0.001*

Table 3.1. Summary statistics for selected analyses. T = number of taxa, C = number of characters, N = number of MPTs, TL = tree length, CI = consistency index. Permutation tail probabilities (PTPs) are for matrix randomization tests using parsimony tree length and pairwise character nesting. * = significant (PTP $\leq$ 0.05).

3.3 Results

3.3.1 A review of aetosaurian phylogenetics

Presented here are reviews of the three previous numerical phylogenetic analyses of aetosaurians, by Parrish (1994), Heckert *et al*. (1996), and Heckert and Lucas (1999). These analyses are treated in chronological order. For each, a summary is given of the published analysis, followed by reports of reanalyses of the data (including any modifications), assessments of support, and discussion of any character construction issues that were identified. Summary statistics for our reanalyses are given in Table 3.1. These reviews form the basis of a combined matrix.

3.3.2 Parrish (1994)

3.3.2.1 Review

Parrish (1994: Table 2) presented a data matrix of eight aetosaurian genera and two outgroups (Prestosuchia and Rauisuchia) scored for 15 binary characters. He reported that parsimony analysis of these data produced three MPTs of TL=16 and CI=0.938, and presented the strict component consensus of these (Fig. 3.2a).

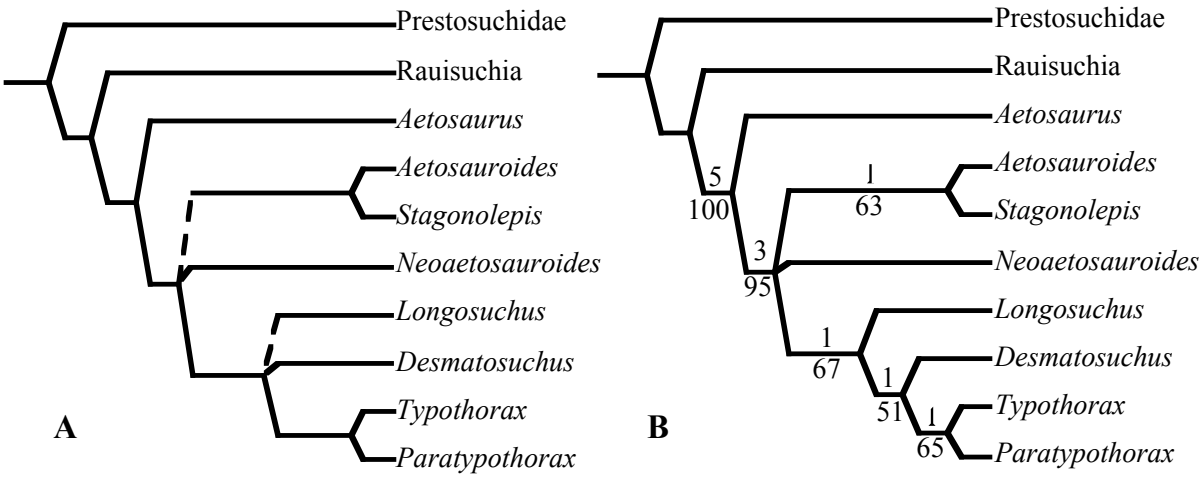


Figure 3.2. (A) Strict component consensus of three MPTs (TL = 16, CI = 0.938) published by Parrish (1994). Dashed lines were unexplained. The same MPTs and consensus were also recovered from analysis of data matrix P94 (CIC = 17.781 bits, CE = 0.968). (B) Strict component consensus tree (CIC = 19.366 bits, CE = 1.000) of the three binary MPTs (TL = 15, CI=1.000) recovered from analysis of rP94 (Table 1). Numbers above and below branches are decay indices and bootstrap proportions respectively.

Consideration of the published matrix (Table 3.2), consensus tree (Fig. 3.2a) and descriptive statistics of the MPTs reveals that the data presented could not be those analyzed. There is no incongruence in the published data (Table 3.2). Consequently, the CI of the MPTs must be one, and the tree length must be equal to the number of (binary) characters (i.e. 15, not 16). Additionally, *Stagonolepis* and *Longosuchus* are scored identically for all characters in the published matrix and must therefore be subtended by the same node in any MPT for these data. This is not true of Parrish's published consensus tree (Fig. 3.2a).

	CHARACTERS		
		1	11111
TAXA	12345	67890	12345
Prestosuchidae	00000	00000	00000
Rauisuchia	11000	00000	00000
Aetosaurus	11111	11000	00000
Stagonolepis	11111	11111	10000
Longosuchus[a]	11111	11111	10000
Longosuchus[b]	11111	11111	00110
Desmatosuchus	11111	11111	01110
Typothorax	11111	11111	?1101
Aetosauroides	11111	111?1	10??0
Neoaetosauroides	11111	11111	000?0
Paratypothorax	????1	11??1	??1?1[c]

Table 3.2. Corrected matrix for Parrish's (1994) data, referred to as "P94". These are the data analysed but not published by Parrish (Parrish pers. comm. 2000). The three underlined character states are changed to "?" in order to remove discrepancies between the matrix and Parrish's (1994) text to produce the matrix referred to as "rP94". *Longosuchus*[a] = scoring of *Longosuchus* in Parrish's published matrix, *Longosuchus*[b] = Scoring of *Longosuchus* in the corrected matrix provided by Parrish (pers. comm.), [c] = Character 15 for *Paratypothorax* was misscored as state 0 instead of 1 in Parrish's published matrix (Parrish pers. comm.).

3.3.2.2 Reanalysis

Reanalysis of the published matrix confirmed its disagreement with the published results, yielding three MPTs of expected length 15 and CI of one. The strict component consensus (and unique SRC) of these trees (not shown) also differs from that published by Parrish, suggesting that the published matrix is not that used in Parrish's analyses. Parrish (pers. comm., 2000) confirmed this and explained that the entire row of data for *Stagonolepis* was inadvertently re-entered for *Longosuchus*, and that *Paratypothorax* was misscored (0 instead of 1) for character 15. The matrix that was originally analyzed and which should have been published (Parrish, pers. comm., 2000) is given in Table 3.2, and

is referred to here as "P94". Analysis of P94 recovers the trees and descriptive statistics reported by Parrish (1994).

There are some discrepancies between P94 and Parrish's (1994) text. *Longosuchus* is scored as having an edentulous anterior premaxilla (character 3), but the description (Parrish, 1994: 196) states that the anterior of the premaxilla is missing, implying that character 3 should be scored as unknown (?) for this taxon. Character 12 concerns the presence or absence of posterior premaxillary teeth. In the matrix, *Longosuchus* is also scored as possessing posterior premaxillary teeth (character 12) and the description of *Longosuchus* supports this, drawing attention to (: 196) "what seems to be a single premaxillary tooth". However, absence of premaxillary teeth (: 207) is described as a synapomorphy of the unnamed clade containing *Longosuchus*, *Desmatosuchus*, *Typothorax* and *Paratypothorax*. Finally, *Aetosaurus* and *Typothorax* are scored as lacking a deep hemispherical fontanelle (state 0, character 14), yet in a list of the synapomorphies (: 207) for the unnamed clade comprising *Longosuchus*, *Desmatosuchus*, *Typothorax* and *Paratypothorax*, this character is described as "indeterminate in *Paratypothorax*, *Typothorax*, *Aetosaurus*, *Aetosauroides* and *Neoaetosauroides*".

A revised matrix ("rP94") that resolves these discrepancies was prepared (see Table 3.2) in which *Longosuchus* was rescored as unknown for character 3, and *Aetosaurus* and *Typothorax* as unknown for character 14. We accepted Parrish's scoring of *Longosuchus* for character 12. Analysis of rP94 recovers two (three binary) MPTs. The rescoring of *Typothorax* removes all conflict from the matrix, and therefore the two trees have a length of 15 and CI of one. The unique SRC (and strict component consensus) (Fig. 3.2b) of the two MPTs is slightly more resolved than Parrish's published tree (Fig. 3.2a).

3.3.2.3 Support

Although P94 and rP94 yield similar phylogenies, they cannot be considered a compelling hypothesis for a number of reasons. Firstly, of the 15 characters in the matrix, two (1 and 2) are parsimony uninformative, and five (3 to 7) provide evidence that only serves to support a split between the outgroup and ingroup taxa. This leaves just eight characters to provide evidence for relationships among the eight included aetosaurians. Of these characters, three (8 to 10) support the ingroup split between *Aetosaurus* and all other aetosaurians, so that just five characters provide evidence for the relationships among the remaining seven genera.

Matrix randomization tests of the full data sets allow rejection of the null hypotheses that P94 and rP94 are no better than comparable random, phylogenetically uninformative data (PTPs < 0.05; Table 3.1), a minimum requirement for phylogenetic data. However, this does not indicate that significant structure is distributed throughout the data (Faith and Cranston, 1991). When the two outgroup taxa are removed and the tests applied to the ingroup-only data (eight taxa and eight characters), only the nesting-based test yields significant PTPs. Alroy (1994) advocated this test for its sensitivity, and our results provide empirical support for this. However, although these data appear to be non-randomly structured, the fact that the ingroup-only data fail the parsimony randomization tests suggests that ingroup trees based on parsimony analysis of these data should be viewed cautiously. With *Aetosaurus* also excluded, matrix randomization tests of the remaining data (seven taxa, five characters), including Alroy's (1994) highly sensitive nesting test, give non-significant results (PTPs > 0.05) despite the complete absence of conflict in rP94. Parrish (1994) did not measure support. With the single exception of the clade comprising all aetosaurians except *Aetosaurus*, all clades within Aetosauria that are supported by the parsimonious interpretation of rP94 (Fig. 3.2b) have minimum decay indices (+1) and low (51 to 69%) bootstrap proportions. We consider rP94 to include too few characters to provide a well supported hypothesis of the relationships of the included aetosaurians.

3.3.3 Heckert, Hunt and Lucas (1996)

3.3.3.1 Review

Heckert *et al*. (1996) scored 23 characters for nine aetosaurian genera, including all genera used by Parrish (1994) plus *Redondasuchus*. Non-aetosaurian outgroups were not included, and trees were rooted on *Aetosaurus*. Exclusion of a non-aetosaurian outgroup meant that Parrish's (1994) characters 1 to 7 were also excluded. Heckert *et al*.'s character 23, describing variation in the gross morphology of the tail region, was not used in all of the analyses they reported. Parsimony analysis of Heckert *et al*.'s restricted data (minus tail character 23) was reported as producing five MPTs, their strict component consensus of which is shown in figure 3.3a. Analysis of the full data set (including character 23) was reported as yielding a single MPT (Fig. 3.3b).

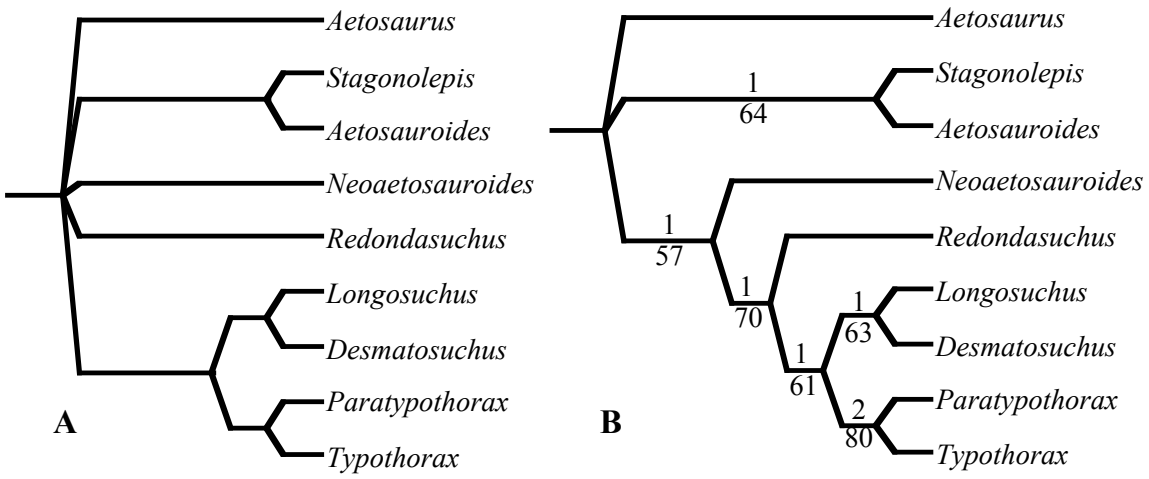


Figure 3.3. (A). Strict component consensus of five MPTs (L=26, CI=0.850) reported by Heckert *et al*. (1996) from analyses excluding their character 23. (B). Single MPT (reported as L = 29, CI = 0.793) from their analysis of their full data. B was recovered in our reanalysis but with L = 27 and CI = 0.815. Numbers above and below branches are decay indices and bootstrap proportions respectively.

3.3.3.2 Reanalysis

Heckert *et al*. (1996) justified the removal of character 23 from some analyses, because "it is not at all clear that reduction of the tail has a single, uniform cause" (: 629), but we have included it in all our reanalyses. Reanalysis of the published matrix yielded Heckert *et al*.'s (1996) MPT (Fig. 3.3b), but with different tree statistics. Heckert *et al*.'s (1996) data matrix was based in part on that presented by Parrish (1994), but several changes in scoring were incorporated. Two of these (from ? to 0 for *Aetosauroides* for character 13, and for *Typothorax* for character 20) were intentional corrections of assumed mistakes in Parrish's scoring (Heckert, pers. comm., 2000). However, three changes (from 0 to 1 for *Aetosaurus* and from ? to 1 for *Paratypothorax* for character 18, and from 0 to 1 for *Longosuchus* for character 21) were accidental (Heckert, pers. comm. 2000). A version of the matrix with these accidental changes rectified was prepared and is referred to as "H96" (Table 3.3). Parsimony analysis of H96 recovers two MPTs, the unique SRC (and strict component consensus) tree of which (Fig. 3.4) differs from the published tree (Fig. 3.3b) in leaving the relationships of *Longosuchus*, *Desmatosuchus*, and (*Paratypothorax* + *Typothorax*) unresolved.

	CHARACTERS				
		1	11111	11112	222
TAXA	12345	67890	12345	67890	123
Aetosaurus	00000	00000	00000	000[a]00	000
Aetosauroides	??000	01?00	000[b]00	001?1	0?0
Stagonolepis	00000	01000	00000	00111	000
Neoaetosauroides	???0?	?1?00	00000	00110	0?1
Redondasuchus	10001	10011	11000	0????	??1
Longosuchus	10000	010?0	00110	01110	0[a]11
Desmatosuchus	11100	01110	00110	01110	111
Paratypothorax	?0010	010?0	00111	11?[a]??	??1
Typothorax	00011	11010	10111	11110[b]	101

Table 3.3. Corrected matrix of Heckert *et al*.'s (1996) data, referred to as "H96". [a] = score as in Parrish (1994) inadvertently changed in Heckert *et al*. (1996), [b] = score as in Heckert *et al*. (1996) intentionally changed from Parrish (1994)

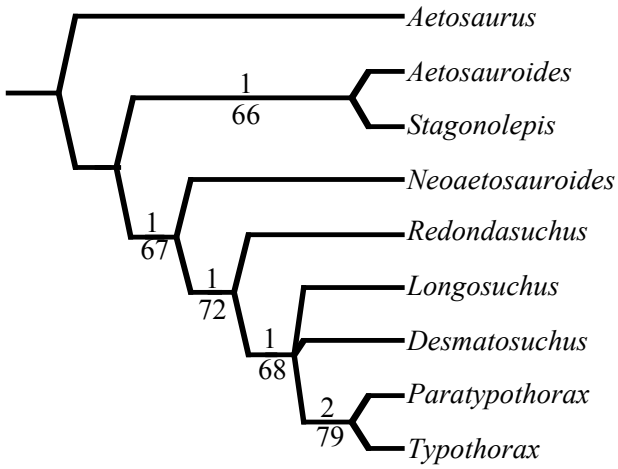


Figure 3.4. Strict component consensus (CIC = 15.460 bits, CE = 0.964) of the two MPTs (L = 29, CI = 0.793) from analysis of data set H96. Numbers above and below branches are decay indices and bootstrap proportions respectively.

3.3.3.3 Support

Randomization tests of H96 yield significant PTPs (Table 3.1), allowing rejection of the null hypothesis that these data are no more structured that would be expected from chance alone. However, decay indices are minimal (+1) for four of the five clades in the strict component consensus tree (Fig. 3.4). The fifth (*Paratypothorax* + *Typothorax*) has a decay index of +2. Bootstrap proportions show only moderate support values (66 to 79%)

for all of the clades. Such low support, coupled with the impact of a few small corrections, indicate that most relationships inferred from H96 cannot be considered robust.

3.3.3.4 Characters

Some aspects of H96 relating either to Heckert *et al*.'s (1996) approach to character construction, or to their scoring of inapplicable characters are problematic. Specifically, we view the independence of one pair of covarying characters to be questionable. In addition, taxa for which particular osteoderms are unknown were nonetheless sometimes scored for characters based on features of those osteoderms.

Characters 4 and 16 describe variation in the shape of the dorsal paramedian osteoderms and the gross morphology of the 'carapace' to which they contribute. Character 4 distinguishes width to length ratios of the dorsal paramedian osteoderms of less than four to one from higher ratios. Character 16 distinguishes the presence or absence of a discoidal carapace. Those taxa with discoidal carapaces also have wide dorsal paramedian osteoderms and vice versa (Table 3.3). Given that broad paramedian osteoderms contribute to a discoidal carapace, the two characters might reasonably be considered to be logically dependent. Heckert *et al*. (1996: 628) acknowledge this linkage, but argue that the two characters are independent "because: (1) it [discoidal carapace] represents a dramatically different body plan amongst the aetosaurs, and (2) it is possible to imagine aetosaurs with narrower paramedians still obtaining a discoidal carapace, or aetosaurs with wide paramedians retaining a more primitive body plan". The latter would seem to require some compensatory reduction in the width of other osteoderms. We also suggest that if similar discoidal carapace shapes reflected dissimilar underlying patterns of osteoderms, then this might reasonably be taken as an indication that the similar carapace shapes were not homologous. Given the absence of aetosaurians with both discoidal carapaces and paramedian osteoderms less than four times as wide as long, or with osteoderms more than four times wider than long and without discoidal carapaces, we consider the hypothesis that the two characters are not independent sufficiently plausible to adopt a more composite character construction. We prepared a revised matrix, referred to here as "rH96" (Table 3.4) in which this pair of reductively coded characters was represented by a single more composite character.

	CHARACTERS				
		1	11111	11122	22
TAXA	12345	67890	12345	78901	23
Aetosaurus	00000	00?00	00000	00000	00
Aetosauroides	??000	01?00	00000	01?10	?0
Stagonolepis	00000	01000	00000	01110	00
Neoaetosauroides	???0?	?1?00	00000	01100	?1
Redondasuchus	10001	10?11	11???	?????	?1
Longosuchus	10000	010?0	00110	11100	11
Desmatosuchus	11100	01110	00110	11101	11
Paratypothorax	?0010	010?0	00111	1????	?1
Typothorax	00011	11010	10111	11101	01

Table 3.4. The data matrix modified from Table 3.3 and referred to as "rH96", incorporating alternative character constructions.

Phylogenetic data matrices often include characters that describe variation with respect to the form of some features that are entirely absent in some of the taxa (e.g. variations in tooth crown morphology in edentulous mammals). Such characters are termed inapplicable. Although the scoring of inapplicable characters and their analytical treatment is controversial (Platnick *et al*., 1991; Maddison, 1993), most workers advocate that where characters are inapplicable, taxa should be scored as unknown, i.e. with missing entries (Hawkins *et al*., 1997; Lee and Bryant, 1999; Strong and Lipscomb, 1999 and see Chapter 2). Heckert *et al*. (1996) do not appear to have adopted any particular convention for the treatment of inapplicable characters. Their data include a number of such characters in which the taxon lacking the feature is scored for one of the character states seen in other taxa. In no case was the scoring justified and the choice of character state therefore appears arbitrary. For example, *Redondasuchus* has been described as lacking lateral osteoderms (Heckert *et al*., 1996) and conventionally would be scored as unknown for characters relating to aspects of the morphology of lateral osteoderms. However, Heckert *et al*. (1996) score *Redondasuchus* as exhibiting state 0 for characters 13, 14, and 15 (Table 3.3), all of which describe variation in lateral osteoderm morphology. This effectively assumes a conditional model of evolution in which state 0 is considered relatively primitive and state 1 derived if lateral osteoderms are themselves derived, or the opposite if it is the absence of lateral osteoderms that is derived. While either model of evolution might be correct, no evidence has been presented to support them and, therefore, we

rescored these three characters in rH96 as unknown (?) for *Redondasuchus* (Table 3.4). A similar treatment of inapplicable character scoring occurs with characters 7 and 8 (Table 3.3), relating to the presence of bosses on the paramedian osteoderms and the position of such bosses, respectively. In Heckert *et al*.'s (1996) matrix (Table 3.3), the position of bosses on osteoderms in taxa which lack bosses (*Aetosaurus* and *Redondasuchus*) has been scored as "on posterior margin". We rescored these taxa as unknown (?) for character 8 in the data matrix rH96 (Table 3.4).

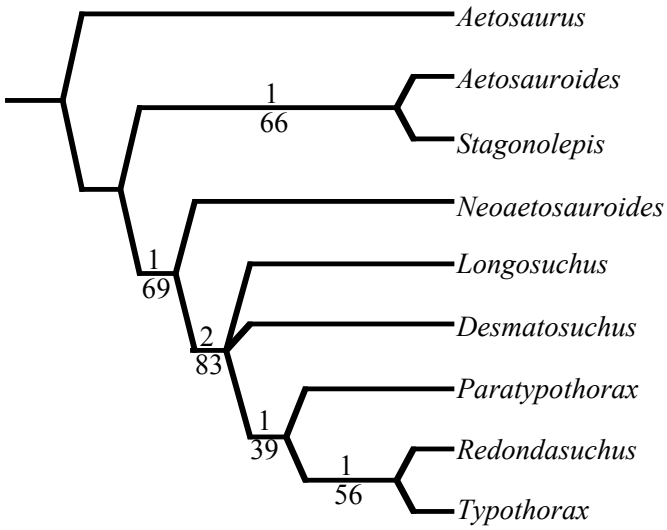


Figure 3.5. Strict component consensus (CIC = 15.459, CE = 0.964) of the two MPTs (L = 26, CI = 0.846) from analysis of data set rH96. Numbers above and below branches are decay indices and bootstrap proportions respectively.

Matrix rH96 (Table 3.4) has significant PTPs (Table 3.1), and parsimony analysis yields two MPTs differing only in their resolution of the positions of *Longosuchus* and *Desmatosuchus*. The relationships in the strict component consensus (Fig. 3.5) differ from those recovered from analysis of H96 (Fig. 3.4) in that *Redondasuchus* is recovered as sister taxon to *Typothorax*. Although the clades (*Paratypothorax* + *Typothorax*) and (*Paratypothorax*, *Typothorax*, *Desmatosuchus*, *Longosuchus*) are lost when the changes in rH96 are implemented, support for the clade (*Paratypothorax*, *Typothorax*, *Desmatosuchus*, *Longosuchus*, *Redondasuchus*) increases. That the position of *Redondasuchus* is affected by our revisions is not surprising given that this taxon was most affected by inapplicable characters because of its apparent lack of lateral osteoderms. The position of *Redondasuchus* proposed by Heckert *et al*. (1996) was thus appears to attributable, in part, to inadvertent errors in the original data.

3.3.4 Heckert and Lucas (1999)

3.3.4.1 Review

Heckert and Lucas (1999) scored 60 characters for 14 taxa, including 11 aetosaurian genera, two species of a twelfth genus, *Stagonolepis*, and an outgroup, *Rauisuchia*. Incorporated among the 60 characters were all but characters 1, 2 and 5 of Parrish's (1994) study, and characters 12, 15 and 23 of Heckert *et al.*'s (1996) study. Heckert and Lucas (1999: 62) reported analyzing "60 characters for *Coahomasuchus* and the 11 taxa listed above". This is confusing, because the listed taxa include *Coahomasuchus* and 12 other taxa, and although they reported recovering 16 MPTs they did not report tree lengths, other descriptive statistics, any of the 16 trees, or their consensus. No reason was given for excluding any taxa at this stage in the analysis. Heckert and Lucas interpreted the consensus tree of their 16 MPTs as confirming their "initial suspicions that *Stagonolepis robertsoni* and *Aetosauroides scagliai* are congeneric, as are *Desmatosuchus* and *Acaenosuchus*, and *Longosuchus* and *Lucasuchus*". On this basis, they removed *Aetosauroides*, *Acaenosuchus* and *Lucasuchus* from subsequent analyses. They also removed *Stagonolepis wellesi* on the grounds that they were sceptical of its distinctiveness from S. robertsoni, and *Redondasuchus* on the basis that it is too incompletely known. A second analysis carried out on the reduced data set of nine taxa was reported as yielding a single MPT (Fig. 3.6).

3.3.4.2 Reanalysis

Our reanalysis of the reduced data set yielded the reported MPT. Analysis of the full published data recovered 10 MPTs, two of three SRC trees of which are shown in figure 3.7. Relationships supported by the full data (Fig. 3.7) conflict with those from the reduced data (Fig. 3.6) in several ways. There is a major shift in the position of *Typothorax*, and smaller differences in the relationships of *Longosuchus* and *Paratypothorax* and of *Stagonolepis* and *Coahomasuchus*. The lack of resolution in the strict component consensus (Fig. 3.7a) is revealed by the second SRC tree (Fig. 3.7b) to be attributable to the instability of *S. wellesi*. Our analysis of the full data does not support the view that *Longosuchus* and *Lucasuchus* are congeneric. The latter is recovered as more closely related to a pairing of *Desmatosuchus* and *Acaenosuchus*. The data matrix also shows that there are character state differences between all of the supposedly synonymous

taxa. In the most extreme case, there are five differences (characters 39, 40, 46, 48 and 52) between *Desmatosuchus* and *Acaenosuchus*, accounting for nearly 25% of those characters that are scored without missing data for both taxa.

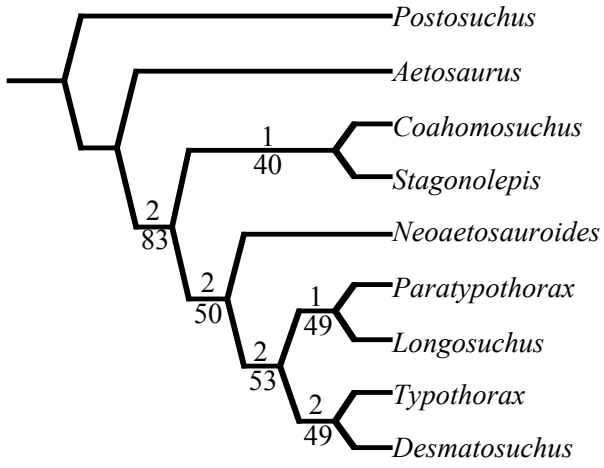


Figure 3.6. Single MPT (L = 76, CI = 0.747) from analysis of Heckert and Lucas's (1999) data with five taxa removed a priori. *Stagonolepis* is represented by *S. robertsoni*, and *Postosuchus* represents the rauisuchian outgroup. Numbers above and below branches are decay indices and bootstrap proportions respectively.

As this example shows, excluding taxa can impact upon the relationships inferred for the remaining taxa. In our view, such exclusion requires justification unless it can be shown to have no impact. In this case, there is no scope for safe taxonomic reduction (Wilkinson, 1995b; Kearney, 2002) and thus we prefer analysis of the full data. There are further reasons in this instance for preferring an analysis of the full data instead of arbitrarily removing taxa. The 10 MPTs it yields are not unmanageable. Two of the three SRC trees for the full data (Fig. 3.7) are more informative than the single MPT for the reduced nine taxon matrix (CIC = 31.488 and 32.094 versus 17.044 bits) and their efficiency is high.

After consultation (Heckert, pers. comm., 2000), character 3 (teeth recurved (0) or conical (1)) was rescored for *Paratypothorax*, from 1 to ? (Table 3.5) to again resolve the inadvertent change from the scoring of this character in Parrish (1994). We refer to this modified data as "H99". The alteration had no effect on the relationships recovered in our reanalyses. Some conflict exists between the matrix and Heckert and Lucas' (1999) text that we were unable to resolve. A maxillary tooth row that does not extend anterior to the posterior end of the external naris (character 5) is scored as present in *Neoaetosauroides* but reported as unknown for that taxon in a list of synapomorphies for all aetosaurians except *Aetosaurus* (Heckert and Lucas 1999: 63). Similarly, *Longosuchus* is scored as

possessing posterior premaxillary teeth (character 6) but reported as unknown in a synapomorphy list (: 64). Heckert and Lucas list state 1 of their character 9 'weakly supporting' the clade comprising *Neoaetosauroides*, *Typothorax*, *Desmatosuchus*, *Longosuchus* and *Paratypothorax* (: 64). However, only *Neoaetosauroides* is scored with this state: *Longosuchus* and *Desmatosuchus* are scored with state 0, and *Typothorax* and *Paratypothorax* as unknown. Finally, for characters 11 and 12 *Coahomasuchus* is scored as possessing state 1, but is listed (: 63) as unknown for both characters. In each of these cases, with the exception of character 9, we adopted a conservative approach and scored taxa subject to contradictory reports with missing entries (Wilkinson, 1997a). For character 9, the contradiction is less clear and we employed the scoring in the original matrix. These discrepancies will need to be addressed in future studies.

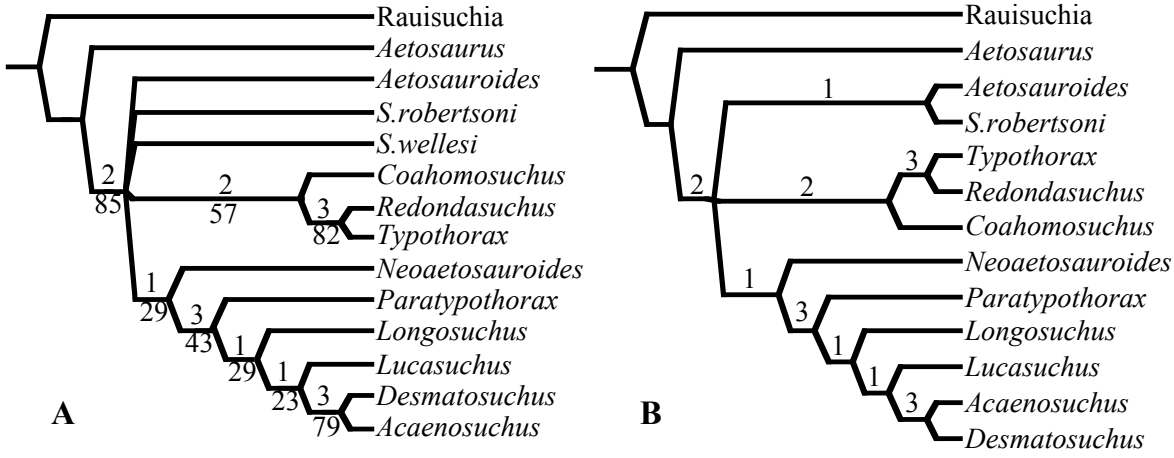


Figure 3.7. RCC profile (in part) of ten MPTs (L = 90, CI = 0.656) from analysis of Heckert and Lucas's (1999) full data. (A) Strict component consensus (CIC = 31.488 bits, CE = 0.903). (B) Most informative RCC tree (CIC = 32.094 bits, CE = 0.920) excluding *Stagonolepis wellesi*. A third and less informative RCC tree (CIC = 11.759 bits, CE = 0.337) completes the profile. Numbers above and below branches are decay indices and bootstrap proportions respectively.

3.3.4.3 Support

Randomization tests of both the complete and reduced (nine taxa) versions of H99 yielded significant PTPs (Table 3.1). With the full data, decay indices are minimal (+1) for three of the clades in the strict component consensus tree (Fig. 3.7a) and no clade has a decay index greater than +3. The additional relationship between *Stagonolepis robertsoni* and *Coahomasuchus* when *S. wellesi* is ignored (Fig. 3.7b) has a minimal decay index. Bootstrap proportions are greater than 50% for only four clades, with strongest support (86%) for *Aetosaurus* lying outside all other aetosaurians. No other clades were recovered in more than 77% of bootstrap trees. With the reduced nine taxon data, bootstrap support is

again highest (83%) for the clade including all aetosaurians except *Aetosaurus*, but no other clades appear in more than 53% of the bootstrap trees, and no clades have decay indices greater than +2. These overall low levels of support for the nine taxon analysis are unsurprising given that over half of the characters (31 of 60) are parsimony uninformative for these restricted data.

3.3.4.4 Characters

Of the 60 characters in H99, 33 relate to variation in osteoderm morphology. We find some aspects of Heckert and Lucas' (1999) character construction to be questionable, leading us to produce a revised matrix, referred to as "rH99" (Characters 1-60 in Table 3.5). There are several instances of taxa seemingly arbitrarily scored with states of characters that are inapplicable and we preferred to rescore these as missing. *Redondasuchus* lacks lateral osteoderms (Heckert *et al*., 1996) but was originally scored as possessing state 0 for the 10 characters (46-53, 55, 57) that relate to variations in the morphology of lateral osteoderms. Character 35 describes the position of bosses on osteoderms. The three taxa (*Coahomasuchus*, *Typothorax* and *Redondasuchus*) that lack bosses on all osteoderms were originally scored as having bosses which are not in contact with the posterior margin of the osteoderm. Similarly, these three taxa were scored as having various forms of the (non-existent) bosses on their dorsal osteoderms (characters 39 and 40). Most of the taxa (Rauisuchia, *Aetosaurus*, *Stagonolepis robertsoni*, *S. wellesi*, *Longosuchus*, *Lucasuchus*, *Desmatosuchus*, *Acaenosuchus*, *Aetosauroides*, *Neoaetosauroides* and *Paratypothorax*) are scored as lacking a ventral keel or strut on dorsal paramedian osteoderms (character 43). However, they are also scored as having ventral keels that are continuous across the width of osteoderms (character 44). We rescored these taxa as unknown for character 44. Characters 49, 50 and 51 describe the presence or absence of lateral spikes on lateral osteoderms, while character 53 describes variation in the angle of spikes on lateral osteoderms. All taxa lacking lateral spikes (Rauisuchia, *Coahomasuchus*, *Aetosaurus*, *Stagonolepis robertsoni*, *Aetosauroides* and *Neoaetosauroides*) were rescored as unknown for character 53.

Randomization tests of rH99, including all taxa, yield significant PTPs (Table 3.1). Parsimony analysis recovered 1 MPT, shown in figure 3.8. Comparison of this tree with the consensus tree in figure 3.7a indicates that our alterations impact upon what can be inferred from the data. In addition to providing resolution, albeit weakly, of the relationships of *Aetosauroides* and the two species of *Stagonolepis*, our alternative

character constructions result in reduced support for the clades (*Redondasuchus*, *Typothorax*, *Coahomasuchus*), and (*Paratypothorax*, *Longosuchus*, *Lucasuchus*, *Desmatosuchus*, *Acaenosuchus*).

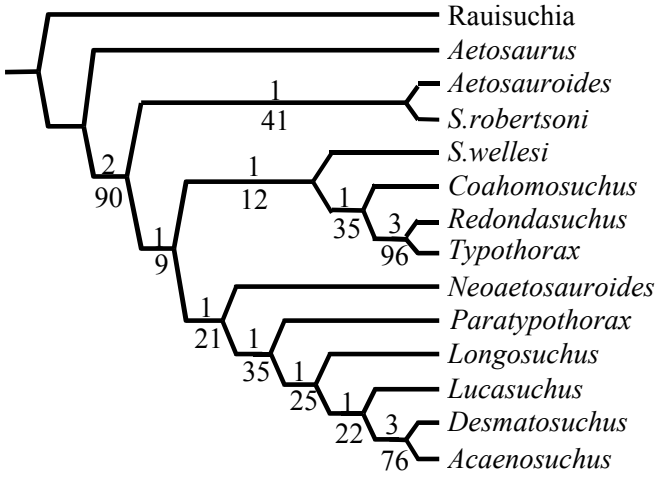


Figure 3.8. Single MPT (L = 86, CI = 0.674) from analysis of data set rH99. Numbers above and below branches are decay indices and bootstrap proportions respectively.

3.3.5 Combined Matrix

The above reviews of the three previous analyses of aetosaurian phylogeny (Parrish, 1994; Heckert *et al*., 1996; Heckert and Lucas, 1999) address a number of character construction and scoring problems. Typographical errors in the published matrices and discrepancies between matrices and character descriptions were resolved, and alternative codings were introduced for some characters. From our reviews, we constructed a combined matrix. This was based primarily on the latest, and most extensive study (Heckert and Lucas, 1999). Characters used in the two earlier studies (Parrish, 1994; Heckert *et al*., 1996) that are not present in the data of Heckert and Lucas (1999) were added to create the combined matrix. The combined matrix (Table 3.5) comprised all 60 characters for the 14 taxa included in Heckert and Lucas' (1999) matrix, plus characters 1, 2 and 5 from Parrish (1994) and 12, 15 and 23 from Heckert *et al*. (1996). Taxa were scored as unknown (?) in the combined matrix for those characters that they had not been scored for in any of the three analyses.

TAXA	CHARACTERS 1 11111 11112 22222 22223 33333 33334 44444 44445 55555 55556 66666 6 12345 67890 12345 67890 12345 67890 12345 67890 12345 67890 12345 67890 12345 6
Rauisuchia	00000 00000 00000 00000 00000 00000 00000 00000 00020 00020 00000 00020 00000 0
Coahomasuchus	????? ?1??? ????1 ?1??? 1?11? ??100 0?00? 111?? 00?20 00000 00?00 01011 ????? ?
Aetosaurus	11000 00101 11111 01?0? 11111 ?1100 00000 00000 00020 00000 00?0? 01011 11100 0
Stagonolepis robertsoni	11111 01101 11111 01100 11111 01100 00001 10000 00020 00000 00?00 11011 11100 0
S. wellesi	????? ????? ????? ?1110 1???? 01100 00001 ?0010 00020 00010 00000 11011 ????? ?
Longosuchus	11111 ?1101 11111 11?0? 11110 0?100 10001 00000 ?0020 10111 11010 01?11 11100 1
Lucasuchus	????? ?1??? ????? ??00? ????? 01100 10000 00010 ?0021 1?111 11000 01??? ????? ?
Desmatosuchus	11110 11101 11111 11011 11111 11110 11100 00010 10021 11111 01011 11??? 11100 1
Acaenosuchus	????? ????? ????? ????? ????? ??110 11100 00001 ?0021 01011 ?0021 11??? ????? ?
Typothorax	11110 101?1 11111 01110 11111 01101 01111 111?? 10100 01010 01110 01??1 11101 1
Aetosauroides	11111 0?1?1 11??? 01100 11111 01100 00001 00000 00020 00000 00?00 11011 11100 0
Neoaetosauroides	11111 0?111 11?11 11?1? 11111 ??100 00000 00010 00020 00000 00?00 01012 11100 1
Paratypothorax	????1 ????? ????1 ??1?? ????? ??101 ?0001 ?0010 10020 00101 11110 11??? ??101 1
Redondasuchus	????? ????? ????? ????? ????? ??100 01111 111?? 11110 ????? ????? 0???? ??210 1

Table 3.5. Matrix combining characters and data from the three previously published studies of aetosaurian phylogeny. Character numbers 1-60 are characters 1-60 of Heckert and Lucas (1999), characters 61to 63 are characters 1, 2 and 5 of Parrish (1994), and characters 64-66 are characters 12, 15 and 23 of Heckert *et al*. (1996). The composite combined matrix is effected by removing characters 33, 38, 47, 55 and 58. Character 3 for *Paratypothorax* was rescored

Within these combined data (referred to here as the "reductive combined" matrix), we identified three sets of covarying characters that describe variation in intraorganismal

homologues. There are alternative, more composite constructions for these sets of characters whereby each set is replaced by a single character, as described below. The more composite constructions for these characters are implemented in a modified version of the combined data (see Table 3.5) referred to here as the "composite combined" matrix.

Heckert and Lucas' (1999) characters 32, 33, 47 and 58 describe variation in the patterning (radiate or random) on the cervical paramedian, dorsal paramedian, lateral, and ventral osteoderms respectively. Excluding missing data, these characters almost all covary. The single exception is that *Redondasuchus* is scored as having a radiate patterning on its lateral osteoderms (character 47) and random patterning on all other osteoderms. However, as noted above, *Redondasuchus* is believed to lack lateral osteoderms (Heckert *et al*., 1996), and we prefer to score this character as unknown for this taxon. In our alternative (see Discussion) construction the four characters are merged into a single character (character 32, Table 3.5). Until aetosaurian specimens are documented that exhibit radiate and random patterning in the different anatomical regions this is, at the very least, a plausible alternative (see Discussion).

Similarly, Heckert and Lucas' (1999) characters 29 and 55 describe the presence or absence of anterior bars on dorsal paramedian and lateral osteoderms. The characters are given as follows: 29) anterior bars on dorsal paramedian osteoderms: present or not applicable (0), absent (1); 55) anterior bars on lateral osteoderms: present (0), absent, replaced by laminae (1). In character 29, "not applicable" is combined in a single character state along with "present" and this unusual construction was not explained. More importantly, the character state distributions for these two characters are virtually identical among the included taxa (except for *Aetosaurus* which is scored '0' for character 29, and '?' for character 55). In the absence of specimens exhibiting anterior bars on either their dorsal or lateral osteoderms only, we merged characters 29 and 55 into a single character (character 29, Table 3.5), maintaining a '0' score for *Aetosaurus*.

Finally, in Heckert and Lucas' (1999) study, all taxa with bosses on their dorsal paramedian osteoderms (character 37) are scored as also having bosses on their caudal paramedian osteoderms (character 38), while taxa that lack bosses on their dorsal osteoderms are scored as also lacking them on their caudal osteoderms. We merged Heckert and Lucas' characters 37 and 38 into a single character in the composite combined matrix (character 37, Table 3.5).

Analysis of the reductive combined matrix produced a single MPT, shown in figure 3.9a. This tree has the same topology, and essentially the same support, as that recovered

from analysis of our altered version of Heckert and Lucas' (1999) data (data set "rH99"). This is unsurprising given that four of the six characters added from Parrish (1994) and Heckert *et al*. (1996) are parsimony uninformative.

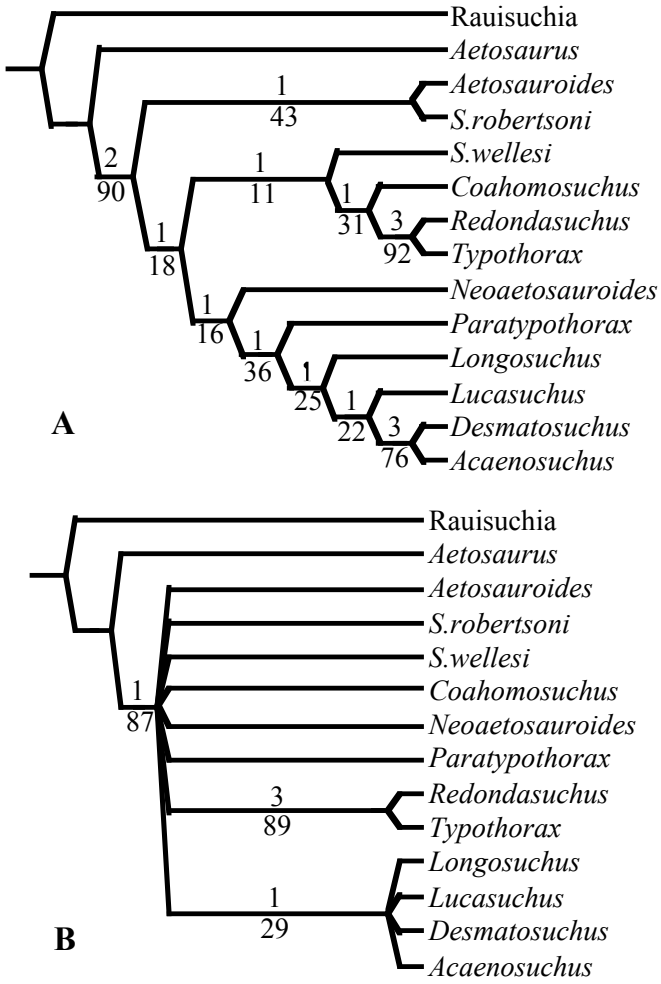


Figure 3.9. (A) Single MPT (L = 91, CI = 0.681) from analysis of reductive version of combined data from the three previous studies of aetosaurian phylogeny. (B) Strict component consensus of nine MPTs (L = 86, CI = 0.682) from analysis of modified (more composite) version of combined data (CIC = 17.251, CE = 0.4924). Numbers above and below branches are decay indices and bootstrap proportions respectively.

Analysis of the composite combined data (see Table 3.5) yielded nine MPTs. The SCC (Fig. 3.9b) of these MPTs is poorly resolved. All nodes supported by a decay value of one in the analysis of the reductive combined data are lost, except for that grouping *Longosuchus, Lucasuchus, Desmatosuchus* and *Acaenosuchus*. The sister-group relationship of *Desmatosuchus* and *Acaenosuchus*, which is supported by a decay value of three in the analysis of the reductive combined data, is lost. Support for *Aetosaurus* as the sister-group to all other aetosaurians is reduced.

	TAXA		
		1	1111
SPLIT	12345	67890	1234
1	.....	....*	...*
2	.....	****.	....
3	.*.**	*****	****
4	.?.*.	.....	*...
5	.....	.**?.	....
6	.?..?	****?	.?**
7	.?..*	????*	.?**
8	.?..*	????*	..?*

Table 3.6. Partition table showing relations (full and partial splits) common to the nine MPTs from analysis of the composite combined matrix. The symbols '.' and '*' indicate the partition of taxa in the corresponding split. '?' indicates exclusion of taxa from partial splits. Taxon abbreviations: 1 = Rauisuchia, 2 = *Coahomasuchus,* 3 = *Aetosaurus,* 4 = *Stagonolepis robertsoni,* 5 = *S. wellesi,* 6 = *Longosuchus,* 7 = *Lucasuchus,* 8 = *Desmatosuchus,* 9 = *Acaenosuchus,* 10 = *Typothorax,* 11 = *Aetosauroides,* 12 = *Neoaetosauroides,* 13 = *Paratypothorax,* 14 = *Redondasuchus*.

The reduced consensus profile of the nine MPTs from analysis of the composite combined matrix produces a profile of six SRC trees, comprising the SCC (Fig. 3.9b) and five other trees. Examination of the SRC trees and their summary partition table (Table 3.6), reveals that the cause of the lack of resolution in the SCC is complicated and can not be attributed to the instability of just one or two taxa. The three reductive sets of intraorganismal homologue characters and their composite alternatives were implemented separately to further explore the cause of loss of resolution. Implementing the composite versions of only characters 32, 33, 47 and 58 or 29 and 55 does not alter the topology of the single MPT recovered from analysis of the reductive combined matrix, but each composite character lowers support for the *Desmatosuchus* + *Acaenosuchus* clade by 1 in decay analyses. In contrast, implementing only the composite version of characters 37 and 38 has a major impact on tree topology, leading to 66 MPTs. The two reductive characters (37, 38) support the clade comprising *Coahomasuchus*, *Typothorax* and *Redondasuchus*, which is present in the MPT recovered from analysis of the reductive combined matrix. Collapse of this clade renders much of the rest of the tree highly unstable, indicating a complex interplay among the remaining data for these taxa.

3.4 Discussion

3.4.1 Aetosaurian phylogeny

All three of the published studies reviewed here (Parrish, 1994; Heckert *et al*., 1996; Heckert and Lucas, 1999) are worthy preliminary investigations into the phylogenetic relationships of a little discussed group. They have provided new morphological data for *Longosuchus, Redondasuchus* and *Coahomasuchus,* and identified potentially useful systematic characters. Despite this, all three previous studies and our combined analyses consistently support only three hypotheses of relationships. These are: that *Aetosaurus* is the sister-group of all other aetosaurians, that *Aetosauroides* is the sister-group of *Stagonolepis* (*robertsoni*), and that *Longosuchus* and *Desmatosuchus* are more closely related to each other than either is to *Neoaetosauroides*. These are the only hypotheses in which we are willing to invest much confidence.

We suggest that the results of previous studies and our own reanalyses should not, for the most part, be accepted as robust hypotheses of aetosaurian interrelationships. This conclusion follows from 1) lack of agreement among different studies, 2) generally low support values in each of the analyses, and 3) sensitivity to alternative character constructions. Much instability is likely to result from abundant missing entries, and less pessimistic assessments of the robustness of some relationships might be achieved using more sensitive methods such as reduced consensus bootstrapping (Wilkinson, 1996) and double decay analysis (Wilkinson *et al*., 2000). However, we conclude that issues of character construction and scoring should be resolved before more extensive investigation of support is merited. Aetosaurian phylogenetics could benefit from better fossils and additional characters from more character systems, but ultimately, with fossil data, there will be an upper bound. This underlines the importance of getting character construction right. As highlighted in the review of aetosaur analyses, future studies of aetosaurian phylogeny must resolve outstanding issues of scoring and should not exclude taxa without good reason. Importantly, such studies will have to address issues of character construction including the treatment of intraorganismal homologues.

3.4.2 Intraorganismal Homology, Character Independence and Character Construction

Our review of aetosaurian phylogenies highlights the potential for alternative approaches to constructing characters from variations in osteoderms in particular, and from systems of intraorganismal homologues in general. Our results demonstrate that alternative

approaches can have a profound impact upon phylogenetic conclusions, both in terms of the relationships that are recovered and the apparent strength of support for those relationships. Osteoderms comprise a system of more or less similar units that we presume are intraorganismal homologues, meaning that they are instances of a repeated pattern that has some common cause (Ghiselin, 1976), or are instances of a repeated or common developmental pattern (Roth, 1984). A moment's reflection on repetition in nature reveals that intraorganismal homology is not a minor phenomenon. Repetition is ubiquitous at all levels of organismal organization and is an important component of much complexity. Wilkinson's (1995a) discussion of composite and reductive coding focused on spatially associated complex structures, such as entire organs, and did not consider intraorganismal homologues per se. However, the distinction between reductive and composite coding applies equally to intraorganismal homologues, which are a special case of complexity built upon similar units that may or may not be spatially or temporally associated.

Heckert and Lucas (1999) constructed a number of sets of characters by treating similar variations in what they viewed as different osteoderm regions as the bases of independent characters. Osteoderm morphology does vary within aetosaurians, making it possible to distinguish paramedian from lateral osteoderms, and often to identify from which approximate region along the axial skeleton isolated osteoderms may originate (e.g. Walker, 1961; Long and Ballew, 1985). However, there can be uncertainty over the regional identity of isolated osteoderms, and similarity among osteoderms from different regions of different taxa (e.g. Hunt and Lucas, 1991: 732). Three of Heckert and Lucas' osteoderm character sets covary in the distribution of their character states. Our reaction to these relatively reductive character constructions in Heckert and Lucas (1999) was that they are too reductive. Naylor and Adams (2001: 450) reacted similarly to several sets of mammalian dental characters used by O'Leary and Geisler (1999) noting that "Because the same underlying genetic architecture generates teeth in a particular tooth group, similar structures on different teeth (e.g. the hypocone) are de facto serially homologous. Therefore, measuring the same feature on multiple teeth in a tooth group represents a redundant and non-independent sampling."

Character independence is considered a fundamental assumption of many phylogenetic methods both for choosing among trees and in assessing support (e.g. Farris, 1973b; Felsenstein, 1985). Characters are logically dependent if the scoring of one or more characters entails some restriction on the coding of another character, and they are biologically dependent if their evolution is causally linked, as might be expected in highly

integrated functional complexes (Wilkinson, 1995a). Logically independent characters may be more or less biologically independent, contingent upon the actual process of evolution. Biological dependence can be viewed probabilistically: if the probability of transformation between the states of one character is conditional upon state changes in one or more others then the characters are dependent (O'Keefe and Wagner, 2001).

Independence is a simplifying assumption that facilitates quantitative evaluation of the weight of evidence and is therefore a desideratum of methods that assume independence. The link between independence and weight of evidence is important because the potential danger in violating the assumption is that too much weight is given to some misleading evidence. For example, the two binary characters 'X wider than long or not' and 'X shorter than broad or not' are simply different ways of expressing the same notion. Using both characters, the underlying variation is given twice the weight (assuming equal weighting) than if just one of these logically dependent characters is used. Similarly, if parallel variations in aetosaurian osteoderms resulted from global changes to the aetosaurian osteoderm system, reductive coding would violate the assumption of biological character independence and overweight the evidence.

Biological dependence and correlated character evolution are believed to be common in morphology (e.g. Emerson and Hastings, 1998). They have been shown through simulation to have the potential to reduce accuracy of parsimony analyses (Wagner, 1998; Huelsenbeck and Nielsen, 1999) and are expected, as found here, to exaggerate support measures (O'Keefe and Wagner, 2001). Detecting and appropriately weighting correlated character evolution resulting from biological dependence is therefore an important issue in phylogenetics (Sneath and Sokal, 1973; O'Keefe and Wagner, 2001). Biological dependence can be complete or partial. For example, if one or more character state transitions entail some other transition (so that the conditional probability of the latter on the former is one), then dependence is complete, whereas if the former merely makes the latter more likely, then the dependence is partial. Several workers have proposed methods for detecting correlated evolution given a phylogenetic tree (Maddison, 1990; Pagel, 1994; Maddison, 2000). O'Keefe and Wagner (2001) have developed very promising statistical tests of correlated character evolution based on patterns of mutual character compatibility that can be used prior to building trees, and which are applicable whether dependence is complete or partial. The reductive codings of aetosaurian osteoderm characters suggest a special case, in which complete dependence results from a single underlying change that produces the same kind of variation in different subsets of

intraorganismal homologues that have been individuated on some other basis. Hecht and Edwards (1977) suggested that suites of characters resulting from change in a single developmental mechanism should be treated as a single character. In such cases, the reductive characters repeat the same pattern of character state distributions (they covary) and have the same patterns of compatibility. The alternative, more composite approach leads to a single character with the same distribution of character states as the reductive characters. The reductive and composite alternatives produce characters having the same phylogenetic significance (in the sense of supporting the same relationships) but they ascribe different weights to the variation.

That Heckert and Lucas' (1999) reductive characters describe the same kind of variation in aetosaurian osteoderms of different regions is key to our preferring a more composite approach. For example, patterning on the osteoderms of the cervical paramedian, dorsal paramedian, lateral and ventral osteoderms is either radial or random. Each reductive character represents hypothesized interorganismal homology of, and explanation for, the similarity of the patterning of the osteoderms of a particular region. Because osteoderms comprise a system of intraorganismal homologues, we consider it plausible that the covarying interorganismal similarity of patterning in different regions can be explained by homology, and the covariation by global change to the system rather than by separate local changes. Reductive coding treats the intra or interorganismal similarity of similar patterning on different regions, such as cervical and dorsal, as coincidental and not homologous. Composite coding, treating variations across the whole osteoderm system as the character, provides a potential explanation for observed similarity of both intra and interorganismal homologies that is more complete, more parsimonious and more plausible.

Covariation of characters does not entail any dependence between them. Conversely, if characters do not covary this does not guarantee their independence, complete or partial, and it does not eliminate concern over appropriate weight (O'Keefe and Wagner, 2001). However, lack of covariation does indicate that not all inter or intraorganismal similarities can be homologous. Characters that do not covary provide evidence for different phylogenetic relationships. Separate character state changes in different parts of the tree must be invoked to explain the observed different distributions whether these events are causally independent or not. We view some degree of independence as a plausible explanation of such separate changes and consequent lack of covariation and homology. Thus, we take lack of covariation in characters as evidence of

character independence and a cause for less concern over potential overweighting by reductive coding. In practice, any overweighting of characters that do not covary is spread across different relationships. With covariation, which is readily identified by inspection of the data, any overweighting is more concentrated. The covariation of reductive characters describing the variation of intraorganismal homologues makes the danger of overweighting particularly severe.

Investigation of character dependence in non-covarying characters requires advanced techniques such as those proposed by O'Keefe and Wagner (2001), whereas the special case we are concerned with here is amenable to simple and routine evaluation. Six sets of dental characters used by O'Leary and Geisler's (1999) were identified a priori as potentially dependently linked as serial homologues by Nalyor and Adams (2001: 450). To test this, Naylor and Adams generated a matrix of pairwise differences among all 45 dental characters and performed a principle coordinates analysis. They reported that four of the six characters sets identified a priori formed distinct clusters, supporting their assessment that "the characters within each of these four sets are not independent". Though not stated explicitly, the failure of the other character sets to cluster together was taken as evidence for their independence. We note that, contrary to Naylor and Adams (2001), one of the four sets of characters accepted as non-independent (characters 74 - 76), does not form an exclusive cluster (see their Fig. 3). Naylor and Adams' (2001) approach agrees with ours in proceeding from an a priori assessment of potential dependence founded on hypotheses of intraorganismal homology to a test of the predicted association of candidate sets of characters. It differs in the use of ordination and clustering to test the association of characters.

We examined the distributions of character states in the six sets of characters identified a priori by Naylor and Adams (2001). The two sets considered by Naylor and Adams to comprise independent characters on the basis of their failure to cluster in ordination have very different character state distributions. In contrast, within all four sets considered to comprise dependent characters by Naylor and Adams, the character state distributions are very similar. In three sets they are identical or identical except for missing entries, and in the fourth (the one that does not form a discrete cluster in the ordination) characters differ in the scoring of a single taxon. Simply on the basis of their covariation (which entails their co-clustering), we accept three sets of characters as comprising potentially redundant characters that would be better represented by a single composite character. On the basis of their lack of covariation, we are more accepting of the reductive

coding of the three remaining sets (notwithstanding additional insights that may be gained through the application of advanced techniques).

Morphologists routinely construct characters from systems of intraorganismal homologues. The approach to character construction adopted appears to us to be mostly influenced by the degree to which subsets of intraorganismal homologues can be individuated based on intrinsic features. Evolutionary differentiation of units or groups of units within a system of intraorganismal homologues must result from local (with respect to the system) evolutionary change, and this lends itself to reductive coding. Although some workers have used relatively reductive codings of parallel variations in extrinsically individuated subsets of intraorganismal homologues (Heckert and Lucas, 1999; O'Leary and Geisler, 1999), there is a clear preference for more composite coding whenever interorganismal variations in intraorganismal homologues could be plausibly explained by a single change. For example, we know of no case where the same variations in the antimeres of bilaterally symmetric organisms have been treated as separate characters. We believe that in adopting composite character construction for parallel variations in intraorganismal homologues, common practice is good practice. The remaining discussion is intended to clarify why this is so.

In the more general context of complexity, Wilkinson (1995a: 307) argued that neither reductive nor composite coding "has a monopoly of advantages or dangers and the task of constructing characters from character complexes or complex characters requires due consideration of these alternative approaches." Choosing between more reductive or composite character constructions turns ultimately on assessments of plausibility and must be made on a case by case basis. To guard against overweighting by excessive reductive coding in cases where the reductive characters covary, Wilkinson (1995a: 302) suggested asking whether covarying reductive characters can be combined into a composite character representing a real unit of biological organization with parts that are plausibly biologically dependent and evolve in concert. He suggested that if answers are affirmative, then the more composite alternative should be considered. Applied to the reductive coding of aetosaurian osteoderms, the answers are affirmative by virtue of the relation of osteoderms as intraorganismal homologues, suggesting that composite coding is sufficiently plausible to warrant consideration in the special case of covarying intraorganismal homologues.

Further guidance on the choice of character construction comes from Hennig's auxiliary principle. Any similarity between organisms may be explained as either homologous or convergent (homoplastic). Confronted with this truism, Hennig (1966)

proposed that similarities should be explained as homologous unless incongruence entails convergence. This is Hennig's auxiliary methodological principle, and he argued that it is needed to establish a link between similarity and phylogeny. If convergence were our preferred explanation, similarities would not be taken as evidence of relationships.

Hennig's auxiliary principle invokes a common cause in preference to separate causes. It can be readily interpreted in the context of character construction as advising phylogeneticists to represent similarities as a priori hypotheses of homology. Typically, this is achieved through character state identity, but in the case of multistate characters the principle can also lead to specific ordering of character states (Wilkinson, 1992b). The reductive approach to aetosaurian osteoderms treats the evolution of, for example, bosses on the lateral and paramedian osteoderms to be independent events. The similarity that exists between bosses on lateral and paramedian osteoderms is therefore interpreted as coincidental and homoplastic, in violation of Hennig's auxiliary principle. With composite coding, the similarity of the parallel variations in different regions is taken as homologous in greater conformity to Hennig's auxiliary principle.

We believe that the conformity to Hennig's auxiliary principle of the composite coding of covarying differences among intraorganismal homologues provides a methodological justification for common practice and the seemingly near universal preference for this sort of coding. However, conformity to Hennig's auxiliary principle may be more or less impressive depending on the plausibility of common cause. Many factors may impact upon this plausibility and decisions on character construction must be made on a case by case basis. We consider the hypothesis of single change implicit in the composite coding of covarying intraorganismal homologues to be at least sufficiently plausible to always warrant explicit consideration. More reductive treatments are not ruled out in specific cases, but these might require some additional justification.

Our discussion of the role of intraorganismal homology in character construction is a cursory foray into an important but under-appreciated topic. We have dealt only with relatively simple cases and expect that biological complexity will confront phylogeneticists with more difficult, but also more interesting gray areas. We do not claim that we are inventing or advocating any novel principles for phylogeneticists. Rather, we hope to have clarified why it is that what the majority of phylogeneticists do in adopting composite character constructions in practice is good practice, and why more reductive codings, although potentially justifiable, seem intuitively problematic.

Chapter 4: Compatibility Methods and Their Uses

4.1 Introduction

Virtually all published morphological phylogenies in recent years have been produced using parsimony analysis. Parsimony is a model of evolution that assumes that the simplest possible explanation for character evolution in terms of tree steps represents the best estimate of phylogeny. Such an assumption, although seemingly rational, is not necessarily true, and therefore may not produce the true phylogeny. Alternative methods, such as maximum likelihood, based on different hypotheses of evolution are also available, and are often employed for analysis of molecular data. Attempts have been made to produce likelihood methods for use with morphological data (Lewis, 2001). Currently such methods are only rarely implemented, although Bayesian inference (which is based on a likelihood foundation) is becoming more popular. One alternative to parsimony and likelihood for investigating both morphological and molecular data is compatibility analysis.

The aims of this chapter are to briefly review current compatibility methods, their uses and discuss their limitations. Two new compatibility methods are introduced:

1) **Fuzzy compatibility**. This is a method for quantifying how incompatible two incompatible characters are based on the minimum number of taxa that must be rescored to make the characters compatible. An algorithm for calculating this measure is presented.
2) **Boildown bootstrap**. This is a method that provides a way of deciding when to halt a boildown procedure. Two types of boildown bootstrap are introduced.

New ideas concerning potential problems for compatibility analyses due to polymorphisms, linked characters and inapplicable data within data matrices, and biases towards unbalanced trees in both compatibility and parsimony are discussed.

4.2 Compatibility

The concept of character compatibility has been integral to the process of phylogenetic systematics since its inception. Darwin (1859) realised the importance of character compatibility with statements such as “The importance, for classification, of trifling characters, mainly depends on their being correlated with several other characters of more or less importance” (:401-402). The first discussion of a true compatibility method

for estimating phylogeny was by Wilson (1965), who described the ideal phylogenetic character as one that "both uniquely defines a set of species and has not been reversed in evolution, so that all existing species which possess this state can be said to have descended from one species in the past that evolved the state". He described a test to identify "unique and unreversed" characters. In the same year the term "character compatibility" was first coined by Camin and Sokal (1965) who independently developed essentially the same concept as Wilson. Both of these methods simply compared the compatibility of individual characters with a tree topology, a method utilised in parsimony analyses. In 1969, Le Quesne defined the 'uniquely derived character' concept with the statement "if one is studying the taxonomy of a group, a character that has evolved only in one direction on a single occasion in its history is more likely to give an unambiguous indication of its phylogeny". Le Quesne (1969) discussed pairwise compatibility of characters, rather than looking at whether single characters can simply be mapped onto a specified tree without homoplasy. Two characters are pairwise compatible if it is possible to draw a tree representing a phylogenetic hypothesis upon which both characters can be mapped without homoplastic changes.

Compatibility methods are useful in attempting to separate phylogenetic signal from noise in a dataset. It is always hoped that in a data matrix the true phylogenetic signal is present in at least some of the characters. Other characters will represent, to varying degrees, nothing more than phylogenetically uninformative noise. Phylogeny reconstruction methods aim to pluck out the true signal from the noise in order to return the true phylogeny. Unlike some other methods, compatibility analysis involves no assumptions about the model of evolution. The only assumption necessary is the most basic assumption that there is a true tree for the taxa under study and that the ancestor-descendant relationships among these taxa may be represented as a tree (Estabrook, 1983). Even this assumption is not always true. Estabrook (1983) identified hybridisation and gene flow as potential pitfalls. In cases where a true phylogeny can be assumed, however, compatibility methods use patterns of compatibility and incompatibility in data to attempt to identify those characters representing the true phylogeny. True characters, as defined by Estabrook (Estabrook *et al.*, 1975; Estabrook, 1984), cannot conflict, and thus all uniquely derived characters must be mutually compatible (see Meacham, 1984). Noisy characters are randomly distributed among taxa, and are likely to be incompatible not only with the true signal, but also with each other. It is hoped that the true phylogenetic signal will be represented by a large block of compatible characters, while noisy characters will have a

generally low compatibility with the rest of the matrix. Large groups of compatible characters do not necessarily represent the true phylogeny, however. Compatibility can also be the result of logical dependence between characters, or functional, ecological or developmental correlation (e.g. see Meacham, 1984). A useful property of compatibility methods is that they are tree-independent, meaning they are an a priori assessment of the data rather than a measure of fit to a phylogenetic hypothesis. This means that their results are partly independent of any parsimony analysis carried out on the same data, so that their results will not necessarily agree. Although it seems that parsimony is generally quite successful at identifying phylogenetic patterns in data (e.g. Wiens and Hillis, 1996; but see Harcourt-Brown, 2002), compatibility, used in conjunction with parsimony can provide useful supplementary information. If the two methods do agree, support for the hypothesis they propose may be considered strong. However, if the two methods disagree, one or both methods must be producing erroneous results, suggesting that the data should be investigated further and homology assessments reviewed.

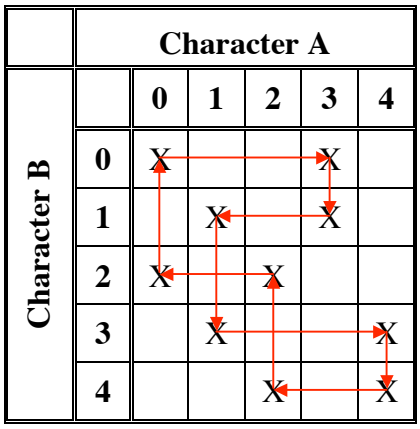
Algorithms for identifying incompatibility in raw data (in the form of taxon x character matrices) have been devised, although some compatibility-based computer programs are useful only with binary data. This is because identifying incompatibility between binary characters is simple. Le Quesne (1969; 1972) observed that, given two binary characters, A and B, each with two character states, 0 and 1, only four combinations of character states are possible (A_0, B_0; A_0, B_1; A_1, B_0 and A_1, B_1). If all four of these state combinations are present in the taxa selected for analysis, then the characters cannot be mapped onto the same tree without at least some homoplasy. So, at least one (or possibly both) of the characters is not uniquely derived. It must be stressed, however, that if fewer than four of the state combinations are present, the two characters are compatible, but neither is necessarily uniquely derived. Identifying incompatibility between multistate characters is slightly more complex. Farris (1973a) introduced a method by which multistate characters could be split into a number of binary characters, allowing the application of Le Quesne's methods (Estabrook *et al.*, 1976; Le Quesne, 1982). However, the simplest method for finding incompatibility in multistate characters, devised by both Estabrook and Landrum (1975) and Fitch (1975) and mathematically proven by McMorris (1975; Estabrook and McMorris, 1977), involves the creation of a character state by character state matrix for the two characters in question (see Estabrook, 1983 for a simplified description). In these matrices, henceforth called state combination matrices, the states of the two characters in question are plotted on the two axes. Cells in the matrix are

marked (here with an X) if the corresponding character state combination is present in at least one taxon. For example, if one or more taxa possess state 0 for one character (character A) and state 1 for another (character B) (both characters having three states), then the following matrix is produced and the cell corresponding to the state combination A_0, B_1 is marked with an X.

		Character A		
		0	1	2
Character B	0			
	1	X		
	2			

The state combinations of all taxa included in the analysis are inserted into the state combination matrix in this way. If at least four of the Xs in the matrix can be joined by horizontal and vertical lines to form the corners of at least one continuous loop (i.e. a path of horizontal and vertical lines that can be followed between Xs and returns to a cell that has previously been visited without reversing direction), then the characters are incompatible. It is not necessary for all Xs to be part of the loop, but all corners of a loop must coincide with an X. Loops may cross themselves in cells without Xs, but these are not regarded as corners. The following are two examples of pairs of incompatible characters with loops present indicated in red.

		Character A				
		0	1	2	3	4
Character B	0				X	
	1			X		X
	2	X				
	3		X			
	4			X		X



Most compatibility measures employ a strict cut-off, so that a pair or group of characters are simply classified as compatible or incompatible. This limits the power of current compatibility tests. Estabrook *et al*. (1975) defined a 'true cladistic character' as a divergent character that is a partial estimate of cladistic history and should meet the following three criteria:

1) A character state should contain its own most recent common ancestor.
2) If one taxon is the ancestor of a second, then the state of which the first is a member must be equal, or should be ancestral in the character state tree (a 'tree' showing the proposed ordering of character states), to the state of which the second is a member.
3) If one character state is ancestral to another in the character state tree, then the most recent common ancestor for the one state should be ancestral to the most recent common ancestor of the other.

As reiterated by Meacham (1984), 'true cladistic characters', as defined by Estabrook *et al*. (1975), cannot conflict with one another, and therefore all true cladistic characters must be mutually compatible. This logic cannot be faulted, but it might be taken to suggest that any characters that conflict with the true signal are not useful in phylogenetic reconstruction. On the contrary, many such are useful indicators of phylogeny. It appears that it is not uncommon for homoplasy (in the form of convergences and/or reversals) to occur in characters that provide phylogenetically useful (and important) information. Such characters, although providing some misleading evidence, are still desirable, because they may provide evidence of relationships that are not identified by other characters. Parsimony analyses aim to use evidence supplied by additional characters to reject homoplastic relationships, so that any useful information in a character is still utilised (e.g. see Farris *et al*., 1996; Källerajö *et al*., 1999). Most methods of compatibility analysis, however, will show such characters as incompatible with many

'true cladistic characters' because of their homoplastic change(s). The hope in such analyses is that if a character possesses few homoplasies, it will still be compatible with many other characters, including some 'true cladistic characters', whereas characters that are no more than noise are unlikely to be compatible with any other characters. However, in practice it is possible that many characters that contain useful phylogenetic information will be classified as 'poor' and will not be distinguished from noise.

In order to attempt to further separate noise from homoplastic characters that contain phylogenetically useful information, it may be useful to quantify how incompatible two incompatible characters are. One possible measure is fuzzy compatibility, described for the first time here.

4.2.1 Fuzzy compatibility

Fuzzy compatibility attempts to quantify the level of incompatibility between two or more characters. It is a measure of the minimum number of taxa that must be rescored in order to make the characters compatible (or the minimum number of taxa that must have been misscored in creating the original data matrix). Guise, Peacock and Gleaves (1982) described a method based on a similar idea that they called "labelling". Their method identified incompatible pairwise binary character comparisons in which only one taxon possessed one of the four possible state combinations. They suspected that the primary homology assessment of one of the characters for that taxon was wrong. All such potentially homoplastic scorings were recorded and at the end of the analysis the particular score most likely to be homoplastic could be identified. Fuzzy compatibility differs from this method in that it calculates the minimum number of taxa that must be rescored to make all incompatible character pairs compatible, not just those in which a single taxon must be rescored. A similar method, called the minflip algorithm (Chen et al., 2003), was devised for supertree construction. Minflip uses heuristic algorithms to flip states in order to find the smallest number of states in the matrix that must be changed in order to make the entire matrix compatible.

In the case of two binary characters, calculating fuzzy compatibility is simple. As stated above, binary characters can only be incompatible if at least one taxon possesses each of the four possible character state combinations for the two characters. Therefore, if the two characters are incompatible, the minimum number of taxa that must be rescored to make the characters compatible is equal to the number of taxa scored as possessing the least common character state combination.

With multistate characters, finding the smallest number of taxa that must be rescored to make the characters compatible is more complicated. Again, the simplest method involves the use of a character state combination matrix as described above. However, in this case, instead of simply marking any state combinations possessed by at least one taxon with an X in the matrix, the number of taxa possessing each state combination is inserted into the cell. For example, consider the following two characters, each of which exhibits four states:

	Char A	Char B
Taxon 1	0	0
Taxon 2	0	0
Taxon 3	0	0
Taxon 4	0	0
Taxon 5	0	1
Taxon 6	1	1
Taxon 7	1	1
Taxon 8	1	2
Taxon 9	2	2
Taxon 10	2	2
Taxon 11	2	3
Taxon 12	3	3
Taxon 13	3	1
Taxon 14	3	3

This combination of character states gives rise to the following state combination matrix:

		Character A			
		0	**1**	**2**	**3**
Character B	**0**	4	0	0	0
	1	1	2	0	1
	2	0	1	2	0
	3	0	0	1	2

In this example, cells containing non-zero values form the corners of a single continuous loop, indicating that the characters are incompatible. Two cells, A_0, B_0 and A_0, B_1 are not integral parts of any loop, and therefore are not the cause of the incompatibility in the data. The taxa possessing these state combinations can be disregarded when identifying the minimum number of taxa that need to be rescored to make the characters

compatible. For the characters to be made compatible, the loop must be broken. It is a rule that **any single loop will be broken by removing any one of its corners**. **Which corner is removed is not important**. Therefore, **the minimum number of taxa that must be rescored to make two characters compatible is equal to the smallest number of taxa present in a cell at the corner of the loop**. In the above example, there are three equally minimal ways to make the characters compatible. State combinations A_1, B_2; A_2, B_3 and A_3, B_1 are all corners of the loop, and are all present in only one taxon. By referring back to the data matrix, it can be shown that taxon 8 (A_1, B_2), taxon 11 (A_2, B_3) or taxon 13 (A_3, B_1) can be rescored to make the two characters compatible. Therefore, the minimum incompatibility value for these two characters is one.

It is possible for multistate characters to produce state combination matrices that contain more than one loop. For example:

		Character 1			
Character 2		**0**	**1**	**2**	**3**
	0	3	1	0	0
	1	2	1	0	1
	2	0	1	2	0
	3	0	0	1	2

In this example, there are two loops that must be broken to make the characters compatible, shown in red and green. In order to break two loops, the smallest number of corners that must be removed is two. Methodologically, this can be achieved by first removing the corner of either loop that corresponds to the lowest number of taxa (i.e. in this example, one taxon must be rescored, either at position A_1, B_0; A_1, B_1; A_1, B_2; A_2, B_3 or A_3, B_1). When one corner is removed, the loop in which it was a corner is broken, and the remaining corners of that loop can be disregarded, provided they are not also corners of further loops. In the above example, if the taxon possessing state combination A_1, B_0 is rescored, the taxa at A_0, B_0 and A_0, B_1 can be disregarded. The taxon at A_1, B_1, however, still forms the corner of the second loop, and therefore cannot be disregarded.

		Character 1			
		0	**1**	**2**	**3**
Character 2	**0**	3	-	0	0
	1	2	1	0	1
	2	0	1	2	0
	3	0	0	1	2

Removal of any corner when two loops are present will always leave a single loop, no matter which corner is chosen. **It is impossible to break two loops by removing only one corner**. In the above example, position A_1, B_1 is a corner of both loops. It therefore may seem that recoding the taxa possessing this state would break both loops simultaneously. However, as shown below, removing this 'corner' simply creates a single loop that crosses itself at the position vacated by the recoded taxon.

		Character 1			
		0	**1**	**2**	**3**
Character 2	**0**	3	1	0	0
	1	2	-	0	1
	2	0	1	2	0
	3	0	0	1	2

Once the first loop is broken, the process can be repeated to break the second loop. Again, the minimum number of taxa that must be rescored to break this loop is one (at position A_1, B_1; A_1, B_2; A_2, B_3 or A_3, B_1). Therefore, the minimum number of taxa that must be rescored to make these characters compatible is two, one to break each of the two loops. This process can be repeated to break any number of loops. **One corner must be recoded for each loop present in a state combination matrix in order to make the two characters compatible**.

4.2.2 Uses of Compatibility

Simply measuring the number of pairwise incompatibilities for each character in a dataset is generally not considered sufficient for most purposes. Often, incompatibility values are used to identify the relative 'strength' of characters in a dataset. This necessitates comparisons of incompatibility values between characters. Simply counting the number of pairwise incompatibilities for a character in a dataset has been recognized as

being problematic for such comparisons (Le Quesne, 1969; 1972). Some characters are more likely than others to be compatible by chance alone, so that they may misleadingly appear 'stronger' than those other characters when in truth they are not. It is possible to calculate the probability of two binary characters with no missing data being incompatible by chance alone using the following formula derived from that published by LeQuesne (1972; see also Meacham, 1984).

$$P = 1 - \frac{n_0! \times (n_T - n_s)!}{n_T! \times (n_0 - n_s)!} - \frac{n_1! \times (n_T - n_s)!}{n_T! \times (n_1 - n_s)!},$$

where 0 and 1 are the two states of the two binary characters, A and B; n_s is the state assigned to the smallest number of taxa (i.e. the smallest number of 0s or 1s in character A or B); n_0 and n_1 are the number of states 0 and 1 in the character not including n_s; and n_T is the total number of taxa.

For example, consider two binary characters scored for ten taxa with no missing data. If both characters have eight taxa scored as state 0 and two taxa scored as state 1, then the probability of their being incompatible by chance alone can be calculated using these values with the formula: $n_s = 2$; $n_0 = 8$; $n_1 = 2$ and $n_T = 10$.

$$P = 1 - \frac{2! \times (10 - 2)!}{10! \times (2 - 2)!} - \frac{8! \times (10 - 2)!}{10! \times (8 - 2)!}$$

$$\therefore P = 1 - \frac{80640}{3628800} - \frac{1625702400}{2612736000}$$

$$\therefore P = 1 - \frac{1}{45} - \frac{28}{45} = \frac{16}{45} = 0.3555.$$

Similarly, if one character has five 0s and five 1s and the other eight 0s and two 1s, then the probability of these two characters being incompatible by chance alone is 0.5555, and if both characters have five 0s and five 1s, the probability of their being incompatible by chance alone is 0.9921. This shows clearly that the characters most likely to be incompatible by chance alone are those with more equal numbers of the two states, and that those with only two taxa possessing one of the states are the least likely to be incompatible by chance alone. This may cause problems when comparing compatibility scores between characters, because characters with an unequal number of each state are more likely to be compatible with other characters by chance alone. Therefore, random, noisy, phylogenetically uninformative characters with unequal numbers of each state are more likely to be wrongly considered 'strong' characters. A number of methods have been devised in an attempt to counteract this problem. Some of these methods are discussed below.

4.2.2.1 The Coefficient of Character State Randomness (CCSR)

The CCSR (Le Quesne, 1969; 1972; 1982) is a measure that attempts to correct the relative 'strength' of characters by taking into account the probability of such a character being compatible by chance alone. The CCSR is simply a ratio of the observed number of pairwise incompatibilities of a character with the rest of a matrix divided by the number of pairwise incompatibilities expected by chance alone for a randomly permuted version of that character. For binary characters it is relatively simple to calculate the expected value mathematically (see above). However, for more complex, multistate characters the exact expected number of incompatibilities is more difficult to calculate, and is therefore often approximated using the average number of incompatibilities of a large number of random permutations of the character. The character must be permuted as opposed to randomised so that the numbers of individual character states in the original character are preserved. In cases where some taxa cannot be scored for a character, it is recommended (Wilkinson, 2001a) that missing data should be held constant during permutation to reduce the number of variables in the permutation test. The CCSR is calculated for each character in the data matrix, and a character-by-character matrix can be constructed containing the pairwise CCSR values (Le Quesne, 1969). The total CCSR value for each character can then be calculated by simply summing the values for all pairwise character comparisons containing that character. CCSR analyses can be carried out using either the normal or fuzzy compatibility methods of measuring pairwise incompatibilities between characters.

4.2.2.2 The Normal Deviate (NDev)

The NDev (Le Quesne, 1972; 1982) is similar to the CCSR, but has the advantage of showing whether the difference between the observed and expected number of incompatibilities for a character is statistically significant. It is a measure of where the observed number of incompatibilities falls in terms of the number of standard deviations away from the mean of the distribution of expected values. If the value is positive, then the observed number is lower than that expected by chance, and vice versa. The NDev for binary characters can be calculated using the following formula:

$$\mathrm{NDev} = \frac{P_s - N_x - 0.5}{\sqrt{\left(\frac{P_s(n_v - P_s)}{n_v}\right)}},$$

where P_s is the sum of all values of P (the probability of two characters being incompatible by chance alone) for comparisons of the character with all other characters; n_v is the

number of valid comparisons (i.e. comparisons with characters that are not phylogenetically uninformative); and N_x is the observed number of character pairings in which all four character state combinations are found. The sign of the last term in the numerator becomes '+' if $P_s < N_x$ (Le Quesne, 1989). This is a modification (Yates' correction) of the formula presented by Le Quesne (1972), which corrects for problems associated with analysing small datasets.

4.2.2.3 Le Quesne Probability (LQP)

The LQP measure was devised independently by both Wilkinson (1992a; 2001a; 1997a) and Meacham (1994), who called it the Frequency of Compatibility Attainment. Here the name LQP is employed because it was proposed first. LQP is a simple randomisation method for testing the null hypothesis that a character is no more compatible with other characters in a data matrix than is a random, phylogenetically uninformative character. The method involves first calculating the number of pairwise incompatibilities in the dataset for each character in its original form, and then doing the same for a number (usually 99 or 999) of random permutations of that character. The LQP is the probability of a randomly permuted character having an equal or lower number of total incompatibilities (the sum of all its pairwise incompatibilities) than the original character. It is calculated by simply counting the number of random permutations of a character that have the same or fewer incompatibilities with the dataset than the original character. This value is then divided by the total number of replicates (number of permutations + 1) to find a p-value between 0 and 1 with which to test the null hypothesis. Characters that are highly compatible with the matrix will have an LQP value close to 0, whereas characters representing random noise should have an LQP value closer to 0.5. The LQP is advantageous over other compatibility measures, such as the CCSR, because, like the normal deviate, it gives a measure of the level at which a character is significantly better (or worse) than random noise. It is especially useful because it sets up a significance test for which the significance level can be defined by the user. As with most randomisation tests, the cut-off between significance and lack of significance is usually taken at the 5% level. The LQP test is a one-tailed test, because its aim is to find characters better than random, not just different from random. This means that any LQP values below 0.05 indicate that a character is significantly more compatible than random at the 5% level.

4.2.2.4 Clique Analysis

A collection of mutually compatible characters that is not a subset of a larger collection is termed a clique (Estabrook *et al.*, 1976; 1977) from the equivalent graph theory concept. Estabrook *et al*. (1976) provided proof that any collection of pairwise compatible, ordered binary characters are mutually compatible. However, this concept does not hold for unordered multistate characters (Fitch, 1975; 1977) or characters that include missing entries (Wilkinson, 1994c) or uncertainties. Wilkinson (1994b) specified the definition of clique analysis as "the discovery of cliques" in order to distinguish it from other compatibility methods, with which the term had often previously been synonymized. The original hypothesis behind clique analysis is that because all 'true' characters must belong to the same collection of mutually compatible characters, then the largest such collection contains them (Estabrook *et al.*, 1977). In other words, the largest clique is assumed to contain the 'true' characters. Estabrook *et al*. (1977) proposed simply identifying the largest clique and using the tree that it supports as the best hypothesis of phylogeny. Unfortunately, as noted by Wilkinson (1994b), largest cliques are often too small and the characters they contain too similar to resolve large sections of the tree (see also Felsenstein, 1982). They can also produce phylogenetic hypotheses that are inconsistent with the results of maximum parsimony methods (Wilkinson, 1994b). However, Wilkinson (1994c) showed that clique analysis and parsimony analysis are equivalent when a dataset does not include any characters with a maximum length greater than two, a situation which can be attained by use of the three-taxon statement method of character representation (Nelson and Platnick, 1991). A number of attempts have been made to devise methods using clique analysis as a starting point to find maximally parsimonious trees for a dataset without the need for conventional, time-consuming parsimony analysis (e.g. Penny, 1982; Lorenzen, 1993). Unfortunately, most of these methods appear to be impractical when large numbers of characters are present in the data, where conventional parsimony analyses are far more efficient (Wilkinson, 1994b).

4.2.2.5 Boildown

Le Quesne (1969; 1972) introduced a number of methods by which noisy characters can be eliminated from a data matrix in an attempt to increase signal. These methods included:

- Eliminating the character or characters with the highest incompatibility count.

- Assuming that the character with the lowest incompatibility count, lowest CCSR or highest positive NDev value is uniquely derived and eliminating characters incompatible with that character.
- Assuming the largest group of completely correlated characters (the largest clique) is uniquely derived and eliminating characters incompatible with it.
- Eliminating all characters with an NDev value below +2 or +3.

Whichever method is employed, once the 'worst' character(s) are identified and eliminated, the procedure can be repeated until all incompatibility has been removed from the data (Le Quesne, 1969; 1972). Gauld and Underwood (1986) first introduced the term 'boil down' for an analogous procedure that they used to identify groups of mutually compatible characters. They calculated CCSR values (which they alternatively called randomness ratios), deleted the character with the worst value (highest CCSR) and then recalculated. This procedure is repeated until no incompatibility remains. All characters left in the dataset are pairwise compatible, and, if all are binary and contain no missing data, mutually compatible, although not necessarily an entire clique. These methods are effectively tree-independent methods of extreme character weighting, in which characters deemed the 'worst' during each repetition are assigned a weight of zero. Many compatibility measures can be used to select the characters for removal at each stage of a boildown procedure, including the CCSR, NDev and LQP. Fuzzy compatibility can also be used in a boildown procedure.

As previously intimated, a boildown procedure is usually halted when no incompatibility remains in the matrix. However, the remaining clique may not be sufficient to resolve the phylogeny of the taxa under study (Felsenstein, 1982; Gauld and Underwood, 1986). Gauld and Underwood (1986) suggested ranking characters not in the maximum clique based on their compatibility with it, and using these to resolve the rest of the tree. As an alternative to using methods for adding characters to a maximum clique to improve resolution (see also Penny, 1982; Lorenzen, 1993), it may be preferable to halt the boildown before all incompatibility is removed, in order to preserve signal that may be present in characters that show low incompatibility with the matrix. This is effectively an attempt to remove noise without removing too much useful phylogenetic signal. One possible way of doing this is by using the LQP as the measure of character strength for the boildown. The boildown procedure can then be stopped when all characters remaining in

the matrix are significantly more compatible with each other than is a random permutation of that character (that is, when the null hypothesis for the LQP test can be rejected for all characters remaining in the matrix).

Alternatively, the effects of every step of the boildown process can be monitored to study how character elimination affects the tree produced by analysis of the data. Assuming that the true phylogeny of the taxa is represented by a large number of relatively compatible characters and that noisy characters are relatively incompatible, the signals provided by noisy characters should be in conflict and should not produce a combined signal strong enough to overpower the true tree. If this is true, by removing incompatible, noisy characters using the boildown, the resulting phylogeny should not change significantly other than possibly to lose resolution in areas of the tree in which initial resolution was provided solely by poor characters. If, however, there is a large amount of change in the topology through the boildown process, this suggests that the characters being removed by the boildown are playing a major role in the topology of the tree produced by parsimony analysis of the complete data. This could be due to the presence of two or more competing signals in the data, possibly caused by convergences due to functional similarities between unrelated taxa. In such cases the original data might be re-examined and hypotheses of homology checked.

Other methods similar to the boildown, by which characters are reweighted based on compatibility values such as the CCSR, have been presented by Penny and Hendy (1985; 1986) and Sharkey (1989). Sharkey (1994) proposed an improved method of character reweighting based on what he termed discriminate compatibility, and described a reduction routine similar to the boildown, which could be used to build trees. His discriminate compatibility method relied on a new procedure for determining the polarity of characters. However, as noted by Wilkinson (1997b), this method is feasible only with a dataset containing a small number of characters. When used to polarise binary characters there are 2^n possible polarisations to consider, where n is the number of characters. Therefore, with a moderately sized matrix of just 100 characters, 2^{100} (= more than 10^{30}) alternative polarisations must be examined, rendering the process impractical even for powerful computers (Wilkinson, 1997b).

4.2.2.6 Boildown Bootstrap

Bootstrap analyses (Felsenstein, 1985) are widely used to investigate the strength of support for clades in a resulting phylogeny. Felsenstein's (1985) bootstrap involves

producing a number of datasets of identical proportions (the same number of characters and taxa) to the original data using a resampling with replacement strategy, in which characters to be included in the bootstrap replicate dataset are selected randomly from the original data. When a character is added to the new dataset it is not excluded from being selected again, so that some characters in the original data can be in the bootstrap replicate dataset more than once, and others not at all. Each bootstrap replicate is then analysed using parsimony, and the MPTs of each replicate are saved. A majority-rule consensus tree is then produced from the MPTs of the replicates, in which the majority-rule values on each node of the tree represent the bootstrap support for that clade.

Here, the same methodology is, for the first time, applied to the boildown procedure in an attempt to answer the unanswered question of when to halt a boildown procedure. This method, which measures the support for clades present in the trees produced at each stage of the process, is named the boildown bootstrap. Two types of boildown bootstrap can be performed.

The first type, called type 1 boildown bootstrap simply involves carrying out bootstrap analyses on the reduced datasets produced at each stage of a boildown procedure. The bootstrap trees produced should include the best-supported compatible groups (in terms of compatibility with the 50% majority rule bootstrap tree) present in less than 50% of the saved trees. This produces one bootstrap tree for each boildown step.

The type 2 boildown bootstrap, is more complex, more labour intensive and far more time consuming. First, a number of bootstrap replicate data matrices of the entire data matrix to be analysed are produced using resampling with replacement. A boildown is then carried out on each of these datasets. After each character removal during these boildowns a parsimony analysis is carried out on the reduced dataset and the resulting MPTs saved. Therefore, at the end of all boildown analyses a set of MPTs for each replicate has been saved for each character removal step. For example, in the first boildown, when one character is removed, the MPTs of the analysis of this reduced dataset are saved to a treefile. In subsequent boildowns, the MPTs of analyses carried out after one character has been removed are appended to this treefile. Each set of MPTs in the tree file is then assigned a weight inverse to the number of trees that it contains. This ensures that there is not a bias towards sets of MPTs containing more trees when a consensus of all sets is taken. For example, each tree in a set of 20 MPTs is assigned a weight of 1/20, while trees in a set of five MPTs are assigned a weight of 1/5 each. Finally, a majority-rule consensus tree with compatible resolutions under 50% included is calculated for each of the treefiles,

so that there is one majority-rule tree for every character removal stage of the boildown. These majority-rule trees are equivalent to bootstrap trees, with the majority-rule values on nodes equivalent to bootstrap values.

Whichever type of boildown bootstrap is used, the point during the boildown at which the tree produced by parsimony is best supported can be assessed. By summing the majority-rule values of all nodes on each bootstrap tree (= majority-rule consensus tree for type 2 boildown bootstraps), a measure of the strength of support for that tree is obtained. This is essentially a measure of the total bootstrap value for the tree. If the bootstrap tree with the highest total bootstrap value is not produced by analysis of the complete data, it can be suggested that up to this point in the process the boildown method has removed noise from the data without sacrificing signal.

4.3 Potential Problems with Compatibility Methods

4.3 1 Uncertainty and Polymorphism Within Leaves

All current compatibility methods suffer from an inability to cope with leaves containing uncertainty (leaves for which the state present for a character is uncertain between two or more possibilities) or polymorphism (or leaves in which taxa are present containing two or more states of a character). These two types of intra-leaf variation cause problems in compatibility analysis for two reasons.

The problem regarding the treatment of uncertainty lies solely in the efficiency of any compatibility analysis program. In theory it is possible to run a compatibility analysis in which taxa coded as uncertain for two or more states can be included in all possible configurations to see which is most compatible and which most incompatible. From this, a range of compatibilities for each character can be recorded. With only a few uncertain taxa this is not a problem, but when a larger number of taxa are coded as uncertain for a number of characters, the number of permutations of possible character states that must be analysed increases exponentially. This causes problems for the efficiency of the analysis, leading to extremely long calculation times for even relatively small datasets. Currently no compatibility programs treat uncertainties in this way. Instead, as an approximation, they generally replace taxa coded with uncertainties with missing data (?), which is equivalent to uncertainty between all states. Although this situation is not perfect, treating uncertainty as missing data should, in most cases, lead to analytical results very similar to those that would be found if uncertainties could be treated individually.

A more difficult problem to solve is that of intra-leaf polymorphisms. These represent instances where the taxa included as a leaf in an analysis possess more than one of the states of a character. This is a major headache in compatibility analysis, because it is possible that a single character containing polymorphic leaves can essentially be incompatible with itself, and therefore in any pairwise compatibility test would seem incompatible with every other character. If a leaf is scored as possessing two states of one character and those states are also both present in one or more other taxa, it is not possible to draw a tree on which the character can be mapped without homoplasy. This phenomenon is here called intra-character incompatibility, and must be the result of errors in homology assessment, errors in the determination of the constituent taxa of leaves or of true homoplasy in the character. There are two ways in which polymorphism within a leaf can occur. The first is a situation where two or more states of a character are identified within a single individual. Such characters fail Patterson's (1982) conjunction criterion of homology assessment and should not be included in analyses (see Chapter 1). The second is a situation where a leaf is made up of more than one taxon, and these groups vary in their scoring for an individual character. By creating a leaf that contains more than one taxon, authors are effectively proposing a hypothesis of relationships that they think is strong enough to assume without further testing. This is necessary when analysing relationships of basal groups, because otherwise an analysis would become too large for current methods to compute. For example, in order to evaluate the relationships within the amniotes it is not feasible to include every known amniote in a phylogenetic analysis. Decisions must be made a priori about which taxa to include as individual leaves and which to group together. Such decisions are based on knowledge gained from previous analyses, so that well-justified clades containing many taxa may be used in more inclusive analyses. In such circumstances, authors are faced with a choice of how to code these multi-taxa leaves. They can be coded either on the basis of a groundplan version of the taxa included in the leaf or using one or more exemplars from the group (see Yeates, 1995). It is only when more than one exemplar is used to code these leaves that polymorphisms can result. If taxa are incorrectly grouped, so that unrelated taxa are forced together into one leaf, then it is not surprising that homoplasy results. A second possible explanation for intra-character homoplasy in characters containing polymorphisms is error in homology assessment. Patterson (1982), in his discussions of primary and secondary homology, described the conjunction criterion, which must be passed in order for the states of a character to be considered homologous. This criterion is that two states of one

character cannot be present in a single taxon. By this, Patterson was suggesting that two states of a character cannot be homologous if a specimen is found possessing both states. Although in the case of polymorphism the two states are not always possessed by a single specimen, the problems caused are similar. The possession of more than one state by a leaf (whether that leaf is a single specimen or a group of taxa) usually leads to homoplasy in that character, so that homology of the states cannot be concluded. The only exception to this occurs when at least all but one of the character states in the polymorphism are unique in that taxon.

At present, in most compatibility analyses, polymorphic taxa are rescored as missing data. This is not a satisfactory method of coping with the problem, because it discounts homoplasy within characters containing polymorphisms, and artificially increases their 'strength'. In extreme cases, where many leaves are polymorphic, a character could contain large amounts of intra-character homoplasy, yet seem compatible with all other characters when these polymorphisms are replaced with missing data. Unfortunately, the alternative is to treat characters containing intra-character homoplasy as incompatible with all other characters. In the case of fuzzy compatibility analyses, a possible alternative is to replace polymorphisms with missing data, but increase the fuzzy incompatibility value to acknowledge the weakness of the character. For example, characters containing intra-character homoplasy could have 1, or the minimum number of recodings necessary to make the character compatible with itself added to each of their pairwise fuzzy incompatibility scores.

4.3.2 Logical linkage and Inapplicable data

Characters that are completely logically linked cannot be incompatible. This can cause problems in compatibility analyses, because, as in parsimony analyses, the linkage of these characters effectively increases the weights given to any hypotheses of relationship they support. In compatibility analyses, any group of completely linked characters will be mutually compatible, which may artificially increase their own compatibility values and decrease those of 'true' characters.

In Chapter 2, I discussed the problem in character construction of characters relating to data that are inapplicable to some taxa under study. With the 'missing' method of coding inapplicable data, which is the method I suggested was preferable for parsimony analyses, there is usually one character describing the presence or absence of a structure and one or more characters describing attributes of that structure in taxa that possess it.

Taxa that do not possess the structure are scored as unknown for these characters. These characters cannot be incompatible with the character describing presence or absence, leading to problems in compatibility analysis similar to those of logically linked characters. Therefore, in compatibility analysis it may be better to include inapplicable data characters using the 'multistate' coding method (see Chapter 2).

Gauld and Underwood (1986) suggested another possible solution. They dealt with multistate characters in their data by splitting them into additive binary characters and labelled each of these additive binary characters so that they would not be compared with one another in analyses. Such a method could also be applied to inapplicable data. However, although their method does stop linked characters upweighting each other's compatibility values, it does not necessarily correct for the effect a large number of additive binary characters would have on other characters in the data. If a large block of linked characters is present in the data, other characters in the data that are compatible with the original multistate character from which the additive binary characters were produced will show an artificially high compatibility relative to a character that is incompatible with the original multistate character. There are, however methods that could potentially negate this bias, for example, by weighting the results accordingly.

4.3.3 Tree Balance

A recurring problem in many phylogeny reconstruction methods is that of a bias towards unbalanced tree topologies. It has been shown in empirical (Mooers *et al.*, 1995) and simulation studies (Huelsenbeck and Kirkpatrick, 1996; Harcourt-Brown, 2002) that parsimony and other reconstruction methods favour unbalanced trees. Guyer and Slowinski (1991; 1993) provided an explanation for this bias in parsimony methods using the "equiprobable model" (that all tree topologies are equally probable). Under this model, they showed that there are more arrangements of taxa on unbalanced tree topologies, effectively meaning that there are more possible unbalanced topologies. Huelsenbeck and Kirkpatrick (1996) provided a simple example. Consider a tree containing eight taxa. There are 40,320 ($8!$) ways that the taxa can be permuted onto (the leaves of) the tree. However, more of these ways of placing taxa onto the leaves are isomorphic (identical topologies, but with nodes rotated) on symmetrical trees. In fact, with 8 taxa, there are 20,160 ($\frac{n!}{2}$, where n is the number of taxa, as long as it is possible to create a completely balanced tree from that number of taxa) different rooted maximally unbalanced trees,

whereas there are only 315 ($\frac{n!}{2^{(n-1)}}$) different maximally balanced trees. With so many more possible unbalanced topologies (64 times as many with 8 taxa), it is highly probable that a tree chosen at random will be unbalanced. This means that any noisy signal that has an effect on the resolution of the most parsimonious tree is likely to introduce more imbalance to that tree.

In recent years, a major emphasis has been placed on attempting to create a tree of life in which all known taxa can be placed and their affinities known. To do this using traditional phylogenetic analysis methods is not currently feasible due to the immense computing time required, so alternative methods have been sought. One solution is to use supertree methods to combine results of smaller analyses. These methods take the consensus trees output by analyses (the source trees) and attempt to join them together, as long as they have at least three (or two plus the root in rooted trees) taxa in common. The most commonly implemented supertree method is matrix representation using parsimony (MRP) (Baum, 1992; Ragan, 1992). This, like many other supertree methods, joins trees together by first representing them as a taxon by 'pseudocharacter' matrix representation (MR) (Farris, 1973a). Each node of each source tree is coded as a single 'pseudocharacter' where taxa within the clade defined by the node are scored as 1 and those outside as 0. Any taxa in the matrix that are not present in the source tree are scored as unknown (?) for all 'pseudocharacters' based on that tree. MRP then employs parsimony to create one or more supertree(s) from the data matrix. However, the MRP method has been much criticised (Purvis, 1995; Wilkinson *et al.*, 2001; Goloboff and Pol, 2002). Among other problems, it suffers from two major biases that make its results unreliable at best. These biases are: that it produces output trees more similar to large source trees (those with more taxa) (Purvis, 1995), and to unbalanced source trees (Wilkinson *et al.*, 2001). The bias towards large source trees is simply due to the larger number of potential nodes that are present in trees with a larger number of taxa. This leads to more 'pseudocharacters' in the MR of large trees, and therefore adds weight to the hypotheses supported by large source trees. The bias towards unbiased trees was illustrated nicely by Wilkinson *et al.* (in prep). They showed that, even using the same taxon list in the source trees, if a completely balanced and a highly incongruent completely unbalanced source tree are submitted to MRP, then the output supertree is more similar to the unbalanced source tree. The reason for this bias is not so simple to understand as the bias towards large source trees. Thorley and Wilkinson (2003) noted that one major problem with the MR method is that the tree-to-supertree

distance that is the basis of the methods objective function does not obey the symmetry axiom. This means that the distance between each source tree and the supertree produced from them is not always the same. In Wilkinson *et al.*'s (in prep) example, if the pseudocharacters produced by MR of the balanced tree are mapped onto the unbalanced tree, the fit (in parsimony steps) is not the same as if the characters produced by the unbalanced tree are mapped onto the balanced tree. Wilkinson *et al.* (in prep) suggested that this asymmetric distance measure is responsible for the bias of MRP towards producing unbalanced supertrees

One artefact of the MR method is that more balanced source trees lead to MRs with a high ratio of primitive states to derived states, whereas unbalanced trees lead to an equal number. This is because balanced trees are formed by bifurcating nodes so that each node contains half the number of taxa contained by the previous node in the tree (the node further down the tree). Unbalanced trees are formed in a hierarchical way, where each node simply contains one taxon less than the previous node. For example, take the two source trees used in the example of Wilkinson *et al.* (in prep) (Fig 4.1a and b). The matrix representations of these trees are shown in tables 4.1 and 4.2.

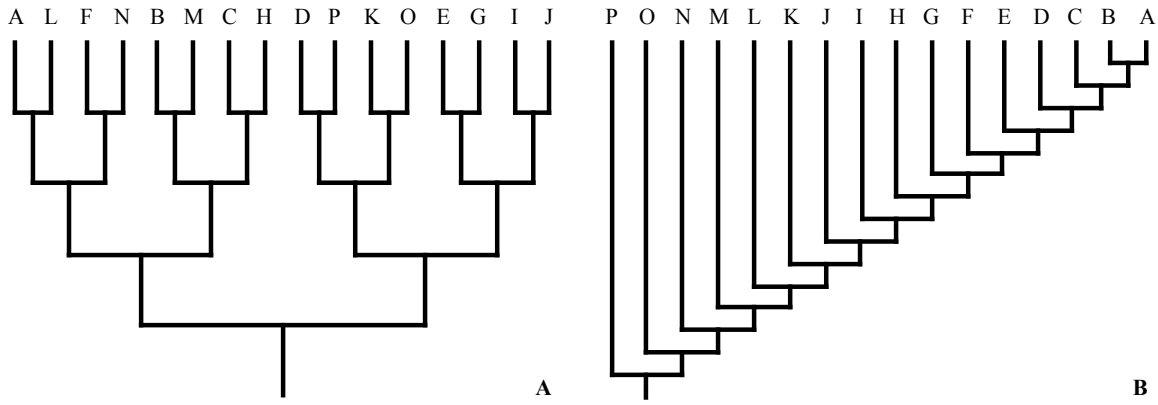


Figure 4.1. The two incongruent source trees used by Wilkinson *et al.* (in prep.). to illustrate the problem of a bias towards unbalanced tree topologies in MRP analysis. A) A completely balanced tree. B) A completely unbalanced tree.

Taxon	1	2	3	4	5	6	7	8	9	10	11	12	13
A	1	1	1	0	0	0	0	0	0	0	0	0	0
B	1	0	0	0	1	1	0	0	0	0	0	0	0
C	1	0	0	0	1	0	1	0	0	0	0	0	0
D	0	0	0	0	0	0	0	1	1	0	0	0	0
E	0	0	0	0	0	0	0	0	0	0	1	1	0
F	1	1	0	1	0	0	0	0	0	0	0	0	0
G	0	0	0	0	0	0	0	0	0	0	1	1	0
H	1	0	0	0	1	0	1	0	0	0	0	0	0
I	0	0	0	0	0	0	0	0	0	0	1	0	1
J	0	0	0	0	0	0	0	0	0	0	1	0	1
K	0	0	0	0	0	0	0	1	0	1	0	0	0
L	1	1	1	0	0	0	0	0	0	0	0	0	0
M	1	0	0	0	1	1	0	0	0	0	0	0	0
N	1	1	0	1	0	0	0	0	0	0	0	0	0
O	0	0	0	0	0	0	0	1	0	1	0	0	0
P	0	0	0	0	0	0	0	1	1	0	0	0	0

Table 4.1. Matrix representation of a maximally balanced tree containing 16 taxa (tree A in Fig. 4.1). Matrix representation contains 160 (77%) primitive states (0s) and 48 (23%) derived states (1s).

Taxon	1	2	3	4	5	6	7	8	9	10	11	12	13
A	1	1	1	1	1	1	1	1	1	1	1	1	1
B	1	1	1	1	1	1	1	1	1	1	1	1	1
C	0	1	1	1	1	1	1	1	1	1	1	1	1
D	0	0	1	1	1	1	1	1	1	1	1	1	1
E	0	0	0	1	1	1	1	1	1	1	1	1	1
F	0	0	0	0	1	1	1	1	1	1	1	1	1
G	0	0	0	0	0	1	1	1	1	1	1	1	1
H	0	0	0	0	0	0	1	1	1	1	1	1	1
I	0	0	0	0	0	0	0	1	1	1	1	1	1
J	0	0	0	0	0	0	0	0	1	1	1	1	1
K	0	0	0	0	0	0	0	0	0	1	1	1	1
L	0	0	0	0	0	0	0	0	0	0	1	1	1
M	0	0	0	0	0	0	0	0	0	0	0	1	1
N	0	0	0	0	0	0	0	0	0	0	0	0	1
O	0	0	0	0	0	0	0	0	0	0	0	0	0
P	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 4.2. Matrix representation of a maximally unbalanced tree containing 16 taxa (tree B in Fig. 4.1). Matrix representation contains 104 (50%) primitive states (0s) and 104 (50%) derived states (1s).

The unbalanced tree (Fig 4.1b) leads to a far more symmetrical MR (Table 4.2). Even if character polarity is unknown, and we code the most common state as state 0 and the least common as state 1 for the unbalanced tree, there are still more 1s in the data than in the MR of the symmetrical tree (Table 4.3). From here on I call this ratio of the less common state to the more common state in a character the state ratio (SR).

An SR of one indicates an equal number of each state in a character, and ratios closer to zero indicate unequal numbers of each state. Characters with a higher SR will also have a higher maximum number of steps that they can take in parsimony analyses. The average SR of the MR of the unbalanced tree is 0.43, compared with 0.3 for the balanced tree, showing that the maximum number of parsimony steps of the MR of the unbalanced tree is greater than that of the balanced tree. However, since all characters are binary, the minimum number of steps each MR can take in parsimony analysis is the same (26 each). Therefore, the difference between the maximum and minimum number of steps is greater for the MR of the unbalanced tree. Figure 4.2 shows the lengths of the two MRs (Tables 4.1 and 4.2) on 1,000,000 randomly generated trees. It is likely that by chance alone the balanced MR will have a lower tree length when plotted onto a random tree. This shows that it is more parsimonious to plot a matrix with a low SR onto a random tree than one with a high SR.

Taxon	1	2	3	4	5	6	7	8	9	10	11	12	13
A	1	1	1	1	1	1	0	0	0	0	0	0	0
B	1	1	1	1	1	1	0	0	0	0	0	0	0
C	0	1	1	1	1	1	0	0	0	0	0	0	0
D	0	0	1	1	1	1	0	0	0	0	0	0	0
E	0	0	0	1	1	1	0	0	0	0	0	0	0
F	0	0	0	0	1	1	0	0	0	0	0	0	0
G	0	0	0	0	0	1	0	0	0	0	0	0	0
H	0	0	0	0	0	0	0	0	0	0	0	0	0
I	0	0	0	0	0	0	1	0	0	0	0	0	0
J	0	0	0	0	0	0	1	1	0	0	0	0	0
K	0	0	0	0	0	0	1	1	1	0	0	0	0
L	0	0	0	0	0	0	1	1	1	1	0	0	0
M	0	0	0	0	0	0	1	1	1	1	1	0	0
N	0	0	0	0	0	0	1	1	1	1	1	1	0
O	0	0	0	0	0	0	1	1	1	1	1	1	1
P	0	0	0	0	0	0	1	1	1	1	1	1	1

Table 4.3. Matrix representation of a maximally unbalanced tree containing 16 taxa (tree B in Fig. 4.1) with the most common state for each character coded as 0 and the least common coded as 1. Matrix representation contains 146 (70%) 0s and 62 (30%) 1s.

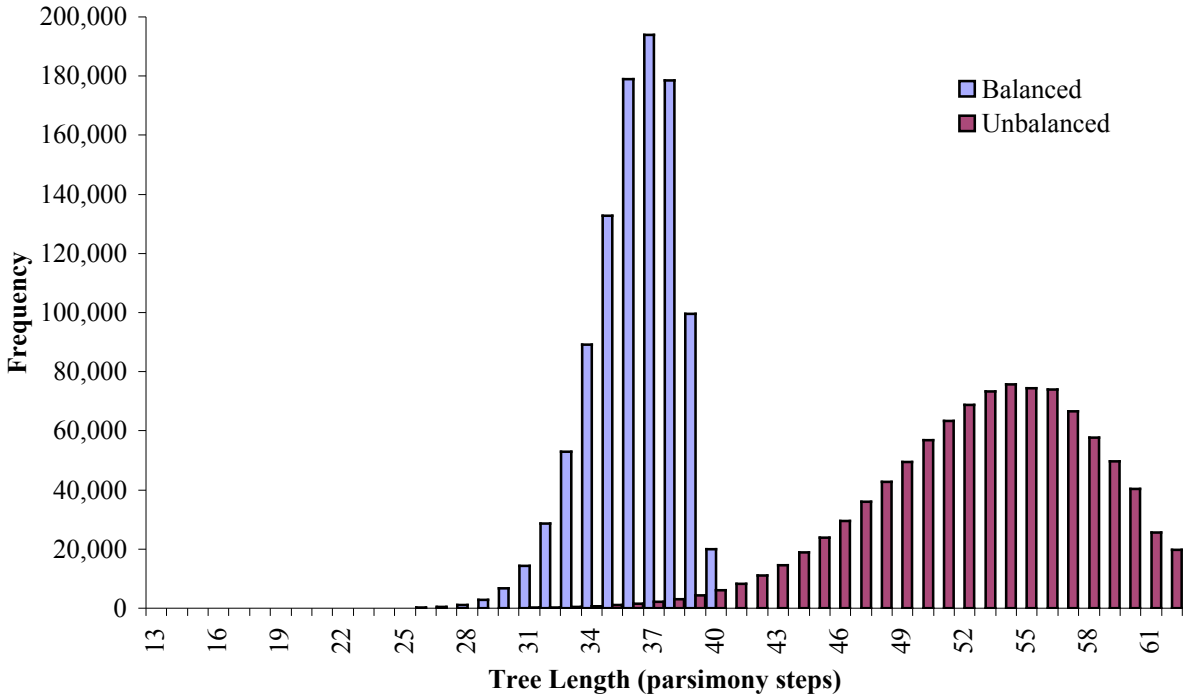


Figure 4.2. Histogram of parsimony lengths of the MRs of the balanced and unbalanced source trees mapped onto 1,000,000 randomly generated trees

Figure 4.3 shows the total tree lengths of the two MRs (Tables 4.1 and 4.2) on 10,000 randomly generated trees against the difference in length between the lengths of the unbalanced and balanced MR on that tree. It can be seen that at high tree lengths the unbalanced MR is much longer than the balanced MR. As the tree length decreases, the difference in length between the two MRs decreases until, on many of the shortest trees, the balanced MR is longer than the unbalanced MR despite its maximum possible length being much greater. This clearly demonstrates that it is more parsimonious to plot the balanced MR (high SR) onto trees more similar to the unbalanced tree than vice versa. This suggests that the bias in MRP towards unbalanced supertrees is caused by the higher SR value of MRs of unbalanced trees.

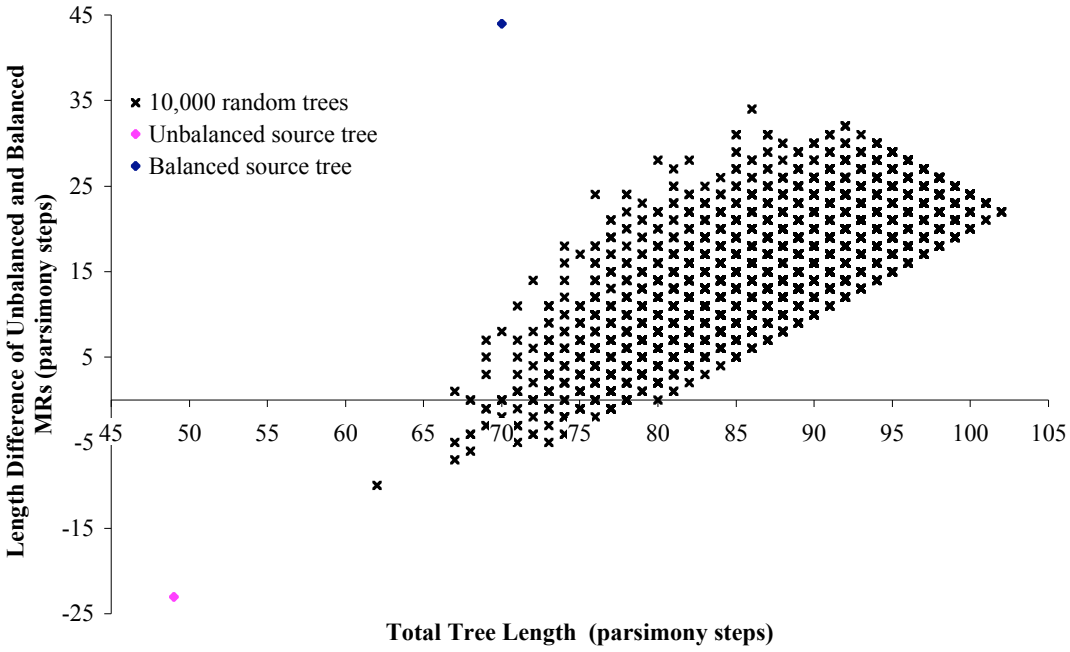


Figure 4.3. Scatter plot of total tree lengths of the MRs of the balanced and unbalanced source trees mapped onto 10,000 random trees. The y-axis represents the difference in length (parsimony steps) of the MRs of the two source trees when plotted onto the random trees. This value is simply the length of the unbalanced MR – the length of the balanced MR. Values on the y-axis greater than 0 indicate that the unbalanced MR is longer than the balanced MR, and vice versa. Points on the plot may represent many occurrences of the same value. The lengths of the two MRs mapped onto the two source trees are also marked.

Interestingly, as pointed out by Wilkinson *et al*. (in prep), a second type of MRP, called Purvis MRP (Purvis, 1995), shows the opposite bias, towards balanced trees. Purvis MRP uses one matrix element to represent each clade splitting its members from the members of its sister-group and the root. Any other leaves in the tree are scored as missing (?). The Purvis MRs of the two source trees are shown in tables 4.4 and 4.5.

With Purvis MRP, the MR of the balanced tree has an SR of 1, whereas the MR of the unbalanced tree has an SR of 0.125. Therefore, with Purvis MRP the bias is still towards the tree with the higher SR, which in this case is the balanced tree.

If unbalanced trees are more likely to be expressed in the results of MRP than balanced trees when an equal number of characters supporting each hypothesis are present, it is unclear what happens in a normal parsimony analysis if competing phylogenetic signals are present. If one signal supports a balanced hypothesis and another supports an unbalanced hypothesis, even if there are equal numbers of characters supporting each hypothesis, it is likely that the output tree will be relatively unbalanced. This could explain

the often reported bias of incorrectly reconstructed phylogenies towards unbalanced topologies (e.g. Mooers *et al.*, 1995).

Taxon	1	2	3	4	5	6	7	8	9	10	11	12	13
A	1	1	1	0	0	?	?	?	?	?	?	?	?
B	1	0	?	?	1	1	0	?	?	?	?	?	?
C	1	0	?	?	1	0	1	?	?	?	?	?	?
D	0	?	?	?	?	?	?	1	1	0	0	?	?
E	0	?	?	?	?	?	?	0	?	?	1	1	0
F	1	1	0	1	0	?	?	?	?	?	?	?	?
G	0	?	?	?	?	?	?	0	?	?	1	1	0
H	1	0	?	?	1	0	1	?	?	?	?	?	?
I	0	?	?	?	?	?	?	0	?	?	1	0	1
J	0	?	?	?	?	?	?	0	?	?	1	0	1
K	0	?	?	?	?	?	?	1	0	1	0	?	?
L	1	1	1	0	0	?	?	?	?	?	?	?	?
M	1	0	?	?	1	1	0	?	?	?	?	?	?
N	1	1	0	1	0	?	?	?	?	?	?	?	?
O	0	?	?	?	?	?	?	1	0	1	0	?	?
P	0	?	?	?	?	?	?	1	1	0	0	?	?

Table 4.4. Purvis matrix representation of a maximally balanced tree containing 16 taxa (tree A in Fig. 4.1). Matrix representation contains 40 (50%) primitive states (0s) and 40 (50%) derived states (1s).

Taxon	1	2	3	4	5	6	7	8	9	10	11	12	13
A	1	1	1	1	1	1	1	1	1	1	1	1	1
B	1	1	1	1	1	1	1	1	1	1	1	1	1
C	0	1	1	1	1	1	1	1	1	1	1	1	1
D	?	0	1	1	1	1	1	1	1	1	1	1	1
E	?	?	0	1	1	1	1	1	1	1	1	1	1
F	?	?	?	0	1	1	1	1	1	1	1	1	1
G	?	?	?	?	0	1	1	1	1	1	1	1	1
H	?	?	?	?	?	0	1	1	1	1	1	1	1
I	?	?	?	?	?	?	0	1	1	1	1	1	1
J	?	?	?	?	?	?	?	0	1	1	1	1	1
K	?	?	?	?	?	?	?	?	0	1	1	1	1
L	?	?	?	?	?	?	?	?	?	0	1	1	1
M	?	?	?	?	?	?	?	?	?	?	0	1	1
N	?	?	?	?	?	?	?	?	?	?	?	0	1
O	?	?	?	?	?	?	?	?	?	?	?	?	0
P	?	?	?	?	?	?	?	?	?	?	?	?	?

Table 4.5. Purvis Matrix representation of a maximally unbalanced tree containing 16 taxa (tree B in Fig. 4.1) with the most common state for each character coded as 0 and the least common coded as 1. Matrix representation contains 13 (11%) primitive states (0s) and 104 (89%) derived states (1s).

4.3.3.1 Bias Towards Unbalanced Trees in Compatibility

Like parsimony, many compatibility analysis methods may show a bias towards unbalanced topologies. The reason is again due to the relatively low SRs of characters supporting unbalanced trees. However, the reason for the bias is slightly different than in parsimony. As pointed out by many compatibility workers, binary characters with equal numbers of 0s and 1s (high SR) are more likely to be incompatible by chance alone (e.g. Le Quesne, 1969; 1972). Thus, if a character with a high SR is randomly permuted it is more likely to be incompatible with another character than is a character with a low SR. Therefore, if two characters are compatible, it is much more likely that the compatibility is attributable to chance alone if the characters have low SR values. This is analogous to a type I error in statistical tests. In this case, the null hypothesis would be that the two characters are random, and therefore unlikely to be compatible. If they are compatible, we can reject the null hypothesis and conclude that the characters are not random, and that they do contain phylogenetic signal. The type I error is that we have rejected the null hypothesis when it is true; i.e. when the character is random but by chance alone is compatible with other characters. To try to correct for this problem, many compatibility tests use random permutations to calculate the probability of the character being compatible by chance alone (calculating the type I error rate). They then use this value to downweight the significance of compatibility between characters that are likely to be compatible by chance alone. However, such methods have disadvantages. By correcting for the possibility that compatibility is due to chance alone, characters that are truly compatible due to phylogenetic signal can be downweighted simply because they have a high SR. As noted by Gauld and Underwood (1986), it is possible for this method to lead to characters with the same number of incompatibilities showing very different CCSR scores, and some characters can even show higher CCSR scores than other characters which have more incompatibilities with the rest of the matrix. For example, imagine a dataset that contained fifteen characters, ten of which were mutually compatible, but incompatible with the remaining five characters, which were themselves incompatible with all other characters in the data. Simply counting the number of pairwise incompatibilities of the characters would show that the first ten characters are all incompatible with five characters, while the last five characters are each incompatible with fourteen characters. Therefore, each of the ten compatible characters is scored equally by simply counting the number of character with which they are incompatible. If it is now assumed that one of the

group of ten characters has an SR of 1 (equal numbers of states 0 and 1) and a second character in the group of ten has a minimal SR (e.g. just two taxa coded state 1 and all others as state 0), then problems arise with some compatibility methods. Simply counting incompatibilities still gives the same result for the two characters: an incompatibility score of five for each. However, if CCSR, NDev or LQP methods are used, an expected incompatibility value is needed for each character. The expected compatibility value for the low SR character will be greater than that of the high SR character. This means that, when using CCSR, NDev or LQP, it is likely that a low SR character will be considered less compatible (weaker) than a high SR character simply because of the distribution of character states it has, when in fact both characters are compatible with an equal number of other characters in the data. In practice, applying methods such as the boildown to these compatibility methods can lead to characters with low SRs being downweighted or removed from the data preferentially. One side-effect of this is that, as previously mentioned, datasets describing asymmetrical trees generally have higher SRs. Therefore, given a dataset containing characters supporting two hypotheses, one being a balanced tree and the other an unbalanced tree, many compatibility methods will favour the unbalanced tree. Methods that involve simply counting the number of incompatibilities or finding maximal cliques, are not subject to this bias. Other methods, such as the CCSR, NDev and LQP, were designed to correct for the chance of compatible characters being compatible simply by chance. However, our goal in carrying out phylogenetic analyses is to find the tree representing true phylogenetic signal, and thus the evolutionary history of a group. To do this we might reasonably assume that the strongest signal in the data will be more likely to be the correct one. Any noisy characters that are compatible with the true signal will not change this result, although they may cause incorrect resolution of parts of the tree that is not resolved by phylogenetically useful data. It seems preferential to me to accept that some noisy characters are compatible with the true tree, and leave such characters in our data than to downweight some characters simply because they describe a synapomorphy of a small number of taxa.

Chapter 5:  The Boildown Computer Program

5.1 Introduction

Boildown is a Mac OSX program written as part of my PhD studies in order to allow a number of compatibility analysis and boildown methods to be used for data exploration. Many programs are available for carrying out tree-based analyses, such as parsimony and likelihood searches, and calculating tree statistics and support measures, such as Hennig86 (Farris, 1988), PHYLIP (Felsenstein, 1995), NONA (Goloboff, 1998), Radcon (Thorley and Page, 2000) and PAUP (Swofford, 2003). PAUP, especially, is a very widely used program that is both user friendly and efficient at performing a large number of phylogenetic functions. Compatibility methods, conversely, have not been as widely used as parsimony and likelihood, and as such have not been the focus of such a great deal of programming work. A number of programs are available, but these have generally been created to carry out a single analysis methods, and have not been designed with a user-friendly interface. For example, clique analysis of binary data is implemented by a number of programs including CLINCH (Fiala, 1984) and the CLIQUE component of the PHYLIP package (Felsenstein, 1995). The most comprehensive compatibility program currently available is PICA (Wilkinson, 2001). PICA is a suite of programs that carry out CCSR and NDev boildowns for binary data, calculate LQP scores for both binary and multistate data, and implement new methods by which compatibility can be used to study the relative stability of taxa. PICA also contains a number of programs that calculate compatibility based PTP tests and split support. Although PICA offers many of the data exploration methods necessary for my work, it does not carry out LQP boildowns, and is restricted to boiling down binary data. The limits on the size of the data matrices it can assess are also smaller than *Boildown*. There were also new methods devised during my work, such as fuzzy compatibility and the boildown bootstrap, which are very labour intensive if carried out manually, so a computer program which automates these methods was very useful. The aims in creating *Boildown* were:

- To create a compatibility analysis program that could carry out 'No.', CCSR, NDev and LQP boildowns on multistate data matrices.
- To make the program user-friendly and compatible with the NEXUS file type used by the most common parsimony analysis software, PAUP.

- To implement algorithms for fuzzy compatibility and the boildown bootstrap, which are introduced for the first time in Chapter 4.

5.2 Specifications

Boildown is available for Apple Macintosh computers using the OS X (OS 10.2 or later) operating system. The compatibility algorithms in *Boildown* were programmed in Mac OS X developer using C, and the user interface using Mac Carbon. The interface of the program is based loosely on those of similar phylogenetic programs, such as PAUP (Swofford, 2003) and Radcon (Thorley and Page, 2000). The main window of the program is the log window (Fig. 5.1), which opens when the program is launched. All results and messages from the program are printed into the log window, which can be edited, printed and saved in the form of a 'TEXT' file. A copy of Boildown is included on CD inside the back cover of this thesis.

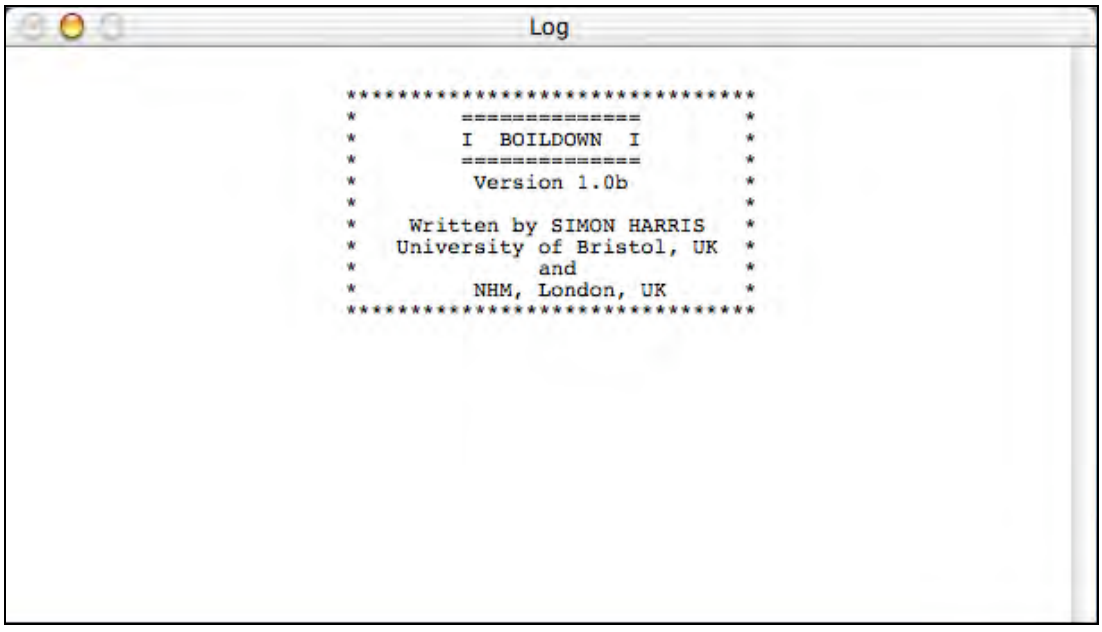


Figure 5.1. The *Boildown* log window.

5.3 Users Manual

5.3.1 The File Menu

Figure 5.2. The File Menu.

5.3.1.1 New

Opens a blank document for editing.

5.3.1.2 Edit

Allows the user to select a text file to open for editing (see Fig. 5.3).

5.3.1.3 Execute

Used to input data in NEXUS format into the *Boildown* program so that it can be analysed. When the execute option is selected the user is asked to select a file to open. This file must be a valid NEXUS file (e.g. see Fig. 5.3), and must conform to the following format:

- The file must start with the line #NEXUS, which identifies it as a NEXUS file.

The number of taxa in the data matrix must be specified in the form: 'Ntax = *x*', where *x* is the number of taxa in the dataset. The maximum number of taxa permitted by the Boildown program is 500.

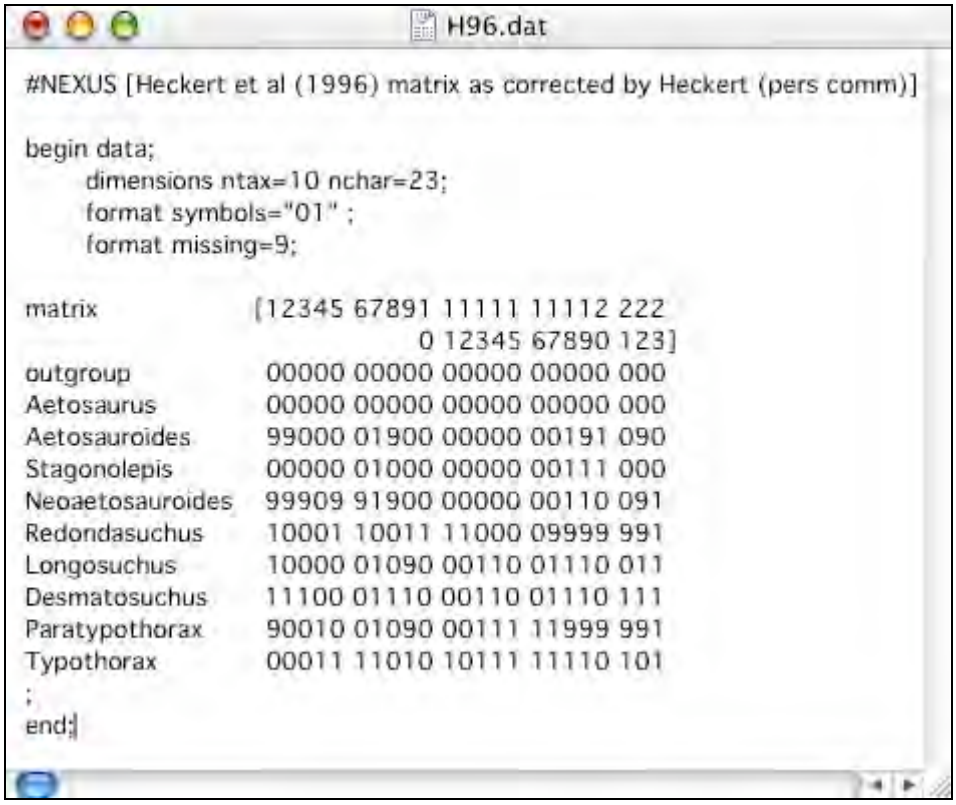


Figure 5.3. Example of a NEXUS file open for editing.

- The number of characters in the data matrix must be specified in the form: 'Nchar = *x*', where *x* is the number of characters in the dataset. The maximum number of characters permitted by *Boildown* is 3500. Some molecular datasets may exceed this limit. However, by removing constant and parsimony uninformative characters from the NEXUS file, the number of characters in the matrix can often be reduced significantly. These characters can be removed before executing the data in *Boildown* without affecting the results of any analyses it carries out. Parsimony uninformative characters can easily be removed by excluding them in PAUP before exporting the data in NEXUS format.
- All symbols used in the data to represent character states must be defined using the statement: 'symbols = "*x y z*"', where *x*, *y* and *z* are character states present in the dataset. *Boildown* can handle data containing up to 22 character state symbols. The symbol '?' is reserved for missing data and cannot be used as a character state symbol.
- To define a symbol to represent missing or unknown character state scores in the data matrix, the following statement can be used: 'missing = *x*', where *x* is the

symbol used to represent missing data. If no symbol is defined to represent missing data, *Boildown* will use the default missing data character, which is '?'.

- To define a symbol to represent gaps in the data matrix, the following statement can be used: 'gap = *x*', where *x* is the symbol used to represent gaps. If the data does not contain gaps, then this does not need to be defined. Gaps are simply treated by the program as missing data.
- The data matrix, which must be situated after the above entries, must be entered in the form of a taxon by character matrix, with characters in columns and taxa in rows, each with the taxon name at the beginning of the row. Taxon names cannot contain blank spaces, and must contain fewer than 30 characters. Other than within taxon names, blank spaces and line breaks are permitted within the data matrix. The data matrix must be preceded by the command: 'matrix', and followed by a semi-colon and the command: 'end;'. Importantly, *Boildown* cannot cope with transposed data matrices in which taxa are in columns and characters in rows.
- Uncertainties and/or polymorphisms in character state scorings must be entered into the data matrix in parentheses and/or curly brackets ({}). There are problems in compatibility analysis associated with both uncertainty and polymorphism within taxa (see Chapter 4). Therefore, both are simply replaced by missing data in *Boildown* before analysis. This means that the number of uninformative characters identified by *Boildown* is likely to be greater than in other programs.
- Any text in the NEXUS file that is enclosed in square brackets ([]) is ignored by the Boildown program. This can be used to enter comments anywhere in the NEXUS file.

If any of the necessary information is missing from the NEXUS file, *Boildown* will return an error message specifying the problem and execution of the file will be terminated. If execution is successful, *Boildown* will report this in the Log window along with some information about the file now in memory (Fig. 5.4).

```
Executing File "H96.dat"...
    Number of taxa = 10
    Number of characters = 23
    2 states (01)
...File successfully executed
```

Figure 5.4. Log output of a successful execution of the data shown in figure 5.3.

Once execution is complete, analyses can be run on the executed data matrix.

5.3.1.4 Close

Closes the highlighted window. If changes have been made to the window, *Boildown* will ask if the user wants to save or discard these changes. The log window cannot be closed.

5.3.1.5 Save, Save As

Saves the highlighted window.

5.3.1.6 Revert

Reverts to the last saved version of the highlighted window.

5.3.1.7 Page Setup

Opens the page setup options window.

5.3.1.8 Print

Prints the selected window.

5.3.2 The Edit Menu

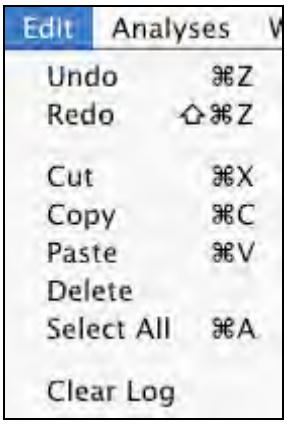


Figure 5.5. The Edit Menu.

5.3.2.1 Undo

Undoes the most recent action.

5.3.2.2 Redo

Redoes the last undone action.

5.3.2.3 Cut, Copy

Cuts or copies highlighted text.

5.3.2.4 Paste

Pastes cut or copied text to the current cursor position.

5.3.2.5 Delete

Deletes highlighted text.

5.3.2.6 Select All

Highlights all text in the selected window.

5.3.2.7 Clear Log

Clears the Log window of all text.

5.3.3 The Analyses Menu

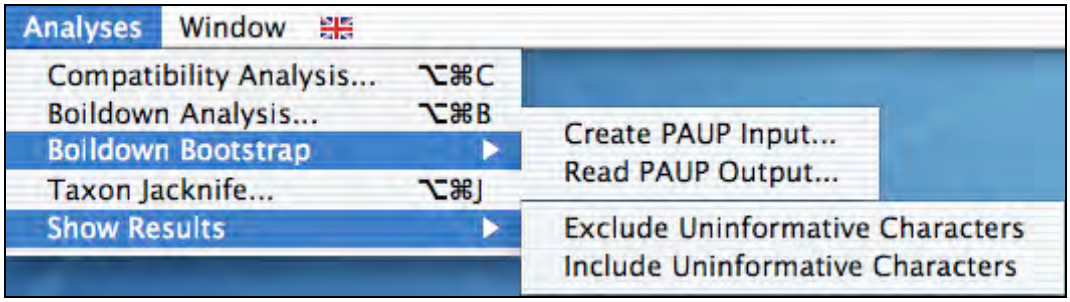


Figure 5.6. The Analyses Menu.

5.3.3.1 Compatibility Analysis

Carries out a compatibility analysis on the data currently in memory. Choosing this option opens the Compatibility Options (Fig. 5.7) dialog box, which allows the user to specify options for the analysis.

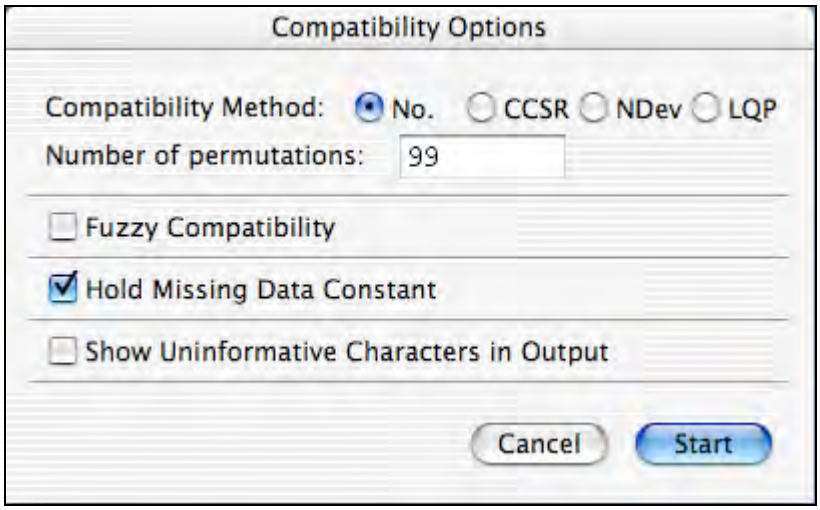


Figure 5.7. The compatibility options dialog box.

- **Compatibility Method**: The type of compatibility analysis desired. The 'No.', or number option is simply a count of the number of other characters in the data with which each character is incompatible. This is the quickest of the available methods, as no random permutations are used. Analyses using all other methods also report the number of incompatibilities as part of their results. The CCSR, NDev and LQP methods are described in detail in Chapter 4. If any of these methods is selected, the desired number of random permutations with which to calculate expected incompatibility values must be specified. It is recommended that at least 99 permutations are used for most analyses. However, increasing the number of permutations increases the duration of the analysis. If zero is entered for the number of permutations, *Boildown* will carry out a 'No.' permutation test.
- **Fuzzy Compatibility**: Selecting this option instructs *Boildown* to calculate fuzzy compatibility values during analyses rather than simply counting the number of incompatible character pairs. Fuzzy compatibility is described in detail in Chapter 4.
- **Hold Missing Data Constant**: If this option is selected missing data and gaps will be preserved in the same positions during permutations of characters. It is generally recommended that this option is selected in order to reduce the number of variables affecting the results of analyses.
- **Show Uninformative Characters in Output**: If this option is deselected, all uninformative characters will be excluded from the output results.

When a compatibility analysis is complete, the results are displayed in the log window. An example of the results reported after a CCSR analysis is shown in figure 5.8.

```
CCSR analysis of "H96.dat":
     99 random permutations
     Missing data points held constant
     5 characters are uninformative
     0 of which are constant
     18 characters are informative

Results:
Char           Obs      Exp         CCSR
   1:            7        9.0       0.775
   2:        UNINFORMATIVE CHARACTER
   3:        UNINFORMATIVE CHARACTER
   4:            3        5.9       0.508
   5:            7        7.2       0.975
   6:EQUIVALENT TO CHARACTER 5       0.975
   7:            6        8.9       0.674
   8:        UNINFORMATIVE CHARACTER
   9:            1        7.2       0.140
  10:        UNINFORMATIVE CHARACTER
  11:            7        6.4       1.098
  12:        UNINFORMATIVE CHARACTER
  13:            4       10.1       0.396
  14:EQUIVALENT TO CHARACTER 13      0.396
  15:EQUIVALENT TO CHARACTER 4       0.508
  16:EQUIVALENT TO CHARACTER 4       0.508
  17:            0        7.9       0.000
  18:            0        5.3       0.000
  19:            0        4.8       0.000
  20:            0        4.6       0.000
  21:            2        5.7       0.352
  22:            1        4.4       0.227
  23:            1        8.5       0.117
```

Figure 5.8. The results of a CCSR compatibility analysis of the data shown in figure 5.3.

For each character in the data the results report the number of other characters with which the character has been observed to be incompatible (Obs), the mean number of characters with which the random permutations of the character are incompatible (Exp) and the CCSR value of the character. If two or more characters show the same distribution of character states, the first (in terms of occurrence in the data matrix) will be reported, and subsequent equivalent characters will be labelled as such. This allows the program to analyse only one of a group of equivalent characters, which increases the speed of the analysis.

5.3.3.2 Boildown Analysis

Carries out a boildown analysis of the data currently in memory. The boildown procedure is discussed in Chapter 4. Selecting the boildown analysis option opens the compatibility options dialog box (Fig. 5.7), as in a normal compatibility analysis (see compatibility analysis for discussion of options). It is important to note that when a boildown is carried out using the LQP method, character comparisons must be made at each stage of the boildown. This means that this procedure can be extremely time consuming. No., CCSR and NDev analyses make character comparisons only once, and are therefore far quicker. When the boildown is complete, the results are output to the log

window. First, the results of the compatibility analysis of the complete data are output, as in a normal compatibility analysis. These are followed by the results of the boildown. Figure 5.9 shows a typical output from a boildown analysis.

```
Boildown:
1) Character(s) removed:
     11 at a CCSR value of 1.098. Matrix CCSR = 0.425.
2) Character(s) removed:
     5 6 at a CCSR value of 1.027. Matrix CCSR = 0.350.
3) Character(s) removed:
     1 at a CCSR value of 0.541. Matrix CCSR = 0.139.
4) Character(s) removed:
     22 at a CCSR value of 0.295. Matrix CCSR = 0.082.
5) Character(s) removed:
     7 at a CCSR value of 0.300. Matrix CCSR = 0.050.
All remaining characters are compatible
```

Figure 5.9. The results of a CCSR boildown analysis of the data shown in figure 5.3.

For each stage of the boildown the characters removed are reported, along with their individual compatibility value (No., CCSR, NDev or LQP) and the average compatibility value of all remaining characters in the data matrix at that stage of the boildown (Matrix No., CCSR, NDev or LQP).

5.3.3.3 Boildown Bootstrap Analysis

Carries out a type 1 boildown bootstrap analysis (see Chapter 4 for a detailed description) of the data currently in memory. This method comprises three parts. First, a boildown analysis is carried out in *Boildown*. However, because *Boildown* cannot carry out parsimony analyses, the bootstrap analyses for each step of the boildown must be carried out in PAUP. Therefore, during the boildown process, *Boildown* creates a file that must subsequently be executed in PAUP, which instructs PAUP to carry out the necessary bootstrap analyses and save the bootstrap trees produced in a tree file. This tree file should then be read back into *Boildown*, which will extract and report the total bootstrap values for each tree.

1) Create PAUP Input

Selection of this option first opens the Bootstrap Options dialog box (Fig. 5.10). This dialog box contains a number of options relevant to bootstrap analyses carried out by PAUP.

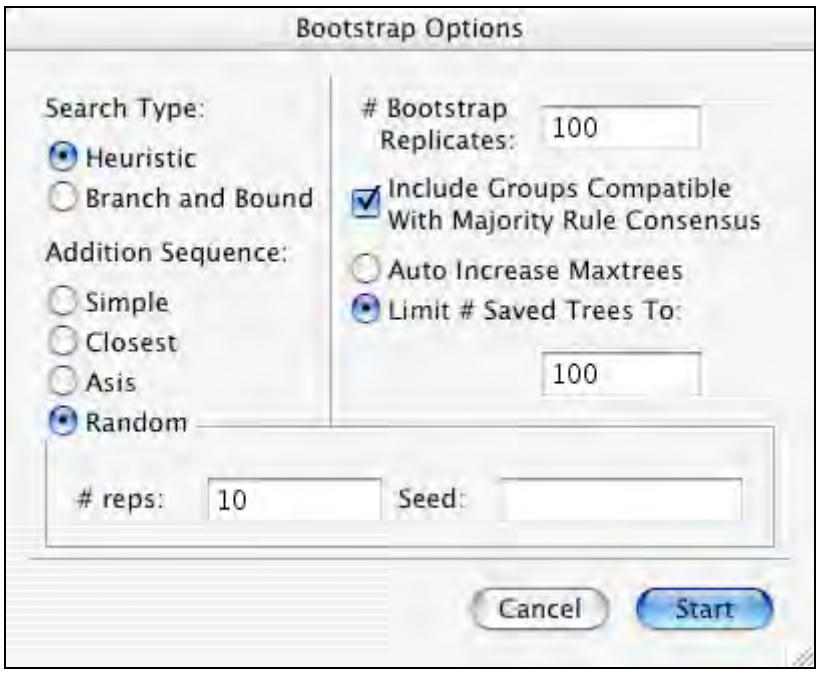


Figure 5.10. The Bootstrap Options dialog box.

- **Search Type**: This option specifies the type of parsimony search that PAUP will use when carrying out bootstrap analyses. The branch and bound option is not available when the dataset in memory contains more than 18 taxa, as this type of analysis is very time consuming with large numbers of taxa.

- **Addition Sequence**: When using a heuristic parsimony search, PAUP offers a number of addition sequence methods for adding taxa to the tree. These choices are also offered in the bootstrap options dialog box in *Boildown*. If a random addition sequence is selected, the **number of replicates** and a **starting seed** can be specified.

- **# Bootstrap Replicates**: This number specifies how many bootstrap replicates will be carried out in the PAUP analysis. It is recommended that at least 100 replicates are used, and more if possible. However, increasing the number of replicates increases the search time.

- **Include Groups Compatible With Majority Rule Consensus**: When this option is selected, all bootstrap trees saved by PAUP will be the 50% majority rule consensus of the bootstrap replicate trees with any compatible groups resolved. When the bootstrap trees are read back into *Boildown*, it calculates a total bootstrap value for each tree. If compatible groups are not included, then this measure will simply sum bootstrap proportions of the nodes in the 50% majority rule tree rather than giving a total bootstrap value as described in Chapter 4. It is therefore recommended that this option is selected.

- **Auto Increase Maxtrees / Limit Maxtrees**: PAUP contains an option for setting the maximum number of trees it will store during a search. If the auto increase maxtrees option is selected, PAUP will store all trees it finds up until its memory allocation is full. If the limit maxtrees option is selected, the user can specify the maximum number of trees PAUP will store during each search. The greater the number of trees saved, the better chance PAUP has of finding all most parsimonious trees in a search. This means the results of the bootstrap may be more accurate if the auto increase maxtrees option is selected, or if a large number of trees are saved. However, saving more trees means more trees must be swapped in the search, so that the search is likely to be more time consuming.

Once all bootstrap options are made, and 'done' is selected in the bootstrap options dialog box, the compatibility options dialog box opens to allow options for the compatibility analysis to be made. Finally, the user is asked to give a name and location to save the PAUP input file. It is important that this file is saved in the same location as the #NEXUS file containing the data being subjected to the boildown bootstrap. The saved input file must then be executed in PAUP.

2) Read PAUP Output

This option should be selected once PAUP has finished running the bootstrap analyses for each stage of the boildown. Selecting this option asks the user to choose a file to open. The treefile produced by the PAUP analysis should be chosen. This file will have the same name as given to the PAUP input file created in *Boildown*, but will have the extension '.tre'. *Boildown* reads the bootstrap values from the tree file and outputs the results to the log window, as shown in figure 5.11.

```
Number of taxa in trees = 10
Maximum number of resolved nodes = 7

Bootstrap Tree*  >0   >50%     Total Bootstrap  Av Bootstrap
1)               7     6           475.49           67.93
2)               7     6           492.34           70.33
3)               7     6           501.33           71.62
4)               7     5           483.76           69.11
5)               7     7           522.63           74.66
6)               7     7           536.43           76.63

...Extraction complete.
*Tree 1 is the bootstrap tree produced when all informative characters
are included. Subsequent trees represent each step in the boildown
process.
```

Figure 5.11. The results of a boildown bootstrap analysis of the file shown in figure 5.3.

For each stage of the boildown the results show the number of nodes in the bootstrap tree that are resolved, the number of nodes with bootstrap support greater than 50%, and the sum and average of the bootstrap supports for all nodes. Note that the average

bootstrap is the sum of the bootstrap support values for all nodes divided by the maximum possible number of resolved nodes in the tree.

5.3.3.4 Taxon Jacknife Analysis

Carries out a taxon jacknife analysis of the dataset currently in memory. The taxon jacknife (Wilkinson, 2001) is a development of a procedure introduced by Guise *et al.* (1982) and Gauld and Underwood (1986) for marking taxa that are particularly responsible for incompatibilities within a dataset. Taxon jacknifing involves calculating the total number of incompatibilities within a dataset and comparing that value with the total number of incompatibilities in the same dataset with each taxon removed. The difference is the number of incompatibilities that no longer occur when the taxon in question is not present. *Boildown* carries out first and second order taxon jacknifing (Wilkinson, 2001). Selecting the Taxon Jacknife Analysis option opens the Taxon Jacknife Options dialog box (Fig. 5.12). This dialog box contains options relevant to the taxon jacknife procedure.

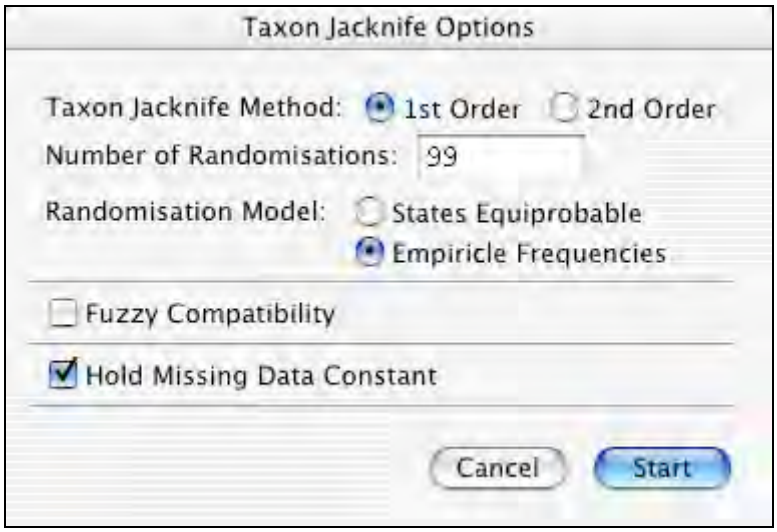


Figure 5.12. The Taxon Jacknife Options dialog box.

Taxon Jacknife Method: First or second order taxon jacknifes can be selected. A **first order** taxon jacknife removes a single taxon and reports the number of incompatibilities with which the taxon is involved. Using this method, a number of randomisations can be specified so that the observed number of incompatibilities a taxon is involved with can be compared with expected value under two models of randomisation of the taxon, as discussed by Wilkinson (2001). The first model, **states equiprobable**, simply creates the random taxon by selecting its characters states for each character equiprobably from the states present in that character. The second model, **empiricle frequency**, selects the character state for each character for the random taxon based on their observed

frequencies in that character (Wilkinson, 2001). The **number of randomisations** and the randomisation model can be specified when first order jacknifing is selected. A **second order** jacknife removes all pairs of taxa including the taxon of interest and calculates the average number of incompatibilities for these pairs. This is done for each taxon in the data.

Fuzzy Compatibility and **Hold Missing Data Constant**: See Compatibility Options above.

When the taxon jacknife is complete, the results are output to the log window. An example of results from first and second order jacknifing is shown in figure 5.13.

1st Order Taxon Jacknife of "H96.dat"

99 randomisations under the empiricle frequencies model
Missing data points held constant
Number of uninformative characters in matrix = 5
Total incompatibility in matrix = 28

Taxon Excluded	Uninf	%Missing	Caused	Exp	P
outgroup	7	0.0	0	12	0.030
Aetosaurus	7	0.0	0	11	0.030
Aetosauroides	6	21.7	0	8	0.060
Stagonolepis	6	0.0	0	10	0.030
Neoaetosauroides	5	30.4	0	6	0.060
Redondasuchus	8	26.1	26	7	1.000
Longosuchus	6	4.3	2	9	0.140
Desmatosuchus	7	0.0	2	10	0.100
Paratypothorax	8	30.4	9	4	0.920
Typothorax	12	0.0	25	9	1.000

2nd Order Taxon Jacknife of "H96.dat"

Number of uninformative characters in matrix = 5
Total incompatibility in matrix = 28

Taxon Excluded	Uninf	%Missing	1st Order	2nd Order
outgroup	7	0.0	0	11
Aetosaurus	7	0.0	0	11
Aetosauroides	6	21.7	0	11
Stagonolepis	6	0.0	0	11
Neoaetosauroides	5	30.4	0	11
Redondasuchus	8	26.1	26	28
Longosuchus	6	4.3	2	12
Desmatosuchus	7	0.0	2	13
Paratypothorax	8	30.4	9	16
Typothorax	12	0.0	25	27

Figure 5.13. The results of first and second order jacknife analyses of the data shown in figure 5.3.

For each jacknife analysis the number of uninformative characters and the total number of incompatibilities in the complete dataset are reported. In first order jacknifing, for each taxon the results shown are: the number of uninformative characters when the taxon is excluded (uninf), the proportion of the taxon data that is missing (%missing), the number of incompatibilities removed from the data by excluding the taxon (caused), the number which would be expected to be removed if the taxon was randomly generated using the selected model (exp), and the proportion of random taxa that caused less incompatibilities than the observed value (p-value). Second order jacknifing results again

show uninf, %missing and caused, but also show the average number of incompatibilities removed from the data when all pairs of taxa including the taxon in question are removed.

5.3.3.5 Show Results

If a compatibility analysis has been carried out on a dataset, this option allows the results to be shown again including or excluding uninformative characters.

Chapter 6: The Stagnation of Phylogenetic Debates and the Affinities of Turtles

6.1 Introduction

One obstacle to phylogenetic reconstruction occurs when two or more competing hypotheses of relationships reach a position of stalemate. Debates over phylogeny reconstruction have been, and still are, abundant in the field of systematics, with prominent examples including the affinities of tetrapods (e.g. Meyer, 1995), snakes (e.g. Rieppel and Kearney, 2001) and scorpions (e.g. Dunlop and Braddy, 2001). Such debates are often a result of large gaps in the fossil record, rapid diversification within a lineage or highly derived morphologies within extant lineages that make phylogeny reconstruction difficult (Meyer and Zardoya, 2003).

Rieppel and Kearney (2002) suggested that stagnation of systematic debates is often a result of an inability to test character hypotheses, and that this is primarily due to the trivialisation of the initial conjecture of homology. They argued that many debates in vertebrate systematics could be resolved by reassessing character constructions and formulating testable character hypotheses. It is true that resolution of conflicting interpretations of morphology (Rieppel and Kearney, 2002) along with the addition of further morphological and/or molecular data (Wilkinson *et al*., 1997) might lead to consensus in such cases, but reassessment of every character utilised in large analyses is laborious and not guaranteed to resolve differences of opinion between authors. Alternatively, in the absence of consensus in morphological assessment, further quantitative analysis can provide useful insights into the reasons for the stalemate.

In instances where different data relating to the same taxa are analysed using the same methods (usually parsimony), any conflict between the resulting phylogenies must be entirely caused by differences in taxonomic and/or character sampling, or in the assignments of character states to taxa. There are a number of ways that resolution of such conflict can be approached. The data could be re-examined until agreement is reached over character sampling and scoring, as suggested by Rieppel and Kearney (2002). Alternatively, a resolution may be sought by judging the value of the data on criteria other than that used during the initial method of analysis.

Conventionally, morphological data matrices are analysed solely using parsimony methods. In these cases, an alternative criterion that may be used to judge between conflicting data is compatibility analysis, as described in Chapter 4.

6.2 A Case Study: The Origin of Turtles

One high profile example of a phylogenetic debate that has raged for well over a century, but is seemingly no nearer a solution, is that of the origin of the Testudines (turtles, terrapins and tortoises, together often grouped under the single common name of turtles). The controversy almost certainly stems from the highly divergent morphology of the group, and the lack of intermediate forms. The earliest known turtle, *Proganochelys quenstedti* Baur, 1887, possesses a derived turtle morphology, extremely similar to extant forms (Fig. 6.1). This makes assessments of homology with potentially close relatives difficult and, to judge from the abundance of alternative interpretations in the literature, open to subjectivity and disagreement. Unfortunately, modern phylogenetic studies have to some extent recapitulated rather than resolved the controversy (e.g. Rieppel and Reisz, 1999; Lee, 2001).

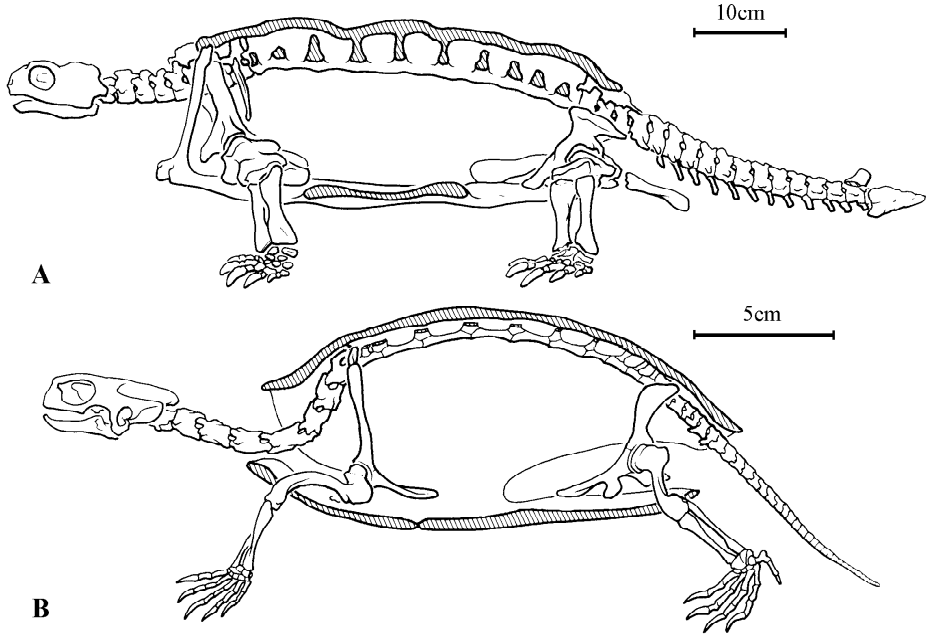


Figure 6.1. The general morphology of turtles, showing the similarity between Triassic and modern forms. (A) The oldest known turtle, *Proganochelys quenstedti*, from the Upper Triassic of Germany. (B) The extant turtle, *Emys orbicularis* (the European pond turtle). Figure modified from Gaffney (1990).

6.2.1 History of the Debate

In the late 19th century, the amniote skull, and the temporal openings it possessed, was identified as a useful tool in the classification of the Reptilia. Günther (1867) noticed that rhynchocephalians differ from Squamates in possessing two temporal arches, while in squamates the lower arch has been lost. Baur (1889) and Cope (1892) developed the "theory of fenestration", that described the evolution of the temporal region from primitively closed, through the opening of a supratemporal fenestra, to finally reach a condition with both supra- and laterotemporal fenestrae present. The use of these fenestrae in classification was taken further by Osborn (1903), who subdivided Reptilia into two classes based solely on this character. He defined the Synapsida as those taxa with a single or undivided temporal arch, and the Diapsida as those taxa possessing double or divided temporal arches. Within his Synapsida, Osborn (1903) included Cotylosauria, Anomodontia (the mammal-like reptiles), Testudinata and Sauropterygia. In 1916, Goodrich proposed a modification to Osborn's (1903) two subclasses. He stated that "as is always the case when we endeavour to classify by a single character, we are liable to confuse forms in which they have been or are being secondarily obliterated, and to misinterpret aberrant modifications" (Goodrich, 1916: 263). More simply, as noted by Darwin, "a classification founded on any single character, however important that may be, has always failed" (Darwin, 1859: 402 in 1968 edition). Goodrich (1916) subdivided the Reptilia into three groups, the Protosauria (Microsauria, Cotylosauria, Pareiasauria and Procolophonia), the Theropsida (those taxa with one lateral temporal fenestra limited below by a single bar, such as mammals) and the Sauropsida (those taxa with two lateral temporal foramina and two bars, including diapsid reptiles and birds). Just one year later the classification scheme was modified again. Williston (1917) renamed the group of reptiles lacking temporal fenestration as the Anapsida and reverted to use the names Synapsida and Diapsida, originally proposed by Osborn (1903), for taxa with one and two temporal openings respectively. This classification is still used today, and is illustrated in figure 6.2 with generalised skull diagrams and examples of each of the three main amniote groups. The figure also includes a generalised skull and example of a fourth group, the euryapsids, which possess a single supratemporal fenestra.

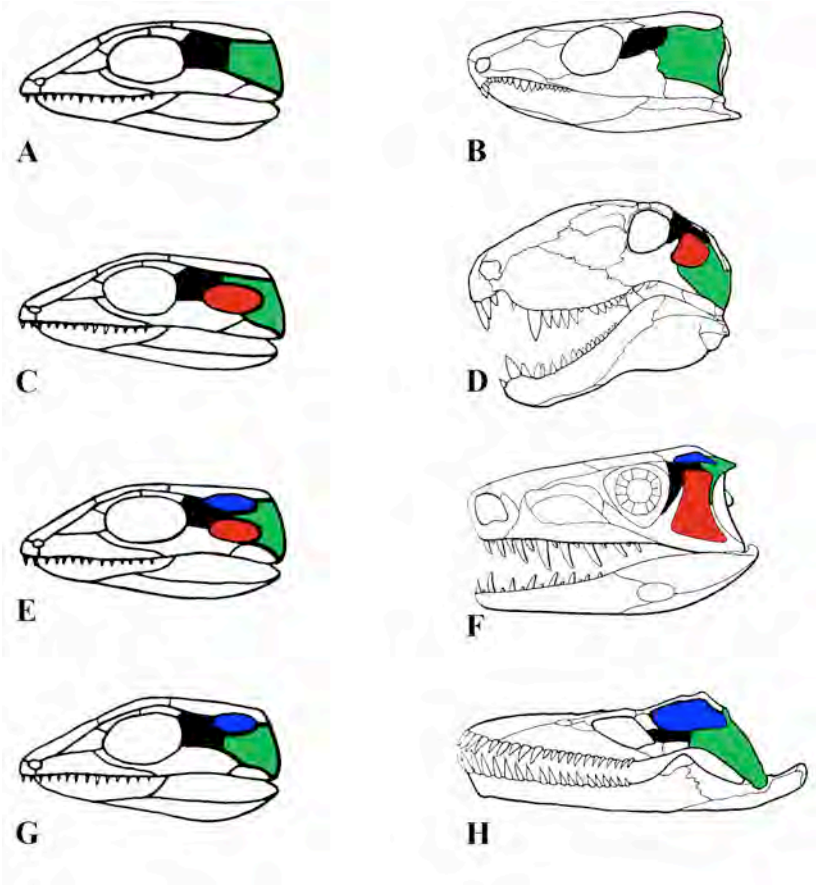


Figure 6.2. Examples of the main amniote groups based on skull fenestration. (A) A generalised anapsid skull with no temporal openings. (B) The skull of *Captorhinus*, a Permian anapsid (captorhinid). (C) A generalised synapsid skull with a single laterotemporal fenestra bounded above by the postorbital and squamosal. (D) The skull of *Dimetrodon*, a Permian synapsid (pelycosaur). (E) A generalised diapsid skull with two temporal fenestrae separated by a bar formed by the postorbital and squamosal. (F) The skull of *Euparkeria*, a Triassic diapsid (thecodont). (G) A generalised euryapsid skull with a single supratemporal fenestra bounded below by the postorbital and squamosal. (H) The skull of *Muraenosaurus*, a Jurassic euryapsid (plesiosaur). Key: black = postorbital, green = squamosal, blue = supratemporal fenestra, red = laterotemporal fenestra. Figure modified from Colbert and Morales (1991).

By the early 1920s, the classification of many amniote groups was becoming well established. The position of turtles, however, remained problematic. Watson (1914: 1011) stated that "the great group Chelonia is of unusual interest, because it is the only example of a persistent and world-wide order whose structure is entirely dependent on a bizarre specialization: the development of the shell". Broom (1924), in his attempt at reptile classification, discussed the problematic nature of turtles and the difficulties involved in identifying their affinities. "There is no order of reptile living or extinct concerning whose affinities greater differences of opinion have been expressed…[They] are so extremely

specialised and in some respects degenerate, so that picking out ancestral characters amid the more recent specialisations is somewhat like the reading of a difficult palimpsest" (Broom, 1924: 48). However, many morphologists have attempted to classify turtles, and three main hypotheses of their origins arose during the early part of the century.

The first of these centred on the Early Triassic reptile, *Eunotosaurus africanus* (Fig. 6.3). In his description of this problematic reptile, Seeley (1892) noted similarities between the vertebrae and expanded ribs of *Eunotosaurus* and chelonians, but concluded that "from the fragmentary condition of the remains, it seems inexpedient to determine absolutely the systematic position of this genus" (Seeley, 1892: 585). Later, Watson (1914) postulated a hypothetical ancestral turtle, which he named "Archichelone". He listed a number of traits that "Archichelone" would have possessed, and suggested that *Eunotosaurus* fitted the bill perfectly, since it exhibited, among other characters, tortoise-like vertebrae, intercentral rib articulations and a short, powerful ulna crest on the humerus. However, Watson did warn that *Eunotosaurus* was not well enough known for him to conclude that it was the ancestor of turtles. Broom (1924) re-evaluated turtle origins and, although admitting that the expanded ribs of *Eunotosaurus* look superficially chelonian, argued that this similarity was purely convergent, because the ribs of *Eunotosaurus* were overlapping and movable unlike those of chelonians. Gregory (1946), conversely, in a detailed comparison of turtles with pareiasaurs and placodonts, included *Eunotosaurus* as a chelonian, implying that the close relationship between these two groups was certain. From this point onwards, however, the *Eunotosaurus* hypothesis of turtle origins lost much of its support. In 1969, Cox asserted that the chelonian carapace evolved from a *Stagonolepis*-like carapace of segmented dermal plates rather than from expanded ribs similar to those of *Eunotosaurus*. He stated that "there are no detailed similarities between *Eunotosaurus* and Chelonia which would unequivocally and convincingly demonstrate a phyletic relationships between them" (Cox, 1969: 191). Finally, in 1993 (a), Lee provided evidence that *Eunotosaurus* was a modified caseid synapsid, making any association with chelonians even more unlikely.

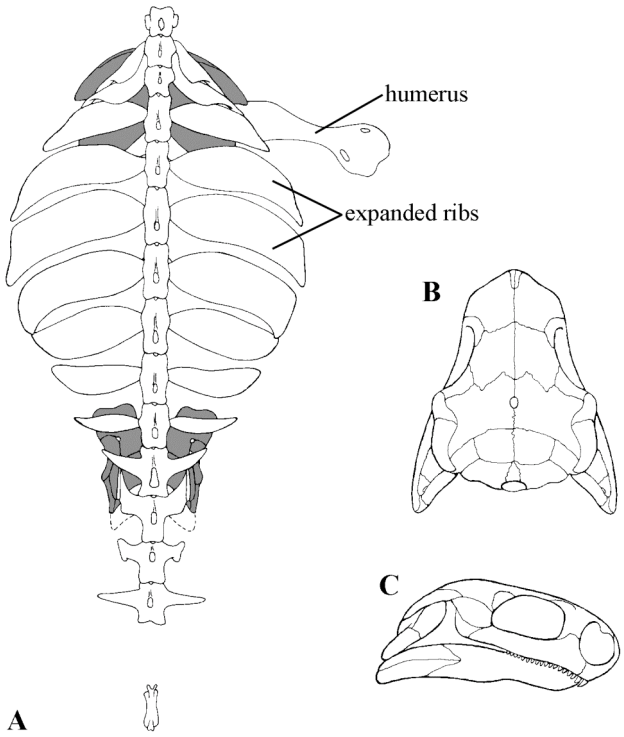


Figure 6.3. The purported turtle relative *Eunotosaurus africanus* from the Upper Permian of South Africa. (A) Postcranial skeleton in dorsal view. (B, C) Skull. Figure modified from Carroll (1988).

The second, and historically most popular, hypothesis for turtle origins is that they arose from an anapsid-skulled "parareptile". At its conception, this hypothesis was largely based on the fenestration of the temporal region of the amniote skull. Turtles, unlike all other living reptiles, have a completely enclosed temporal skull roof (Fig. 6.4), and as such were classified along with fossil taxa lacking fenestrae completely, or possessing just a single temporal opening (Baur, 1889; Cope, 1892; Osborn, 1903; Williston, 1917).

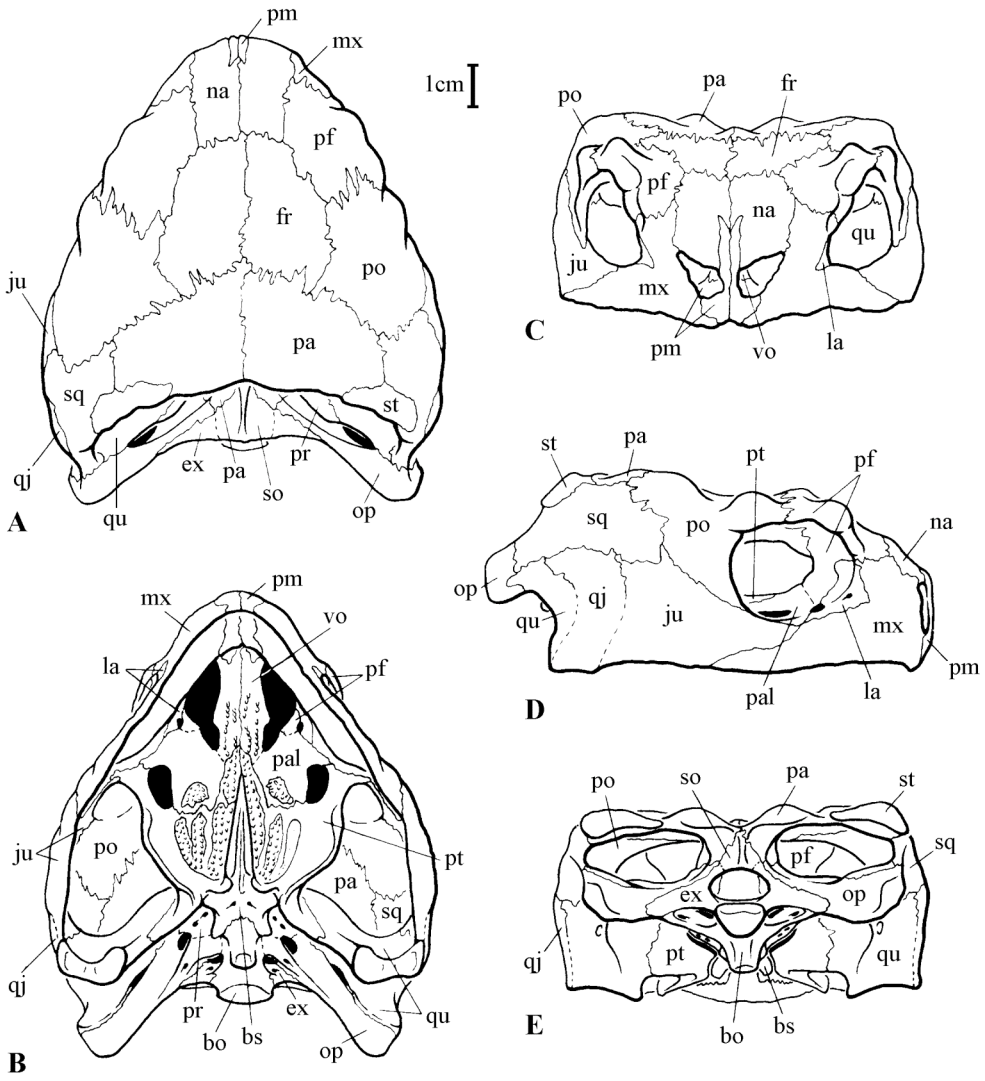


Figure 6.4. The skull of *Proganochelys quenstedti* in (A) dorsal, (B) ventral, (C) anterior, (D) lateral, and (E) occipital views. bo = basioccipital, bs = basisphenoid, ex = exoccipital, fr = frontal, ju = jugal, la = lacrimal, mx = maxilla, na = nasal, op = opisthotic, pa = parietal, pal = palatine, pf = prefrontal, pm = premaxilla, po = postorbital, pr = prootic, pt = pterygoid, qj = quadratojugal, qu = quadrate, so = supraoccipital, sq = squamosal, st = supratemporal, vo = vomer. Figure modified from Gaffney (1990).

For example, Cope (1892) linked turtles with cotylosaurs, and Williston (1917) included them in his new group, the Anapsida. Williston (1917) agreed with Baur (1889) that the turtle skull could not have originated from a reptile with a perforated temporal region. He stated that "chief reliance must be placed upon the skull structure, especially that of the cranial and temporal regions…[because]…these parts are the most conservative, and least liable to homoplastic duplication" (Williston, 1917: 420). Broom (1922; 1924) suggested that there were at least two other possibilities: that turtles derived from an ancestor with at least one temporal opening, such as plesiosaurs or lepidosaurs, where the openings had

been secondarily lost, or from an ancestor that had a single temporal opening and whose descendents lost the squamoso-pareital bar. He also pointed out (Broom, 1922; 1924) that only a few chelonians have a roofed skull, most having an open temporal region, and that *Hydromedusa* even possesses a completely enclosed temporal foramen (see Fig. 6.5).

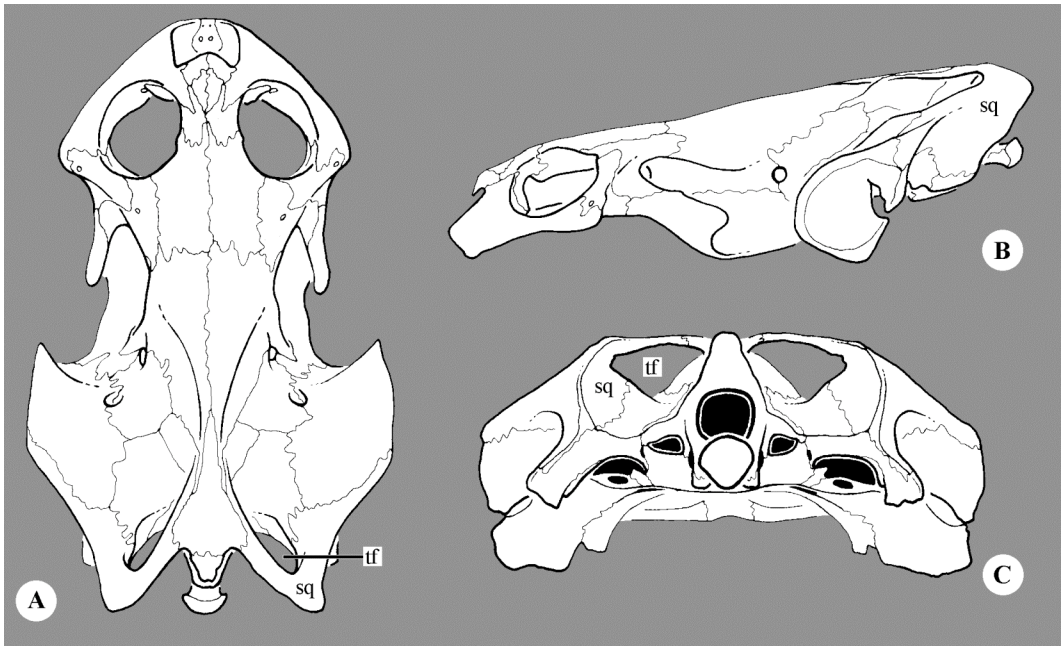


Figure 6.5. The skull of *Hydromedusa tectifera* showing a completely enclosed temporal foramen, which is probably the result of secondary growth of the squamosal. Skull shown in (A) dorsal, (B) lateral and (C) occipital views. sq = squamosal, tf = temporal foramen. Figure modified from Gaffney (1979a).

Olson (1947: 37) rejected Broom's alternate explanations and declared that in chelonians "there seems to be little doubt…that no true temporal fenestra has ever been developed, and that the broad opening in the temporal region of more advanced skulls has resulted from recession of the posterior margin of the squamosal. The closed opening, such as that in *Hydromedusa*, probably resulted from secondary growth of the squamosal, although this has not been clearly demonstrated". In 1946, Gregory suggested the Permian parareptiles, pareiasaurs (see Fig. 6.6), as possible candidates in the search for a turtle ancestor. "When I turned my attention to the pareiasaurs I was surprised to find that apart from their gigantic size they seemed on the whole to afford an excellent starting point for the chelonian line, both in their general construction and in many features of the skull, vertebrae, ribs, girdles, limb bones, hands and feet" (Gregory, 1946: 86).

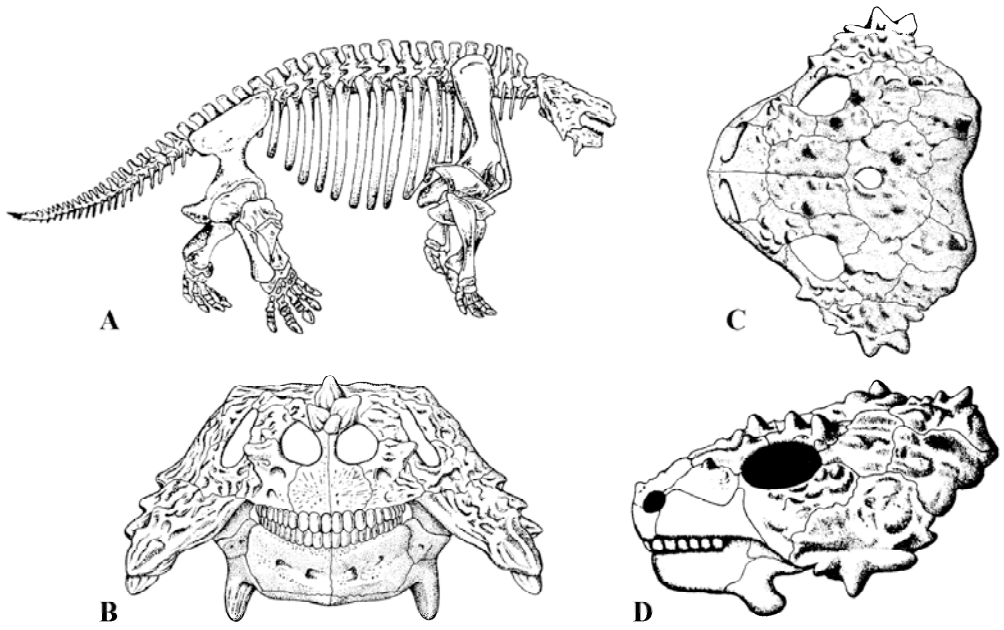


Figure 6.6. Pareiasaur morphology. (A) Lateral reconstruction and (B) anterior view of the skull of *Scutosaurus*. (C) Dorsal and (D) lateral views of the skull of *Pareiasaurus*. A, B modified from Carroll (1988), C, D, modified from Romer (1956).

At this time pareiasaurs were generally grouped with procolophonids and the basal amniote diadectids and seymouriids (Goodrich, 1916; Olson, 1947). In 1964 Romer split this group so that pareiasaurs and procolophonids were separated from diadectids and seymouriids, which were themselves no longer considered putative turtle ancestors (Olson, 1965). Turtles were left grouped with the pareiasaurs and procolophonids, although Olson did not discuss the interrelationships of these three groups (Olson, 1965). Clark and Carroll (1973) championed a further group of anapsid reptiles as possible turtle ancestors. During work on captorhinids they were struck by the similarities of the occiput of this group with that of *Proganochelys*, the most primitive chelonian. They also claimed that pareiasaurs and procolophonids were "particularly inappropriate ancestors for turtles, because they have a totally different configuration of the occiput" (Clark and Carroll, 1973: 403). The captorhinomorph hypothesis was supported by Gauthier *et al*. (1988a; 1988b) both in a systematic analysis of 112 characters derived from the literature for 14 taxa (Gauthier *et al*., 1988a) and in a reanalysis of previous data complied by Gardiner (1982) with the addition of data from fossil taxa (Gauthier *et al*., 1988b). They noted (Gauthier *et al*., 1988a) that parareptiles (in which they included pareiasaurs, millerettids, procolophonids and mesosaurs) accounted for a disproportionate share of the homoplasy. Within three years, turtle origins within the Anapsida were on the move again. Reisz and Laurin (1991;

Laurin and Reisz, 1993) presented evidence from new specimens of *Owenetta*, the oldest known procolophonid, and provided ten synapomorphies uniting procolophonids with turtles. In 1995, a systematic analysis of 124 characters for 13 taxa (including two diapsid groups) by the same authors (Laurin and Reisz, 1995) again supported a procolophonid sister-group for turtles. However, throughout this time Lee (1993a; 1993b; 1994) began endorsing not only a sister-group relationship between turtles and pareiasaurs, but even claimed that pareiasaurs were the paraphyletic proximal outgroup to turtles (Lee, 1993b). While Laurin and Reisz were producing their cladistic analysis supporting procolophonids as the sister-group of turtles, Lee (1995; 1996a) produced his own analyses supporting the pareiasaur-turtle hypothesis.

In recent years a controversial, yet strong, contender for the resolution of turtle relationships has appeared. This hypothesis, championed of late by Olivier Rieppel among others, suggests that turtles are in fact diapsids like all other living reptiles, and that the anapsid condition of their skull is due to the secondary loss of temporal fenestrae, as proposed by Goodrich (1916). That turtles may have a diapsid ancestor is, however, not a new idea. As stated by Gregory (1946: 281): "…it has been generally and indefinitely assumed from Buckland's time to the present that the plesiosaurs were related to the turtles". Jaekel (1902) was the first to propose placodonts as turtle ancestors, based on similarities in the skull and the presence of a carapace, a view supported by Broom (1922) among others. However, at this time, in the early twentieth century, neither placodonts nor plesiosaurs were considered diapsids as they generally are today. For example, in his reclassification of the reptiles, Osborn (1903) included both turtles and placodonts within his Synapsida, and Broom (1924) included them as a separate group of their own outside diapsids, synapsids and cotylosaurs. Goodrich (1916) was the first to propose that turtles were sauropsidans (his group containing living reptiles and birds) based on the hooked fifth metatarsal (see Fig. 6.7) and heart morphology shared by the two groups.

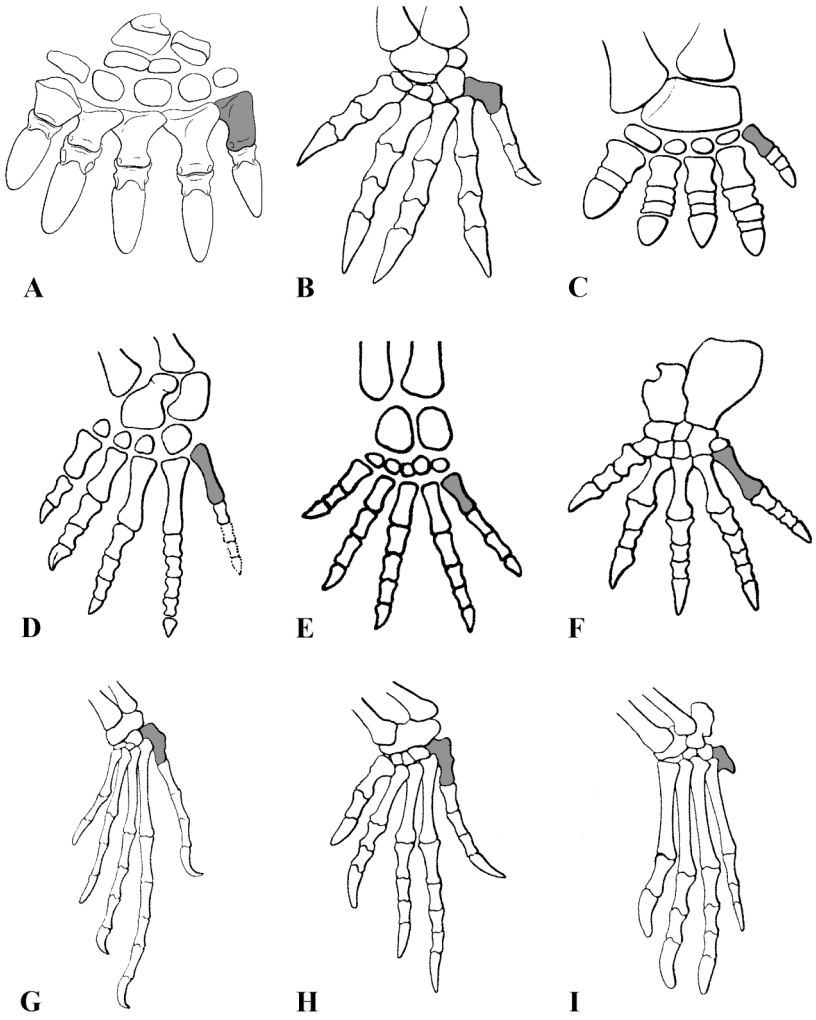


Figure 6.7. Left feet of (A) The Triassic turtle *Proganochelys quenstedti*, (B) The extant snapping turtle, *Chelydra serpentina*, (C) The pareiasaur, *Bradysaurus baini* (anapsid), (D) *Procolophon pricei* (anapsid), (E) The pelycosaur, *Haptodus longicaudatus* (synapsid), (F) *Edaphosaurus* sp. (synapsid), (G) *Iguana tuberculata* (diapsid), (H) *Sphenodon punctatus* (diapsid), and (I) *Caiman sclerops* (diapsid). Metatarsal V, which is hooked in turtles and diapsids, is shaded. A modified from Gaffney (1990), B, E, F, G, H and I modified from Goodrich (1916), C and D modified from Romer (1956).

Lakjer (1926), led by a study of jaw musculature, was the first to actually place turtles within the clade Diapsida, although Broom (1924) had earlier proposed a lepidosaurian, *Sphenodon*-like ancestor. Following this, the diapsid hypothesis became unpopular for over fifty years until a spate of new papers were published, which examined data derived from sources wider than just osteology. Løvtrup (1977) proposed a turtle-crocodile link based on evidence from osteology, bone and soft tissue histology, penis morphology, eggshell structure and blood proteins. Many of his synapomorphies were later shown to be present in unrelated taxa or were flawed because of an inappropriate choice of outgroups (Gardiner, 1982), but further studies on similar characters also supported a turtle origin

either within or as the sister-group to the Diapsida (Gardiner, 1982; Gaffney and Meylan, 1988; Gauthier *et al*., 1988a; 1988b). Evidence from purely osteological analyses also began to be presented that weakly supported the diapsid hypothesis (Gaffney and Meylan, 1988; Rieppel, 1990; 1995). In a study of skeleton formation in reptiles, Rieppel (1995: 298) criticised many previous analyses, because they "were predicated upon the assumption that the Testudines are, in fact, anapsids, and taxa for comparison were chosen accordingly". He suggested that many of the characters used to support sister-group relationships of turtles with anapsid taxa also occur in some diapsid taxa not included in the analyses. He presented an analysis of 70 characters for 15 taxa that placed turtles as the sister-group of lepidosaurs.

Since 1995 the turtle origins debate has heated up. In 1996, Rieppel and deBraga produced a much larger analysis of osteological characters of amniotes than had been attempted before. Their data matrix comprised 168 characters for 33 taxa, and included anapsids, diapsids, synapsids and basal amniotes, using the Diadectomorpha and Seymouriidae as outgroups. They claimed that their results "robustly" supported the diapsid affinities of turtles, placing them as the sister-group of the sauropterygians (placodonts + nothosaurs and plesiosaurs). However, in reply, Wilkinson *et al*. (1997) showed that this claim of robust support was unfounded. Using Templeton tests (Templeton, 1983), Wilkinson *et al.* (1997) demonstrated that an anapsid placement of turtles was not a significantly less parsimonious fit to the data of Rieppel and deBraga (1996). Despite this, the data matrix produced by Rieppel and deBraga (1996) has since become the crux of the osteological evidence for the position of the turtles. Various rescorings of the same 168 characters have been shown to place turtles both within the Anapsida (Lee, 1997b; 2001; Motani *et al*., 1998) or the Dipsida (Rieppel and Reisz, 1999). In the most recent of this string of reanalyses, Lee (2001) compiled a combined matrix of 176 osteological characters (the 168 characters from Rieppel and deBraga (1996) plus eight new characters), 40 soft tissue characters derived from Gauthier *et al.* (1988a) and 1783 alignable sites taken from the complete 12S and 16S mitochondrial rRNA sequences published by Zardoya and Meyer (1998). Lee found that the anapsid hypothesis (placing turtles as the sister-group to pareiasaurs) was a significantly better fit to the osteological data, and the combined morphological and molecular data than the diapsid hypothesis according to Templeton tests (Lee, 2001), despite the fact that the molecular evidence alone supported a diapsid origin (Zardoya and Meyer, 1998).

Virtually all analyses of molecular data have shown turtles to nest within the living Diapsida, using mitochondrial genes (Zardoya and Meyer, 1998; Kumazawa and Nishida, 1999; Cao *et al*., 2000; Rest *et al*., 2003) and nuclear genes (Bishop and Friday, 1988; Hedges and Poling, 1999; Mannen and Li, 1999; Cao *et al*., 2000). Most of these analyses have shown turtles to be close relatives of the archosaurs (crocodiles and birds), with lepidosaurs outside. However, there appears to be some disagreement between the results of analyses of mitochondrial genes, which tend to position turtles as the sister-group of all other living archosaurs (e.g. Zardoya and Meyer, 1998; Kumazawa and Nishida, 1999; Zardoya and Meyer, 2001; Rest *et al*., 2003), and nuclear genes, which tend to position them inside the Archosauria as the sister-group of crocodilians (e.g. Hedges and Poling, 1999; Cao *et al*., 2000; and see Meyer and Zardoya, 2003). Based on their analysis of nuclear genes, Hedges and Poling (1999) suggested that aetosaurs may be a possible relative of turtles, because they are close relatives of crocodiles, but also possess a number of morphological similarities to turtles that crocodiles do not, including beak-like jaws with reduced teeth, body armour similar in structure to the turtle carapace (cf. Cox, 1969) and similarities between the neck spines of the aetosaur *Desmatosuchus* (Fig. 6.8) and the turtle *Proganochelys*.

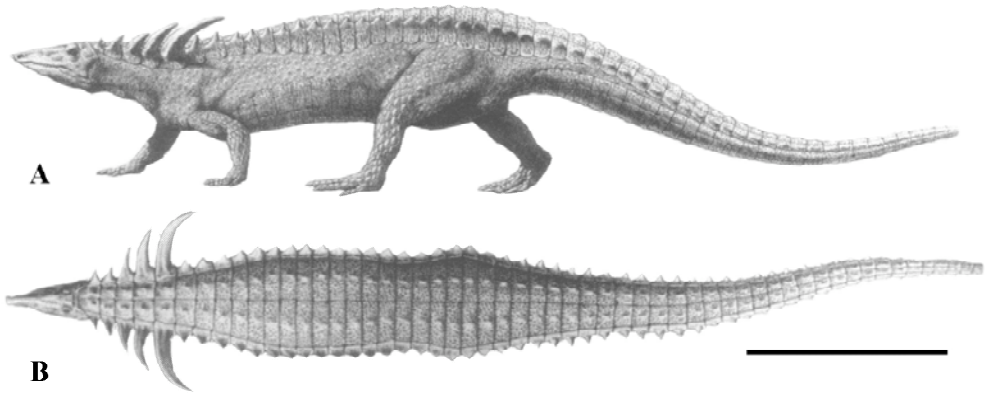


Figure 6.8. A reconstruction of the Triassic aetosaur, *Desmatosuchus haploceras*, in (A) left lateral and (B) dorsal views, showing the carapace made up of dermal scutes and the cervical spines that have been suggested to be homologous to similar structures in *Proganochelys*. Scale bar = 1m. Figure modified from Long and Murray (1995).

Opponents of the diapsid hypothesis could argue that (other than turtles) there are no living anapsid reptiles from which to extract DNA for sequencing, so it is not possible for turtles to plot with other anapsids in molecular analyses. This is reminiscent of Rieppel's (1995) complaint about early cladistic analyses of turtle positions using morphological data that included no diapsids.

6.2.2 Data Exploration Methods

The debate over turtle affinities is an ideal case study for the application of data exploration techniques, because it possesses a number of interesting properties. First, the position of turtles is of great interest and importance in the resolution of the phylogeny of amniotes, and the greater goal of the production of an entire tree of life. Second, the debate has an explicit area of conflict and the two major alternative hypotheses are clear-cut. Third, the two alternative hypotheses disagree only in the positioning of turtles. The positions of all other taxa are constant in the most parsimonious phylogenies produced by the two datasets (Wilkinson *et al*. 1997). This reduces the number of variables and simplifies the interpretation of results.

The data matrices published in the two most recent osteological analyses concerning the affinities of turtles, which provide support for both the diapsid (Rieppel and Reisz, 1999) and anapsid (Lee, 2001) hypotheses, were explored in detail. Both datasets were first reanalysed using parsimony methods to ensure that the published results could be replicated. All parsimony analyses were carried out using PAUP 4.10b (Swofford, 1999) by employing heuristic searches with 100 random addition sequences and TBR branch swapping.

Clade support was quantified using the bootstrap (Felsenstein, 1985) and decay indices (Bremer, 1988). Bootstrap proportions (BPs) were based on analyses of 1,000 replicates. Decay indices were calculated for each node using reverse topological constraints. Leaf stability (Thorley and Wilkinson, 1999) using BPs was calculated in Radcon (Thorley and Page, 2000). Leaf stability is based on the idea that stable taxa (leaves) will occur in stable triplets or quartets. Three measures of leaf stability are presented, (1) maximum, which is the BP of the resolution of the quartet which occurs in the greatest proportion of bootstrap trees, (2) difference, which is the difference between the BPs of the resolutions of the quartet which occur in the greatest and second greatest proportion of bootstrap trees and (3) entropy, which takes into account the BPs of all three possible resolutions of the quartet and assesses their deviation from the expectation that they are equal (Wilkinson pers. comm. 2001). Backbone constraints were applied to find the shortest trees for each dataset with turtles positioned within the Anapsida and Diapsida, and Templeton tests (Templeton, 1983) employed to test the null hypothesis that there is no significant difference between the fit to the data of the best trees with turtles within the Anapsida and Diapsida. Sequential taxon removal and reverse successive weighting

(Trueman, 1998) were used, as illustrated by Rieppel (2000b), in an attempt to identify the presence of any subsignals in the datasets. Sequential taxon removal involves removing the sister-group of a problematic taxon (in this case turtles). If the nesting of the group is strong, its position should not change relative to the remaining taxa when its sister-taxa are removed. If, however, removal of the sister-taxon leads to a drastic repositioning of the taxon in question, it is possible that its position in the initial tree was caused by convergence rather than a signal representing the true affinities of the group. The greater the number of successive sister-taxa that can be removed from a tree without changing the position of the taxon in question relative to the remaining taxa, the stronger its position in the tree appears to be. Reverse successive approximations character weighting (RSACW, Trueman, 1998) is, as the name suggests, the converse of the usual method of successive approximations character weighting (SACW, Farris, 1969). Generally in SACW analyses, a parsimony analysis is first carried out. Characters that fit well to the resulting tree are assigned high weights, while those that fit poorly are assigned low weights, usually based on their consistency index (CI), retention index (RI) or rescaled consistency index (RC), which is the product of the CI and RI. SACW is employed in an attempt to increase resolution in poorly resolved consensus trees, based on the assumption that characters that fit best to a tree are the best characters for resolving controversial areas of that tree. The weighted dataset is then reanalysed and the process repeated until a tree is found in which the weights are the same as the previous weighting, so that the analysis has reached stability. RSACW, conversely, downweights characters that fit the most parsimonious trees well, and upweights those that fit poorly, often by simply excluding characters that show no homoplasy. In theory, this means that subsignals in the data that are hidden in the most parsimonious tree may be shown when the weighted data are reanalysed. Unlike SACW, RSACW cannot be repeated until the resulting tree stabilises (Trueman, 1998). Instead, each reweighting leads to a new subsignal in the data being revealed in a hierarchical way, so that the most parsimonious signal appears first, and as the data are successively weighted the signals shown are less strong in the original analysis. The process is repeated until the tree produced by analysis of the reweighted data breaks down completely to an unresolved bush, indicating that little signal remains in the reweighted data.

6.2.2.1 Consensus Matrix

The similarity of character constructions used by Rieppel and Reisz (1999) and Lee (2001) makes comparison of their interpretations of morphology relatively simple. Character states assigned to taxa were compared, and all scoring differences recorded. In the case of turtles, disagreements over morphology are not unexpected given the highly derived morphology of the group. However, across all taxa only 77 scoring differences were identified in the 168 characters shared by the two matrices, equating to less than 2% of the total data. Thus, 98% of the data are agreed upon by the opposing authors, and might therefore be assumed less likely to be subject to human error in assessments of morphology. It is logical that if two workers code a taxon as possessing mutually exclusive states for a character, then at least one of those interpretations of morphology must be incorrect. From our list of scoring differences, it was possible to create a new data matrix in which such conflicts were removed. In essence, this is a consensus of the original matrices, a seemingly logical method of exploration of conflicting data, yet as far as I know this is the first true application of such a technique. In order to facilitate production of the consensus matrix, a number of alterations had to be made to the matrices of Rieppel and Reisz (1999) and Lee (2001) (referred to hereafter as the input matrices).

First, input matrices must only contain taxa that are either the same as, or independent of, taxa in the other input matrices. So, a taxon in one input matrix cannot be subsumed into a more inclusive taxon in a second matrix. However, slightly different taxon lists were used by Rieppel and Reisz (1999) and Lee (2001), because Lee (2001) combined groups of taxa present in the analysis of Rieppel and Reisz (1999) into single, more inclusive, taxa. For example, Rieppel and Reisz (1999), coded three pareiasaur taxa (*Anthodon*, *Bradysaurus* and *Scutosaurus*), while Lee (2001) coded a single Pareiasauridae. In order to construct the consensus matrix, taxa present only in the matrix of Rieppel and Reisz (1999) were combined to produce taxa equivalent to those in Lee (2001). In the above example, the three pareiasaur taxa were combined into a single Pareiasauridae equivalent to that coded by Lee (2001) (See Table 6.1 for a list of taxa that were combined).

Rieppel and Reisz (1999)	**Lee (2001)**
Seymouriidae + Diadectomorpha	Outgroup = Diadectomorpha only, Seymouriidae excluded
Owenetta + *Procolophon*	Procolophonoidea
Bradysaurus + *Scutosaurus* + *Anthodon*	Pareiasauridae
Placodus + *Cyamodus*	Placodontia

Table 6.1. Taxa included in the analysis of Rieppel and Reisz (1999) and the equivalent, more inclusive taxa included in the analysis of Lee (2001).

If taxa to be combined in this way exhibited different states for a character, then the combined taxon was scored as possessing both states (polymorphism). Similarly, if one taxon was scored as unknown for a character, but a second taxon with which it was to be combined had been assigned a score, then the more informative score was assigned to the combined taxon.

Similarly, all characters in the input matrices must also be either identical to or independent of characters in the other input matrices. Therefore, Rieppel and Reisz's (1999) character 51, which was split into two characters (51 and 169) by Lee (2001), was reconstructed into a single complex character to facilitate production of the consensus matrix. All characters not present in all input matrices were excluded from the consensus so that all characters included in the consensus had been subjected to the scrutiny of the two rival sets of authors (Rieppel and Reisz and Lee). All of the characters excluded during this process had been added by Lee (2001), and therefore it is not possible to know if Rieppel and Reisz (1999) agree with the scorings assigned to the characters. Alternative, less strict (semi-strict), methods of consensus matrix construction might include such characters on the basis that they are not contradicted in any of the other input matrices.

We employed a strict consensus method, in which all scoring differences between input matrices were replaced by missing data (?). Semi-strict methods of consensus matrix production are possible. For example, in cases where one author assigns a state to a taxon for a character that a second author has scored as missing, our strict method retains the uncertain scoring, whereas a semi-strict approach might choose to keep the more informative scoring, since it is not contradicted. Similarly, for multistate characters, opposing authors may agree that a taxon does not possess one or more states of the character, but may not be able to agree which of the alternatives it does possess. Using our strict consensus method, the data point would be scored as ? in the consensus matrix. A semi-strict method may score such disagreements as polymorphic (uncertain) for all states other than those that the authors agree are not possessed by the taxon. Majority rule

consensus matrices can also be envisaged, in which the score assigned in the consensus matrix is that which is present in the majority of studies, or supported by the majority of authors in the field.

6.2.2.2 Compatibility Methods

In an attempt to shed further light on the reasons for the deadlock between the anapsid and diapsid hypotheses, a number of the compatibility tests and boildown methods described in Chapter 4 were employed. These tests aim to identify and remove characters that cause intra-matrix conflict, and were used as a means of exploring subsignals held within data matrices that have been used as evidence in favour of both anapsid (Lee, 2001) and diapsid (Rieppel and Reisz, 1999) turtle affinities. Removal of incompatible characters from the data is effectively a form of *a posteriori* character weighting and may provide evidence of underlying support for one or other of the competing hypotheses. Differential weighting of characters is not widely used, because of a general scepticism of subjectivity in *a priori* weighting methods. However, methods in which *a posteriori* weighting is applied to characters in a justifiable and objective way, such as here, should not be totally disregarded. They may constitute an important means of data exploration that could be especially useful in situations where the data appear to contain multiple strong signals. The compatibility methods employed here may also be questioned by many cladists because they lead to the removal of data from the matrix. Some workers believe that the best way to obtain a true phylogeny is to include all available data and rely on the parsimony method to pick out signal from any noise and subsignals. However, in cases where parsimony fails to do this, the use of other methods to attempt to identify and remove noise in order to help parsimony find a resolution seems reasonable.

6.3 Results

6.3.1 Rieppel and Reisz (1999)

6.3.1.1 Reanalysis

Attempts to reproduce the results published by Rieppel and Reisz (1999) by analysing their published matrix failed, suggesting typographic errors may be present in the published matrix. Therefore, a corrected version of the matrix was obtained in electronic form from Olivier Rieppel. Analysis of this corrected version of the data

produced the results reported by Rieppel and Reisz (1999). Two MPTs were found (L=792, CI=0.505, RI=0.695) that differ from each other only in the relative positions of the Archosauriformes and Prolacertiformes. The strict consensus of these trees is shown on the left in figure 6.9. Average bootstrap and decay values for all nodes were 64.2 and 4.3 respectively, with many of the groupings within the Diapsida having low support. Turtles plotted within the Diapsida, as the sister-group of the aquatic sauropterygians (*Placodus*, *Cyamodus* and the Eosauropterygia). However, as shown by Wilkinson *et al*. (1997), and re-established here, Templeton tests illustrate that this diapsid placement is not a significantly better fit to the data (Templeton p-values range from 0.4697 to 0.4880) than the most parsimonious anapsid placement for turtles, which positions turtles as the sister-group of the pareiasaurs, the topology championed by Lee (2001) among others.

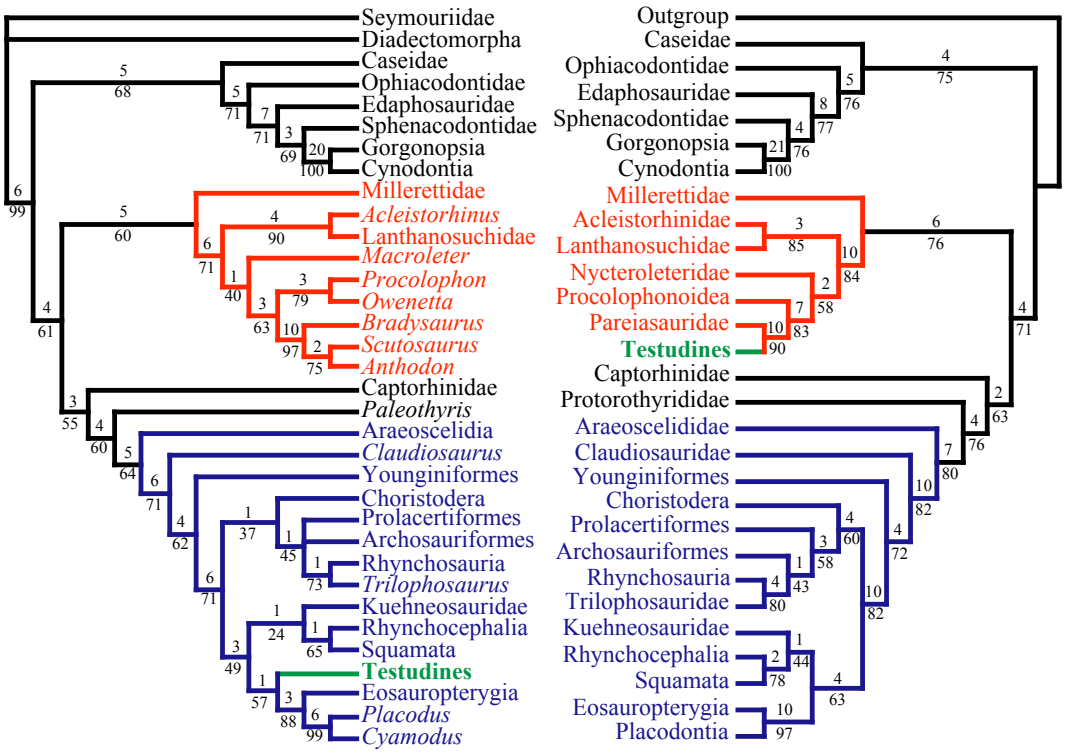


Figure 6.9. Strict consensus of the two MPTs for the data of Rieppel and Reisz (1999) on the left, and the single MPT for the osteological data of Lee (2001) on the right. Numbers above nodes represent decay values, and those below nodes represent bootstrap proportions. The Anapsida are highlighted in red, the Diapsida in blue and turtles in green.

Leaf stability analysis of the bootstrap trees showed that turtles, cynodontians and gorgonopsians were the least stable taxa (See Table 6.2), although due to the stability in most of the tree, all taxa had relatively high values.

Leaf	Maximum	Difference	Entropy
Younginiformes	0.9194	0.8531	0.7687
Claudiosaurus	0.9188	0.8524	0.7660
Bradysaurus	0.9124	0.8443	0.7579
Scutosaurus	0.9124	0.8443	0.7579
Anthodon	0.9124	0.8442	0.7578
Ophiacodontidae	0.9118	0.8396	0.7623
Edaphosauridae	0.9118	0.8396	0.7623
Sphenacodontidae	0.9118	0.8396	0.7626
Araeoscelidia	0.9088	0.8310	0.7455
Placodus	0.9088	0.8398	0.7607
Cyamodus	0.9088	0.8398	0.7607
Caseidae	0.9080	0.8331	0.7515
Eosauropterygia	0.9077	0.8377	0.7561
Procolophon	0.9073	0.8345	0.7397
Owenetta	0.9070	0.8342	0.7387
Seymouriidae	0.9060	0.8308	0.7501
Diadectomorpha	0.9060	0.8308	0.7501
Paleothyris	0.9040	0.8241	0.7325
Acleistorhinus	0.8963	0.8179	0.7104
Lanthanosuchidae	0.8962	0.8180	0.7099
Squamata	0.8955	0.8160	0.7313
Macroleter	0.8936	0.8137	0.7122
Rhynchocephalia	0.8929	0.8135	0.7261
Average	0.8929	0.8094	0.7176
Captorhinidae	0.8916	0.8048	0.7038
Trilophosaurus	0.8868	0.8001	0.7153
Rhynchosauria	0.8854	0.7950	0.7182
Millerettidae	0.8834	0.7971	0.6805
Archosauriformes	0.8831	0.7929	0.7092
Prolacertiformes	0.8786	0.7862	0.7047
Kuehneosauridae	0.8756	0.7875	0.6915
Choristodera	0.8675	0.7701	0.6901
Gorgonopsia	0.8179	0.6784	0.5396
Cynodontia	0.8175	0.6777	0.5391
Testudines	**0.8119**	**0.6566**	**0.5354**

Table 6.2. Maximum, difference and entropy leaf stability measures for the bootstrap trees from the analysis of the data of Rieppel and Reisz (1999). Taxa are listed in order of decreasing stability. Turtles are highlighted in bold.

6.3.1.2 Sequential Sister-Group Removal

Removal of the sauropterygians, which were the sister-group of turtles in the initial analysis, did not change the relative position of the turtles, which remained within the Diapsida. In the new tree, turtles were the sister-group of the clade comprising the

Kuehneosauridae, Squamata and Rhynchocephalia. Removal of this new sister-group caused turtles to move to the Anapsida, where they formed a clade with the pareiasaurs (*Anthodon*, *Bradysaurus* and *Scutosaurus*). Removal of the pareiasaurs moved turtles back within the Diapsida, as the sister-group of the Rhynchosauria. Continuing the process they then became sister-group of *Trilophosaurus*, followed by *Procolophon* back in the Anapsida, before the strict consensus was collapsed, with turtles plotting with both the Archosauriformes and the synapsid Gorgonopsia and Cynodonita. By this time, however, 12 taxa had been removed from the analysis.

6.3.1.3 RSACW

Figure 6.10 shows the results of reverse successive approximations character weighting (RSACW) of Rieppel and Reisz's (1999) data. In all trees the turtles nested within the Diapsida. The first RSACW tree (Fig. 6.10a) differed from the strict consensus tree of all data in that the synapsid Gorgonopsia and Cynodontia moved into the Diapsida.

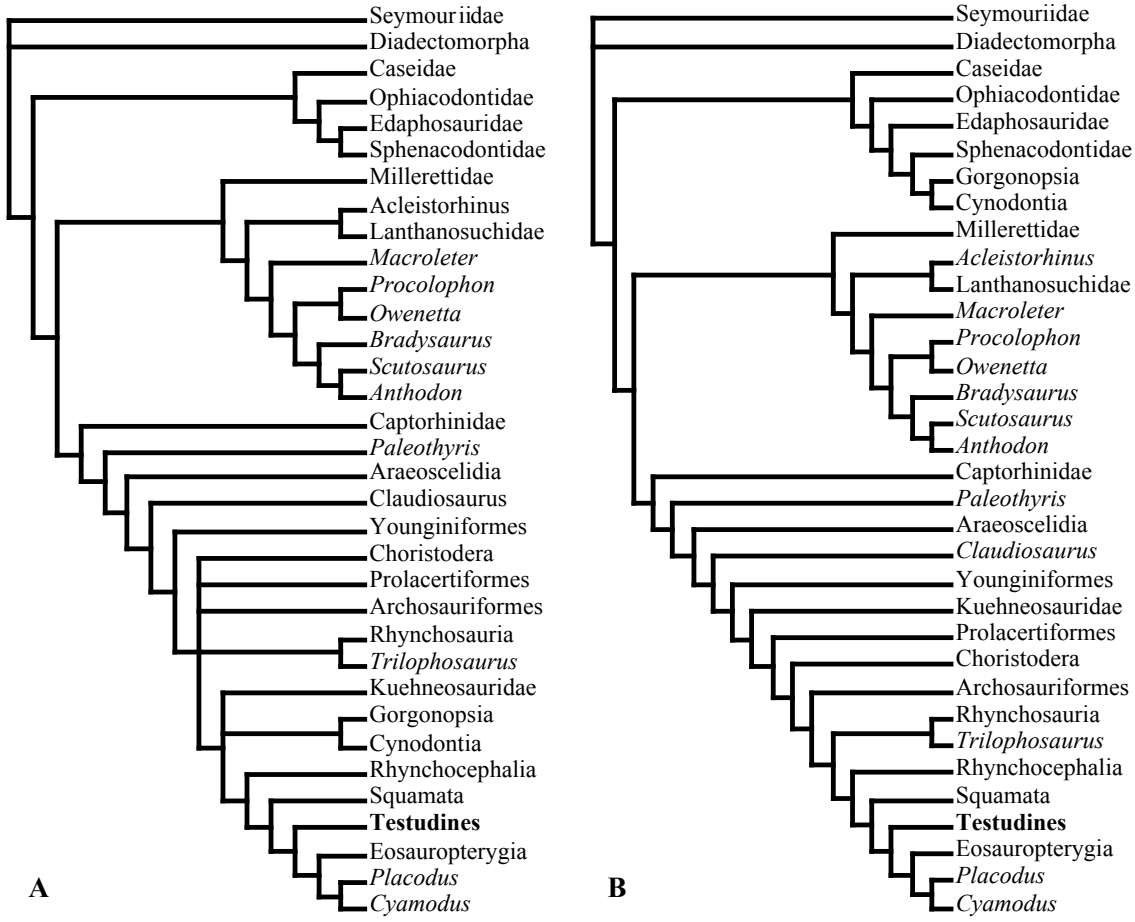


Figure 6.10. (A) First and (B) second RSACW trees of the analysis of the data of Rieppel and Reisz (1999).

In the second RSACW tree (Fig. 6.10b), the only change from the consensus tree of all data was in the resolution of the Diapsida. No remaining characters had a CI of one on this tree, so the process was halted.

6.3.1.4 Compatibility Methods

6.3.1.4.1 Taxon Jacknife Analysis

The matrix of Rieppel and Reisz (1999) contains 8022 pairwise incompatibilities, with five of the 168 characters being uninformative (Note that due to the replacement of polymorphic scorings with missing data in the computer program *Boildown*, the number of uninformative characters and the percentage of missing data in compatibility-based analyses in this study may be higher than the true value). Results of first- and second-order taxon jacknife analyses are shown in table 6.3. The taxa that are involved in the highest number of first- and second-order incompatibilities are the cynodontians, followed by the turtles and gorgonopsians, the same three taxa that were the least stable in leaf stability analysis. Along with these taxa, the randomisation test also showed the Lanthanosuchidae, Rhynchosauria, Kuehneosauridae, *Acleistorhinus* and *Cyamodus* (in order of decreasing p-value) to be no less incompatible than would be expected by chance alone. The Cynodontia and Testudines are actually involved in significantly more incompatibilities than would be expected by chance from a randomly permuted version of the taxon.

Taxon Excluded	Uninf.	% Missing	Caused	Expected	p	2nd Order
Cynodontia	6	3.0	730	391	0.999	1071
Testudines	**6**	**17.9**	**476**	**255**	**0.995**	**819**
Gorgonopsia	6	3.6	293	361	0.234	660
Rhynchosauria	5	5.4	262	330	0.243	623
Placodus	6	8.9	169	288	0.049	533
Choristodera	6	13.7	162	276	0.044	521
Procolophon	6	4.2	158	344	0.006	521
Claudiosaurus	5	10.7	152	282	0.024	513
Trilophosaurus	5	15.5	147	277	0.028	514
Rhynchocephalia	6	10.1	139	295	0.010	502
Squamata	7	17.9	134	261	0.019	495
Cyamodus	6	19.6	121	221	0.054	484
Kuehneosauridae	6	27.4	116	178	0.144	478
Scutosaurus	8	2.4	111	322	0.006	474
Edaphosauridae	5	12.5	99	275	0.003	462
Younginiformes	5	10.1	92	299	0.001	456
Eosauropterygia	5	25.6	88	195	0.018	453
Lanthanosuchidae	6	45.8	87	82	0.590	447
Archosauriformes	5	16.7	81	246	0.005	447
Captorhinidae	6	2.4	80	345	0.001	439
Araeoscelidia	5	4.8	79	329	0.001	442
Sphenacodontidae	5	1.8	78	348	0.001	445
Prolacertiformes	5	16.7	76	260	0.001	441
Anthodon	6	6.0	65	295	0.001	431
Macroleter	5	13.7	64	247	0.003	427
Bradysaurus	5	4.2	48	307	0.001	415
Acleistorhinus	5	49.4	46	75	0.159	409
Millerettidae	5	14.9	38	241	0.001	402
Ophiacodontidae	5	4.2	37	334	0.001	403
Owenetta	6	26.2	33	163	0.002	398
Caseidae	5	1.2	19	348	0.001	385
Seymouriidae	6	0.6	18	358	0.001	385
Paleothyris	5	6.0	11	303	0.001	376
Diadectomorpha	6	4.2	9	340	0.001	376

Table 6.3. Results of first- and second-order taxon jacknife analyses of the data of Rieppel and Reisz (1999). Taxa are listed in order of decreasing number of caused incompatibilities. Uninf. = number of uninformative characters in matrix when the taxon is removed, % missing = percentage of taxon character scorings that are missing, caused = change in the number of incompatibilities in the data when the taxon is removed, expected = average number of incompatibilities caused by randomised versions of the taxon, p = proportion of randomised versions of the taxon that cause the same or fewer incompatibilities than the actual taxon, 2nd order = average number of incompatibilities caused by all pairs of taxa including the taxon in question. Turtles are highlighted in bold.

6.3.1.4.2 Compatibility Analysis

The results of 'Number', CCSR and LQP compatibility analyses and subsequent boildowns of the data of Rieppel and Reisz (1999) are shown in table 6.4. The average (per character) number of incompatibilities, CCSR value and LQP value of the entire data were 47.8, 0.751 and 0.0631 respectively. The most incompatible character in the 'Number' analysis was 14, with 150 incompatibilities, the worst in the CCSR analysis was character 81, with a CCSR value of 1.253, and in the LQP analysis character 22 was the worst, with an LQP value of 0.7. In the LQP analysis, 122 characters showed an initial LQP value of less than 0.05, and were therefore a significantly better fit at the 5% level to the other data in the matrix than randomly permuted versions of the same character. A further 14 characters had LQP values between 0.05 and 0.1. 27 characters had LQP values greater than 0.1, and are therefore not significantly more compatible than a randomly permuted character.

Removal of all characters with an initial LQP value greater than 0.1 produced three MPTs (L=638, CI=0.502, RI=0.727) when the reduced data were analysed using parsimony. The strict consensus of these trees, which only differ in the position of the Kuehneosauridae, is shown in figure 6.11. Turtles still plot as the sister-group of the eosauropterygians and placodonts. Enforcing a backbone constraint positioning turtles in the Anapsida yields two best trees (L=647 CI=0.495 RI=0.719), both of which place turtles as the sister-group of the pareiasaurs. Templeton tests show that these trees are not a significantly worse fit to the data than the three MPTs (Templeton p-values range from 0.3376 to 0.3429).

Char.	No.	Fuzzy No.	CCSR	Fuzzy CCSR	LQP	Fuzzy LQP	No. Boil	Fuzzy No. Boil	CCSR Boil	Fuzzy CCSR Boil	LQP Boil	Fuzzy LQP Boil
1	86	168	0.619	0.561	**0.001**	**0.001**	-	108=	-	-	-	-
2	U	-	-	-	-	-	-	-	-	-	-	-
3	62	64	0.736	0.733	0.15	0.145	117=	120=	125	112	23	25
4	53	63	0.525	0.492	**0.015**	**0.019**	-	133=	-	-	82=	66=
5	68	69	0.812	0.802	0.225	0.198	-	-	-	37	17	19
6	142	413	0.932	0.756	**0.014**	**0.005**	5=	6	19	47	73=	74=
7	52	54	0.62	0.624	0.075	0.077	120=	124=	104	98	32	30
8	125	243	0.917	0.836	0.122	0.095	34=	41	26	20	27	28
9	42	42	0.847	0.831	0.4	0.414	135=	133=	99	24	9	8
10	122	257	0.917	1.006	0.14	0.502	38=	30	36	5	29	5
11	104	210	0.73	0.584	**0.002**	**0.001**	59	57	59	60	73=	83=
12	30	30	0.592	0.573	0.161	0.186	137=	133=	118	111	18	24
13	84	174	0.596	0.519	**0.001**	**0.001**	124	124=	137	108	111=	-
14	150	454	1.002	0.949	0.522	0.243	1	1	6	9	8	16
15	53	55	0.673	0.682	0.137	0.138	-	-	131	50	21	26
16	57	74	0.484	0.424	**0.004**	**0.001**	125=	127	-	113	89=	93
17	136	391	0.875	0.646	**0.001**	**0.001**	15	9=	49	80	-	-
18	78	107	0.597	0.44	**0.001**	**0.001**	103=	113	110	138	-	-
19	133	318	0.917	0.817	**0.039**	**0.033**	22=	26	24	23	54	40
20	109	218	0.791	0.644	**0.003**	**0.004**	86=	58=	78	59	89=	66=
21	10	10	0.199	0.191	**0.013**	**0.015**	-	-	-	-	50=	48
22	140	296	1.033	1.03	0.7	0.603	8=	25	4	4	1	4
23	121	205	0.985	0.989	0.337	0.421	38=	55	7	6	11	6
24	81	105	0.7	0.621	**0.013**	**0.014**	111=	112	92	71	62	65
25	135	341	0.955	0.968	0.157	0.313	16=	18	10	7	15	10
26	101	150	0.796	0.69	**0.032**	**0.021**	85	87	76	56	46=	50=
27	69	103	0.566	0.51	**0.003**	**0.003**	125=	133=	133	126	116	-
28	30	30	0.356	0.347	**0.006**	**0.009**	-	138=	138	119	67=	55
29	72	98	0.737	0.76	0.088	0.118	125=	124=	95	67	31	22
30	110	275	0.737	0.582	**0.001**	**0.001**	66	45=	97	101	114	-
31	112	213	0.802	0.654	**0.006**	**0.004**	75	80=	83	82	84=	88=
32	140	411	0.935	0.865	**0.039**	**0.027**	8=	5	13	16	39	34
33	99	120	0.957	0.917	0.361	0.324	64=	84=	9	13	10	11
34	82	122	0.698	0.711	**0.022**	**0.035**	-	-	-	64	60=	49
35	123	252	0.86	0.755	**0.015**	**0.019**	44	42=	60	62	82=	94
36	115	291	0.772	0.601	**0.001**	**0.001**	116	42=	126	105	-	-
37	120	236	0.911	0.937	0.174	0.261	45=	39=	29	10	19	13
38	103	252	0.677	0.469	**0.001**	**0.001**	102	76=	130	133	-	-
39	102	213	0.68	0.453	**0.001**	**0.001**	119	119	-	146	-	-
40	16	16	0.328	0.324	**0.026**	**0.04**	-	-	-	-	35	37=
41	118	318	0.785	0.635	**0.001**	**0.001**	67=	27	81	81	-	-
42	146	376	0.955	0.73	**0.047**	**0.001**	2	11	11	55	20	86=
43	35	39	0.337	0.289	**0.001**	**0.001**	130=	132	135	139	111=	-
44	54	59	0.524	0.451	**0.011**	**0.011**	-	-	-	124	56=	77
45	95	145	0.716	0.476	**0.002**	**0.001**	120=	120=	-	134	115	-
46	87	126	0.658	0.373	**0.001**	**0.001**	111=	114=	119	147	-	-
47	90	93	1.104	1.11	0.617	0.6	90	100=	2	2	4	3
48	111	246	0.736	0.495	**0.001**	**0.001**	91=	106=	105	-	-	-
49	126	270	0.838	0.559	**0.001**	**0.001**	34=	39=	62	85	99=	-
50	108	253	0.718	0.51	**0.001**	**0.001**	137=	138=	134	128	-	-
51	142	453	0.905	0.666	**0.001**	**0.001**	5=	2	32	63	89=	-
52	124	302	0.797	0.479	**0.001**	**0.001**	38=	38	84	132	-	-
53	112	289	0.736	0.536	**0.001**	**0.001**	62	52=	86	102	-	-
54	U	-	-	-	-	-	-	-	-	-	-	-
55	105	198	0.704	0.462	**0.001**	**0.001**	117=	100=	129	123	111=	-
56	115	257	0.82	0.781	**0.013**	**0.029**	61	33=	70	39	84=	63
57	89	175	0.645	0.575	**0.002**	**0.001**	-	-	-	120	110	102=
58	85	138	0.668	0.625	**0.009**	**0.01**	-	-	-	86=	89=	74=
59	90	157	0.642	0.498	**0.001**	**0.002**	130=	-	-	118	99=	102=
60	80	112	0.64	0.529	**0.009**	**0.003**	125=	114=	-	-	-	-
61	126	307	0.888	0.909	**0.042**	0.124	33	23	40	14	52	27

62	133	366	0.896	0.791	**0.006**	**0.011**	22=	15	35	34	66	45=
63	104	163	0.801	0.642	**0.01**	**0.002**	76	88=	53	53	50=	56
64	134	257	1.049	1.05	0.695	0.651	20=	35=	3	3	3	2
65	144	383	0.93	0.656	**0.007**	**0.001**	4	8	23	68	-	-
66	128	376	0.845	0.732	**0.001**	**0.003**	32	12	42	38	71	61
67	124	241	0.877	0.744	**0.012**	**0.018**	37	51	47	51	53	57=
68	74	102	0.554	0.394	**0.002**	**0.001**	106	103=	107	140	-	-
69	113	282	0.745	0.544	**0.001**	**0.001**	107=	70	124	121	-	-
70	96	184	0.662	0.477	**0.001**	**0.001**	100=	90=	109	137	-	-
71	28	28	0.504	0.506	0.161	0.172	-	-	-	100	16	20
72	121	282	0.846	0.775	**0.014**	**0.029**	47=	31=	55	31	76	50=
73	135	247	0.991	0.939	0.466	0.367	16=	37	8	12	6	12
74	135	417	0.87	0.714	**0.001**	**0.001**	16=	4	51	45	-	86=
75	107	258	0.709	0.509	**0.001**	**0.001**	60	54	69	84	-	-
76	141	428	0.931	0.808	**0.007**	**0.003**	7	3	20	27	99=	52
77	130	329	0.908	0.851	0.076	0.069	28=	22	34	18	41	33
78	137	341	0.896	0.678	**0.005**	**0.001**	14	24	33	70	-	-
79	139	403	0.929	0.858	**0.034**	**0.032**	11=	7	16	22	36	43
80	99	160	0.756	0.657	**0.009**	**0.008**	97	84=	120	104	-	-
81	65	65	1.228	1.233	0.672	0.699	125=	128=	1	1	2	1
82	132	272	0.9	0.639	**0.02**	**0.001**	27	44	31	66	72	95=
83	128	285	0.87	0.657	**0.004**	**0.002**	30=	35=	50	69	99=	95=
84	42	63	0.351	0.348	**0.001**	**0.001**	122=	120=	122	-	-	-
85	109	234	0.767	0.66	**0.001**	**0.004**	79=	56	79	61	89=	88=
86	133	305	0.899	0.668	**0.016**	**0.001**	22=	29	28	57	56=	95=
87	118	204	0.891	0.821	0.093	0.069	43	58=	39	35	40	41
88	140	347	0.949	0.804	0.087	**0.011**	8=	17	12	29	26	53=
89	85	118	0.687	0.583	**0.008**	**0.005**	107=	102	123	117	-	95=
90	85	138	0.668	0.625	**0.009**	**0.01**	-	-	-	86=	89=	74=
91	112	234	0.78	0.568	**0.002**	**0.001**	79=	78=	82	89	89=	-
92	76	134	0.555	0.444	**0.002**	**0.001**	103=	96=	100	91	-	-
93	U	-	-	-	-	-	-	-	-	-	-	-
94	30	33	0.376	0.347	**0.002**	**0.009**	-	-	-	-	99=	59
95	111	248	0.751	0.534	**0.001**	**0.001**	79=	80=	96	127	-	-
96	77	114	0.722	0.814	0.095	0.265	-	-	-	36	33	17
97	50	53	0.466	0.376	**0.008**	**0.004**	-	-	102	88	63	83=
98	48	48	0.91	0.908	0.444	0.438	-	-	18	11	5	7
99	112	227	0.776	0.596	**0.003**	**0.001**	67=	72	85	107	-	-
100	U	-	-	-	-	-	-	-	-	-	-	-
101	78	81	0.951	0.966	0.349	0.361	95	105	17	8	12	9
102	83	117	0.687	0.528	**0.008**	**0.001**	137=	138=	-	144	84=	95=
103	99	248	0.672	0.576	**0.001**	**0.001**	91=	65	112	110	-	-
104	98	118	0.866	0.716	0.097	**0.038**	79=	92=	66	79	64	60
105	94	176	0.688	0.616	**0.001**	**0.001**	107=	108=	113	129	99=	-
106	70	132	0.485	0.382	**0.001**	**0.001**	130=	128=	136	115	-	-
107	116	222	0.829	0.693	**0.015**	**0.005**	53	61	65	78	81	-
108	65	78	0.642	0.619	**0.028**	**0.026**	113=	114=	115	122	88	78=
109	134	369	0.894	0.763	**0.006**	**0.003**	16=	9=	37	33	99=	66=
110	103	198	0.72	0.503	**0.001**	**0.001**	77	74=	77	92	-	-
111	90	202	0.602	0.42	**0.001**	**0.001**	100=	88=	106	142	-	-
112	50	53	0.467	0.375	**0.004**	**0.006**	-	-	103	90	89=	71=
113	109	200	0.74	0.454	**0.001**	**0.001**	79=	94=	91	145	-	-
114	109	218	0.754	0.55	**0.001**	**0.001**	91=	106=	98	-	-	-
115	98	187	0.662	0.417	**0.001**	**0.001**	67=	73	73	97	-	-
116	120	235	0.875	0.803	0.053	**0.031**	47=	47	52	40	44=	47
117	139	354	0.936	0.805	**0.014**	**0.013**	11=	16	15	28	37	45=
118	65	69	0.582	0.459	**0.012**	**0.006**	107=	114=	89	95	67=	71=
119	70	125	0.509	0.347	**0.001**	**0.001**	-	-	-	-	-	-
120	116	315	0.773	0.618	**0.001**	**0.001**	67=	28	90	106	-	-
121	57	63	0.578	0.51	**0.019**	**0.01**	113=	120=	127	103	78	78=
122	55	55	0.652	0.625	0.072	0.054	122=	128=	108	96	34	29
123	134	216	1.014	0.855	0.486	0.083	20=	48=	5	25	7	31
124	126	345	0.852	0.799	**0.004**	**0.006**	34=	19	58	32	99=	78=
125	118	253	0.807	0.6	**0.001**	**0.001**	67=	83	94	130	99=	-

126	146	368	0.948	0.657	**0.034**	**0.001**	3	13	14	65	56=	95=
127	129	259	0.923	0.845	0.12	0.113	30=	33=	22	19	24	21
128	110	217	0.756	0.55	**0.001**	**0.001**	98	92=	121	135	-	-
129	59	69	0.586	0.54	**0.016**	**0.016**	113=	114=	111	116	56=	78=
130	129	275	0.872	0.596	**0.003**	**0.001**	28=	45=	44	77	99=	-
131	118	229	0.815	0.585	**0.002**	**0.001**	52	62	67	99	-	-
132	106	183	0.765	0.587	**0.001**	**0.001**	84	84=	80	94	99=	95=
133	118	238	0.835	0.725	**0.01**	**0.002**	51	52=	64	75	-	-
134	103	210	0.743	0.633	**0.003**	**0.006**	130=	128=	101	72	84=	-
135	78	79	0.851	0.818	0.294	0.284	78	98=	27	21	13	15
136	47	54	0.456	0.415	**0.004**	**0.003**	-	-	-	-	89=	-
137	120	216	0.807	0.467	**0.001**	**0.001**	47=	71	68	131	-	-
138	116	223	0.837	0.705	**0.019**	**0.014**	54=	48=	57	49	67=	66=
139	112	261	0.742	0.504	**0.001**	**0.001**	94	103=	114	143	-	-
140	115	213	0.849	0.739	**0.024**	**0.011**	54=	64	61	76	79=	66=
141	71	80	0.629	0.524	**0.02**	**0.022**	-	-	88	73	65	44
142	97	216	0.66	0.489	**0.001**	**0.001**	99	94=	117	-	-	-
143	121	197	0.918	0.792	0.09	**0.025**	45=	68	46	52	49	71=
144	98	199	0.67	0.464	**0.001**	**0.001**	96	80=	116	-	-	-
145	117	229	0.826	0.633	**0.002**	**0.001**	64=	96=	93	114	-	-
146	113	206	0.853	0.807	0.057	0.072	54=	60	63	58	48	39
147	122	222	0.916	0.853	0.12	0.082	38=	48=	43	30	30	32
148	3	3	0.047	0.047	**0.005**	**0.003**	-	-	-	-	73=	62
149	133	333	0.883	0.704	**0.005**	**0.003**	22=	21	38	54	89=	83=
150	115	200	0.913	0.908	0.193	0.302	47=	66	25	15	14	14
151	109	184	0.853	0.828	0.073	0.088	73=	67	74	41	44=	35=
152	41	41	0.551	0.536	**0.049**	**0.046**	135=	133=	128	136	46=	53=
153	82	174	0.566	0.443	**0.001**	**0.001**	88=	76=	87	93	-	-
154	101	179	0.723	0.552	**0.001**	**0.001**	103=	98=	139	141	-	-
155	115	194	0.842	0.662	**0.016**	**0.003**	63	74=	71	109	77	88=
156	11	11	0.222	0.222	**0.016**	**0.02**	-	-	-	-	67=	57=
157	97	131	0.821	0.745	0.078	**0.048**	86=	108=	56	43	28	35=
158	123	283	0.809	0.549	**0.001**	**0.001**	38=	31=	75	83	-	-
159	96	121	0.8	0.65	0.062	**0.024**	72	90=	45	48	38	42
160	95	189	0.667	0.495	**0.001**	**0.001**	130=	138=	132	125	-	-
161	105	166	0.817	0.708	**0.039**	**0.047**	73=	78=	54	44	43	37=
162	138	375	0.919	0.768	**0.01**	**0.001**	13	14	21	42	55	88=
163	133	335	0.899	0.742	**0.013**	**0.004**	22=	20	30	46	60=	88=
164	91	111	0.752	0.584	**0.037**	**0.023**	88=	108=	72	74	42	64
165	116	187	0.897	0.842	0.118	0.208	54=	63	48	26	25	18
166	U	-	-	-	-	-	-	-	-	-	-	-
167	6	6	0.096	0.096	**0.01**	**0.007**	137=	-	-	-	79=	78=
168	113	188	0.889	0.859	0.13	0.148	54=	69	41	17	22	23

Table 6.4. Table showing the results of compatibility tests and the order of removal in subsequent boildowns of each character in the data of Rieppel and Reisz (1999) using CCSR, "Number" and LQP methods with and without the fuzzy compatibility. C = constant character and U = uninformative character. LQP results that are significant at the 5% level are highlighted in bold.

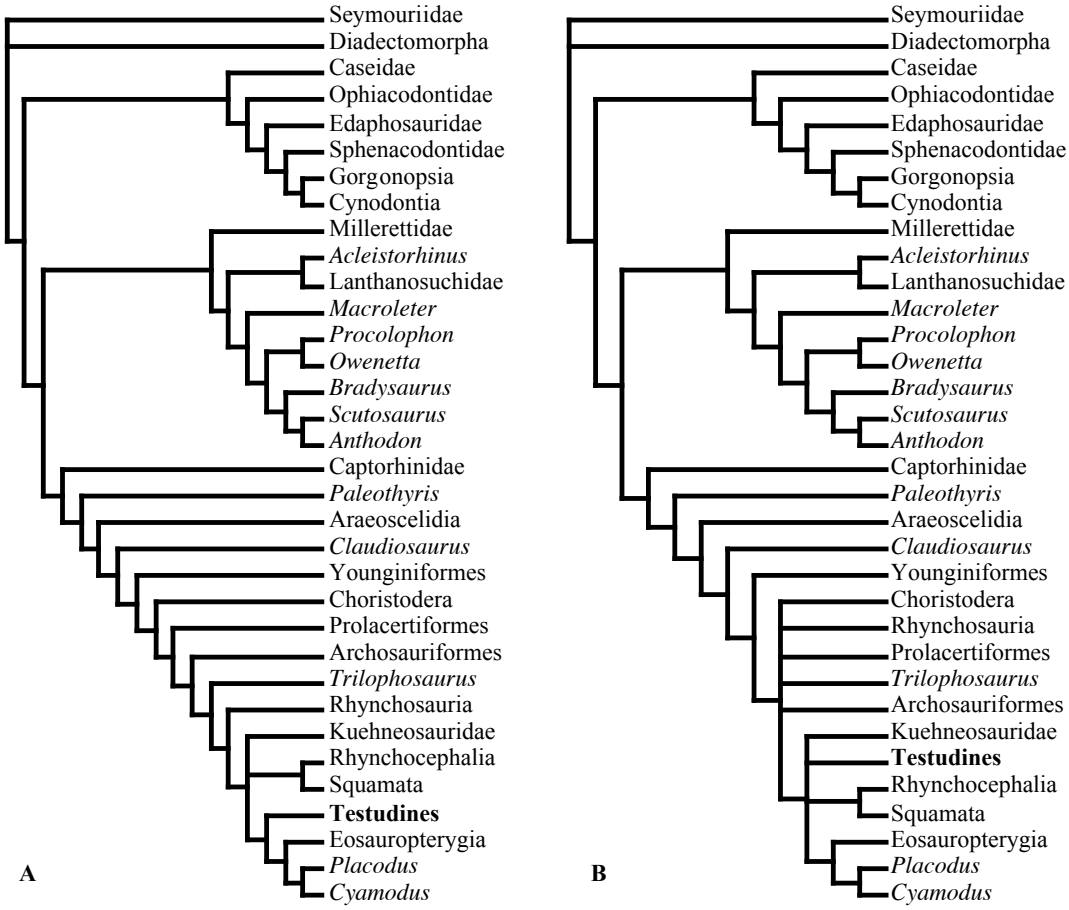


Figure 6.11. Strict consensus trees of the MPTs for the data of Rieppel and Reisz (1999) with characters excluded that are not significantly better than random at the (A) 10% and (B) 5% level of LQP tests.

Removal of all characters with an initial LQP value greater than 0.05 produced 12 MPTs (L=577, CI=0.508, RI=0.739). Resolution within the Diapsida was poor compared with that using the complete data. Turtles were nested deep within the Diapsida, but were not the sister-group to the Eosauropterygia and Placodontia. Enforcing a backbone constraint placing turtles within the anapsids produced eight MPTs (L=583 CI=0.503 RI=0.734), all of which placed turtles as the sister-group of the pareiasaurs. However, Templeton tests showed that these trees are not a significantly worse fit to the data than the three unconstrained MPTs (Templeton p-values range from 0.3961 to 0.4227).

6.3.1.4.3 Boildown Analysis

Results of Templeton tests of length differences between the most parsimonious trees with turtles constrained within the anapsids and diapsids, at each step of the boildown processes, are shown in figure 6.12. During the 'Number' boildown there was support, but not significantly so, for the diapsid hypothesis until around 115 characters had been removed. At this point, there was a sudden switch towards support for the anapsid

hypothesis. This support reached significance when 134 characters had been removed. During the CCSR and LQP boildowns, the diapsid hypothesis was more parsimonious throughout, and its support increased until 5% significance was reached when 57 and 70 characters were removed, respectively. In both cases, significance was lost and regained on a number of occasions as the boildown progressed further.

6.3.1.4.4 Fuzzy Compatibility

Results of 'Number', CCSR and LQP fuzzy compatibility tests and subsequent boildowns are shown in table 6.4. The results of Templeton tests at each stage of the boildowns are shown in figure 6.13. Generally the results using fuzzy compatibility were very similar to those using the non-fuzzy compatibility method. However, it appears that in all cases the initial trends towards diapsid support were stronger when the fuzzy method was used. The trend late on in the 'Number' analysis of a movement towards support for the anapsid hypothesis also occurred more quickly, reaching a p-value of 0 when 103 characters were removed and reaching significant support for the anapsid hypothesis with 132 removed. The CCSR analysis reached 5% significance in favour of the diapsid hypothesis when 50 characters had been removed, while the LQP boildown only reached significance when 85 characters were removed.

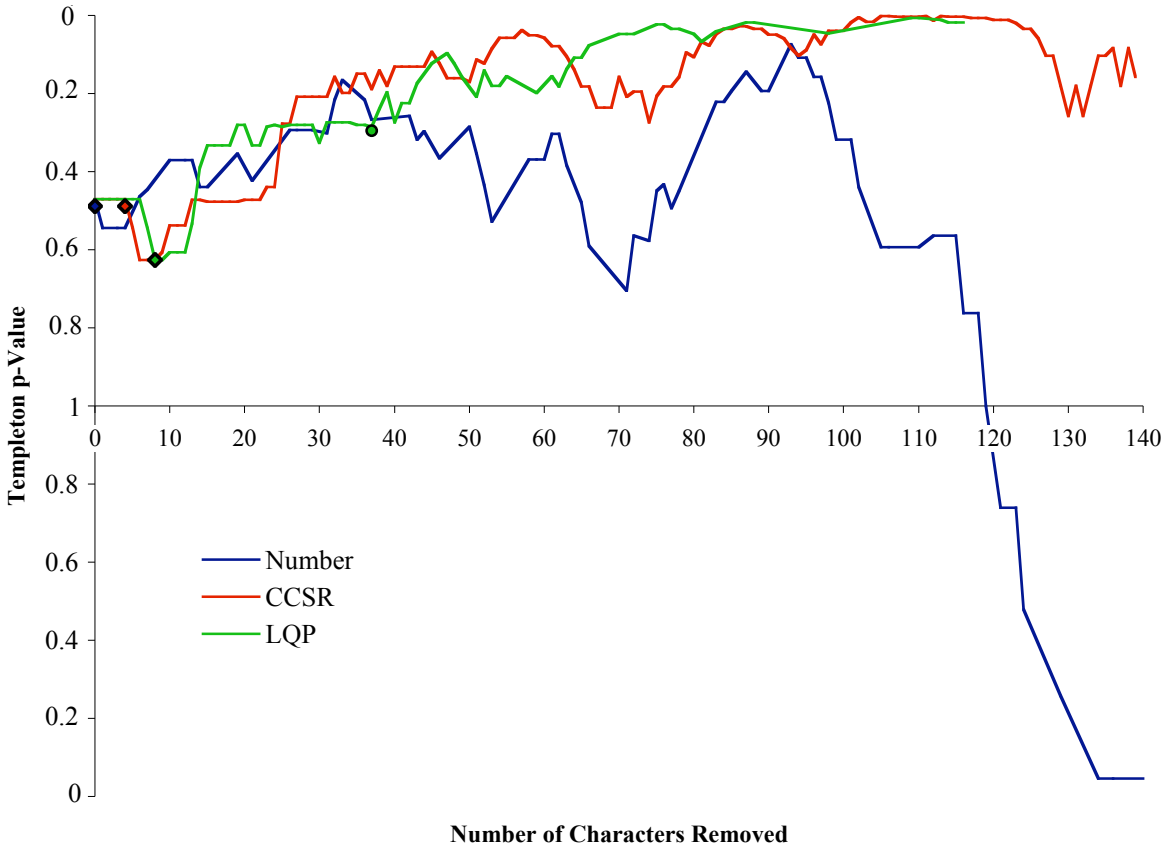


Figure 6.12. Line chart showing the Templeton test p-value when comparing the fit of the data of Rieppel and Reisz (1999) to trees with backbone constraints forcing turtles within the Anapsida and Diapsida at each stage of number, CCSR and LQP boildowns. The x-axis represents the number of characters removed by the boildown. The y-axis is the Templeton test p-value. Values above the x-axis indicate that the diapsid hypothesis is more parsimonious than the anapsid hypothesis, whereas those values below the x-axis indicate that the anapsid hypothesis is more parsimonious. Coloured diamonds indicate the point at which the boildown bootstrap identified maximum bootstrap values for each boildown. The green circle is the point at which characters removed in the LQP boildown have an LQP value lower than 0.05, and are therefore significantly more compatible than random characters.

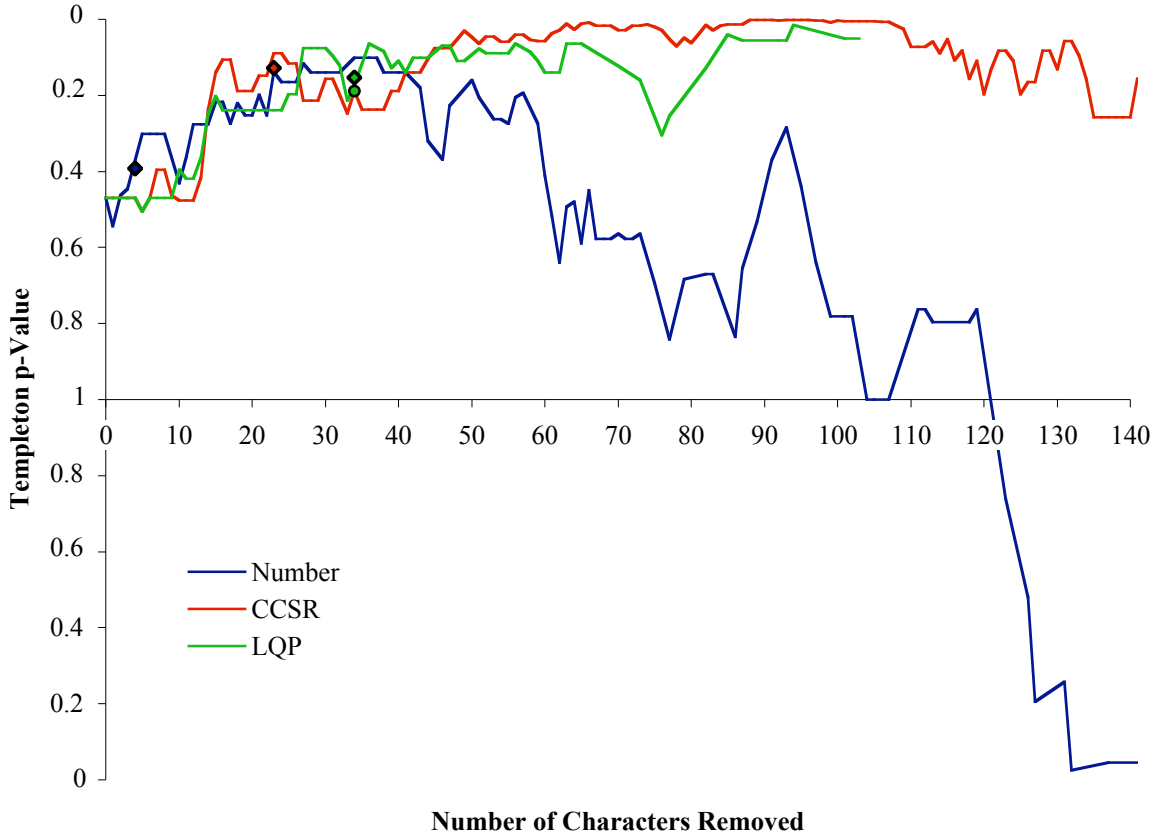


Figure 6.13. Line chart showing the Templeton test p-value when comparing the fit of the data of Rieppel and Reisz (1999) to trees with backbone constraints forcing turtles within the Anapsida and Diapsida at each stage of fuzzy number, CCSR and LQP boildowns. The x-axis represents the number of characters removed by the boildown. The y-axis is the Templeton test p-value. Values above the x-axis indicate that the diapsid hypothesis is more parsimonious than the anapsid hypothesis, whereas those values below the x-axis indicate that the anapsid hypothesis is more parsimonious. Coloured diamonds indicate the point at which the boildown bootstrap identified maximum bootstrap values for each boildown. The green circle is the point at which characters removed in the LQP boildown have an LQP value lower than 0.05, and are therefore significantly more compatible than random characters.

6.3.2 Lee (2001)

6.3.2.1 Reanalysis

Leaf	Maximum	Difference	Entropy
Araeoscelidida	0.9451	0.8993	0.8315
Younginiformes	0.9437	0.8974	0.8289
Claudiosauridae	0.9427	0.8959	0.8243
Protorothyrididae	0.9416	0.8942	0.8190
Procolophonoidea	0.9375	0.8899	0.8049
Sphenacodontidae	0.9366	0.8860	0.8165
Edaphosauridae	0.9365	0.8859	0.8160
Ophiacodontidae	0.9363	0.8855	0.8152
Lanthanosuchidae	0.9360	0.8879	0.8063
Acleistorhinidae	0.9356	0.8872	0.8050
Caseidae	0.9352	0.8838	0.8109
Pareiasauridae	0.9348	0.8854	0.7928
Nycteroleteridae	0.9315	0.8816	0.7960
Millerettidae	0.9282	0.8728	0.7806
Outgroup	0.9269	0.8701	0.7872
Eosauropterygia	0.9269	0.8760	0.7836
Placodontia	0.9268	0.8758	0.7833
Squamata	0.9256	0.8724	0.7803
Rhynchosauria	0.9249	0.8711	0.7771
Trilophosauria	0.9249	0.8724	0.7764
Rhychocephalia	0.9245	0.8712	0.7778
Average	0.9239	0.8670	0.7754
Archosauriformes	0.9214	0.8662	0.7680
Captorhinidae	0.9135	0.8470	0.7506
Prolacertiformes	0.9131	0.8522	0.7525
Kuehneosauridae	0.9110	0.8512	0.7484
Choristodera	0.9085	0.8470	0.7417
Testudines	**0.9036**	**0.8315**	**0.6986**
Gorgonopsia	0.8595	0.7533	0.6076
Cynodontia	0.8594	0.7531	0.6073

Table 6.5. Maximum, difference and entropy leaf stability measures for the bootstrap trees from the analysis of the osteological data of Lee (2001). Taxa are listed in order of decreasing stability. Turtles are highlighted in bold.

Reanalysis of the osteological portion of the data matrix of Lee (2001), an electronic version of which was provided by Mike Lee, produced the results he reported. A single MPT (shown on the right in Fig. 6.9) was recovered (L=771, CI=0.541, RI=0.666), in which turtles were nested within the Anapsida, as the sister-group of the pareiasaurs. The average bootstrap and decay values over the whole tree were 74.2 and 5.8 respectively. Templeton tests showed that the anapsid placement for turtles is a significantly better fit to

the data than the most parsimonious diapsid placement (p = 0.045), which positions turtles as the sister-group of rhynchosaurs.

Results of a bootstrap leaf stability analysis (Table 6.5) show that turtles, gorgonopsians and cynodontians are the least stable leaves, although all taxa have high values, and the difference between these least stable taxa and the remainder of the leaves is very small.

6.3.2.2 Sequential Sister-Group Removal

Removal of the Pareiasauridae, the initial sister-group of turtles, caused turtles to move to become the sister-group of the Rhynchosauria, within the Diapsida. Removal of the Rhynchosauria caused the strict consensus to collapse, producing ten MPTs with turtles as sister-group to procolophonoids, sauropterygians, squamates and lepidosaurs in at least one.

6.3.2.3 RSACW

Figure 6.14 shows the results of RSACW of Lee's (2001) data. Turtles remained within the Anapsida in all RSACW trees. Exclusion of the 36 characters that have a maximal CI on the MPT of all data led to the Gorgonopsia and Cynodontia moving into the Anapsida as the sister-taxon to the turtles (Fig. 6.14a). The second RSACW tree (Fig. 6.14b), with three further characters removed, differed from the strict consensus with all characters included only in the internal resolution of the Diapsida. No remaining characters had a CI of 1.0 on this tree.

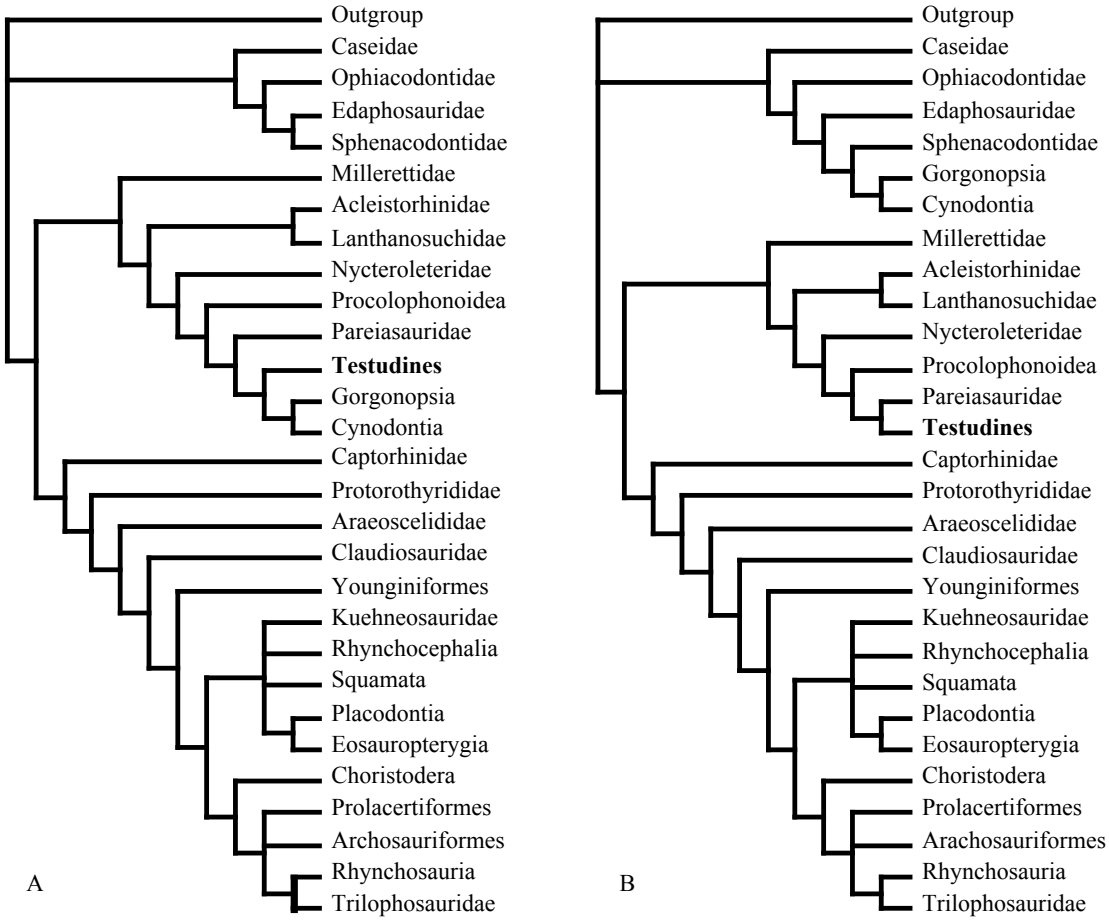


Figure 6.14. (A) First and (B) second RSACW trees of the analysis of the data of Lee (2001).

6.3.2.4 Compatibility Methods

6.3.2.4.1 Taxon Jacknife Analysis

The matrix of Lee (2001) contained 7943 pairwise incompatibilities, and 13 characters were uninformative (in the *Boildown* program). Results of first and second order taxon jacknife analyses are shown in table 6.6. As with the data of Rieppel and Reisz (1999), the taxon involved in most incompatibilities is the Cynodontia, followed by the Testudines, Pareiasauridae, Placodontia and Gorgonopsia. Along with these taxa, the randomisation test suggests that the Rhynchosauria, Lanthanosuchidae, Kuehneosauridae, Acleistorhinidae and Trilophosauridae (in order of decreasing p-value) are no less incompatible than would be expected by chance. In fact, both the Testudines and Cynodontia are involved in significantly more pairwise incompatibilities than would be expected by chance alone.

Taxon Excluded	Uninf.	% Missing	Caused	Expected	p	2nd Order
Cynodontia	16	2.8	763	435	0.991	1188
Testudines	**17**	**15.3**	**515**	**288**	**0.996**	**947**
Pareiasauridae	16	10.2	443	342	0.878	883
Placodontia	17	8	406	336	0.789	839
Gorgonopsia	15	3.4	332	395	0.278	786
Rhynchosauria	13	5.1	311	370	0.279	759
Trilophosauridae	13	15.3	183	303	0.061	637
Choristodera	14	13.6	182	301	0.052	626
Claudiosauridae	14	10.8	171	314	0.030	615
Rhynchocephalia	14	9.1	161	329	0.009	609
Procolophonoidea	14	11.4	152	312	0.008	604
Nycteroleteridae	14	12.5	126	279	0.008	573
Kuehneosauridae	14	27.3	125	195	0.111	573
Edaphosauridae	13	11.9	114	309	0.002	562
Squamata	15	17.6	113	290	0.002	563
Outgroup	13	4.5	111	383	0.001	558
Younginiformes	13	9.7	111	329	0.001	560
Archosauriformes	13	16.5	102	279	0.003	552
Araeoscelididae	13	4.5	98	353	0.001	547
Eosauropterygia	14	24.4	98	221	0.009	550
Prolacertiformes	13	15.9	95	291	0.001	546
Sphenacodontidae	13	1.7	90	385	0.001	544
Lanthanosuchidae	14	42.6	81	106	0.271	529
Millerettidae	13	14.8	65	262	0.001	515
Acleistorhinidae	13	50.6	40	82	0.070	490
Ophiacodontidae	13	3.4	37	368	0.001	490
Caseidae	13	1.1	27	388	0.001	480
Captorhinidae	13	2.8	23	362	0.001	474
Protorothyrididae	13	6.8	7	315	0.001	458

Table 6.6. Results of first and second order taxon jacknife analyses of the osteological data of Lee (2001). See caption of Table 6.3 for key. Taxa are listed in order of decreasing number of caused incompatibilities. Turtles are highlighted in bold.

6.3.2.4.2 Compatibility Analysis

Results of 'Number', CCSR and LQP compatibility and boildown analyses of the osteological data of Lee (2001) are shown in table 6.7. Average 'Number', CCSR and LQP compatibility values for the whole dataset were 90.26, 0.722 and 0.1072, respectively. In the initial analysis, character 14 exhibited the highest number of pairwise incompatibilities with 142. In both the CCSR and LQP tests, character 22 was the worst character initially, with values of 1.257 and 1, respectively.

Char.	No.	Fuzzy No.	CCSR	Fuzzy CCSR	LQP	Fuzzy LQP	No. Boil	Fuzzy No. Boil	CCSR Boil	Fuzzy CCSR Boil	LQP Boil	Fuzzy LQP Boil
1	88	150	0.648	0.591	**0.002**	**0.001**	127=	119=	138	-	-	-
2	U	-	-	-	-	-	-	-	-	-	-	-
3	67	69	0.81	0.8	0.209	0.213	119=	130=	129	102	42	35
4	57	66	0.576	0.531	**0.037**	**0.024**	-	-	-	-	109	108=
5	66	67	0.79	0.775	0.181	0.175	-	-	-	66	38	38
6	139	348	0.932	0.835	0.057	**0.026**	9	3	35	39	72	72
7	55	57	0.66	0.66	0.12	0.116	114	117	112	108	51	51
8	125	174	1.078	1.041	0.685	0.563	37=	58=	4	6	4	9
9	44	44	0.883	0.878	0.414	0.417	117=	119=	94	29	15	18
10	124	260	0.923	1.034	0.177	0.593	34=	19	43	8	40	11
11	106	197	0.755	0.637	**0.003**	**0.004**	60	45	51	64	73	83
12	U	-	-	-	-	-	-	-	-	-	-	-
13	75	100	0.589	0.472	**0.002**	**0.003**	-	-	131	115	108	104=
14	150	389	1.021	1.02	0.764	0.613	1	1	8	10	2	5
15	54	55	0.719	0.714	0.178	0.171	134=	128=	119	54	33	33
16	49	49	0.583	0.568	0.077	0.075	134=	135=	130	101	70	64
17	136	286	0.888	0.615	**0.003**	**0.001**	13=	16	50	87	-	-
18	73	86	0.608	0.459	**0.006**	**0.001**	101=	110=	107	113	112=	111=
19	131	290	0.95	0.956	0.163	0.335	23=	13	22	18	37	22
20	109	190	0.82	0.705	**0.01**	**0.017**	80	56=	77	69	89	74
21	U	-	-	-	-	-	-	-	-	-	-	-
22	145	218	1.273	1.33	0.996	0.995	2=	33	1	1	1	1
23	127	210	1.006	1.025	0.468	0.514	29=	36	9	9	11	14
24	85	106	0.728	0.631	**0.019**	**0.027**	105=	107=	81	76	62=	71
25	140	316	0.996	1.015	0.386	0.565	6=	8	12	12	13	6
26	100	123	0.957	0.935	0.341	0.381	82	87=	41	33	28	27
27	72	102	0.595	0.519	**0.003**	**0.005**	121=	119=	125	132	-	111=
28	17	17	0.324	0.327	**0.043**	0.051	134=	135=	-	139	50	53
29	61	78	0.728	0.739	**0.031**	**0.046**	134=	-	83	74	62=	52
30	107	193	0.749	0.582	**0.004**	**0.004**	85=	56=	114	103	110	104=
31	114	223	0.794	0.639	**0.003**	**0.002**	69=	82=	103	121	-	111=
32	136	341	0.922	0.89	**0.029**	0.067	11=	4	30	25	46	42
33	102	123	0.978	0.953	0.406	0.385	56=	79=	14	17	14	16
34	80	119	0.678	0.687	**0.018**	**0.048**	-	-	-	82	88	70
35	122	212	0.895	0.853	0.099	0.125	43=	42=	54	70	76	67=
36	107	209	0.741	0.566	**0.001**	**0.003**	-	-	135	140	-	-
37	121	171	0.964	0.829	0.309	0.117	40=	73	20	34	17	41
38	106	202	0.708	0.458	**0.001**	**0.001**	85=	71=	-	-	-	-
39	103	196	0.694	0.479	**0.001**	**0.001**	121=	-	-	-	-	-
40	15	15	0.326	0.316	**0.026**	**0.029**	-	-	-	137	64	66
41	110	222	0.782	0.63	**0.002**	**0.001**	88=	53	95	84	111	98=
42	144	328	0.941	0.713	**0.016**	**0.001**	4=	5	29	63	85	96=
43	27	27	0.52	0.516	0.144	0.151	134=	135=	123	126	52	49
44	54	60	0.519	0.462	**0.008**	**0.013**	-	-	-	-	71	77
45	97	150	0.714	0.498	**0.002**	**0.001**	121=	133=	-	-	-	-
46	93	137	0.666	0.39	**0.001**	**0.001**	98=	106	98	130	-	-
47	92	94	1.13	1.117	0.679	0.669	85=	95	3	3	6	4
48	113	214	0.756	0.495	**0.001**	**0.001**	93=	101=	101	128	-	-
49	130	256	0.873	0.615	**0.001**	**0.001**	25=	31	58	91	98=	107
50	108	221	0.726	0.514	**0.001**	**0.001**	134=	123=	122	135	-	-
51	127	262	0.907	0.898	0.093	0.176	29=	20	46	30	56	43
52	118	239	0.769	0.476	**0.001**	**0.001**	54=	66	116	136	-	-
53	109	240	0.739	0.556	**0.001**	**0.001**	69=	47	89	96	112=	-
54	C	-	-	-	-	-	-	-	-	-	-	-
55	108	184	0.728	0.484	**0.001**	**0.001**	88=	84=	-	-	-	-
56	112	195	0.884	0.942	0.124	0.31	64	40	70	28	61	29
57	86	151	0.641	0.586	**0.003**	**0.003**	-	-	-	-	105=	110
58	81	133	0.633	0.606	**0.006**	**0.009**	-	-	-	-	93=	89=
59	91	154	0.677	0.593	**0.004**	**0.005**	131=	130=	-	-	-	103
60	63	76	0.542	0.446	**0.007**	**0.002**	134=	135=	-	-	-	-
61	125	250	0.934	0.995	0.185	0.42	37=	27	34	15	30	19

62	127	260	0.891	0.77	**0.037**	**0.021**	29=	21	53	47	81	60
63	97	159	0.699	0.474	**0.001**	**0.001**	75=	74	86	105	-	-
64	132	252	0.997	1.028	0.443	0.567	20=	29	13	11	12	8
65	130	273	0.885	0.707	**0.01**	**0.01**	25=	22	59	83	-	98=
66	128	286	0.87	0.719	**0.002**	**0.007**	28	14	39	55	43	79
67	122	196	0.945	0.912	0.181	0.191	37=	52	26	24	36	32
68	70	86	0.556	0.43	**0.001**	**0.002**	105=	99	104	107	107	106
69	U	-	-	-	-	-	-	-	-	-	-	-
70	94	145	0.685	0.489	**0.001**	**0.001**	98=	84=	110	109	-	-
71	23	23	0.449	0.448	0.2	0.206	-	-	-	110	32	30
72	118	165	1.013	0.989	0.578	0.507	46=	58=	11	16	9	15
73	42	44	0.499	0.437	**0.026**	**0.018**	-	-	-	-	97	85
74	131	272	0.869	0.649	**0.004**	**0.001**	20=	23=	55	81	112=	-
75	109	230	0.736	0.555	**0.001**	**0.002**	56=	41	68	86	-	111=
76	140	362	0.952	0.904	0.111	0.134	6=	2	25	23	68=	36
77	131	256	0.994	1.049	0.435	0.608	23=	26	19	7	18	13
78	138	289	0.933	0.752	**0.047**	**0.013**	10	15	27	68	74=	84
79	144	320	0.982	0.832	0.295	**0.035**	4=	9	17	43	26	93
80	120	212	0.901	0.864	0.093	0.094	50=	42=	71	53	83	82
81	U	-	-	-	-	-	-	-	-	-	-	-
82	135	262	0.954	0.808	0.197	0.05	15=	30	21	41	31	63
83	132	268	0.906	0.74	**0.025**	**0.007**	20=	23=	45	71	112=	86
84	30	43	0.278	0.311	**0.001**	**0.001**	127=	126=	126	133	112=	108=
85	108	175	0.849	0.826	**0.044**	0.06	75=	68=	75	60	79	65
86	127	233	0.885	0.647	**0.011**	**0.003**	29=	39	42	77	68=	-
87	110	162	0.921	0.939	0.271	0.348	62=	62=	48	22	39	24
88	136	240	0.951	0.702	0.119	**0.005**	11=	32	24	78	44	96=
89	89	135	0.673	0.56	**0.005**	**0.002**	101=	107=	121	119	-	-
90	81	133	0.633	0.606	**0.006**	**0.009**	-	-	-	-	93=	89=
91	114	212	0.808	0.618	**0.001**	**0.002**	62=	67	76	90	98=	111=
92	77	121	0.582	0.483	**0.001**	**0.003**	88=	87=	90	85	105=	98=
93	C	-	-	-	-	-	-	-	-	-	-	-
94	32	36	0.401	0.38	**0.006**	**0.01**	-	-	-	141	-	-
95	112	203	0.783	0.533	**0.001**	**0.001**	69=	82=	93	125	-	-
96	76	112	0.712	0.808	0.084	0.283	-	-	136	46	57	34
97	31	31	0.581	0.577	0.192	0.185	121=	122	82	52	22	28
98	U	-	-	-	-	-	-	-	-	-	-	-
99	112	192	0.799	0.588	**0.003**	**0.001**	73=	62=	87	100	-	-
100	U	-	-	-	-	-	-	-	-	-	-	-
101	79	82	0.946	0.966	0.358	0.328	91=	96	49	21	21	23
102	89	125	0.718	0.562	**0.01**	**0.007**	134=	-	137	145	98=	-
103	58	86	0.447	0.383	**0.002**	**0.001**	105=	97=	105	106	-	-
104	91	103	0.916	0.838	0.198	0.134	91=	91=	63	62	55	54
105	91	151	0.689	0.627	**0.003**	**0.003**	108=	107=	115	124	-	-
106	60	72	0.499	0.392	**0.001**	**0.002**	121=	123=	120	118	112=	-
107	115	209	0.838	0.758	**0.022**	**0.011**	54=	50=	65	73	95=	92
108	60	60	0.754	0.726	0.114	0.112	110	112=	99	94	54	45
109	134	301	0.934	0.897	0.12	0.152	15=	10	37	32	66	47
110	97	170	0.692	0.507	**0.001**	**0.001**	79	70	80	93	-	-
111	87	154	0.593	0.392	**0.001**	**0.001**	98=	89	109	120	-	-
112	29	29	0.545	0.542	0.167	0.199	133	126=	91	59	27	26
113	95	140	0.67	0.407	**0.001**	**0.001**	97	105	100	129	-	-
114	112	193	0.791	0.559	**0.004**	**0.001**	93=	102=	97	123	-	-
115	101	176	0.689	0.449	**0.001**	**0.001**	65=	62=	72	98	-	-
116	119	217	0.874	0.788	**0.048**	**0.041**	43=	37=	56	51	80	78
117	135	267	0.956	0.844	0.162	0.084	15=	23=	36	35	53	56
118	66	66	1.179	1.18	0.687	0.699	101=	110=	2	2	5	3
119	70	116	0.52	0.362	**0.001**	**0.001**	-	-	-	-	-	111=
120	110	230	0.758	0.581	**0.001**	**0.001**	115=	118	127	134	-	-
121	14	14	0.307	0.305	**0.023**	**0.029**	-	-	-	-	91=	75=
122	44	44	0.831	0.83	0.315	0.342	127=	128=	113	36	20	20
123	111	140	0.955	0.849	0.305	0.147	56=	75=	28	44	25	48
124	117	278	0.809	0.797	**0.002**	**0.009**	43=	18	78	37	98=	67=
125	107	177	0.767	0.548	**0.001**	**0.001**	131=	133=	132	143	-	-

126	145	324	0.967	0.773	0.101	**0.017**	2=	6	18	50	49	75=
127	127	197	0.992	0.901	0.342	0.186	34=	46	16	27	19	31
128	99	178	0.705	0.572	**0.001**	**0.002**	111=	114=	134	-	-	-
129	53	55	0.659	0.669	0.068	0.083	111=	114=	108	97	48	46
130	130	244	0.902	0.651	**0.012**	**0.001**	25=	37=	38	80	60	111=
131	117	214	0.81	0.582	**0.001**	**0.001**	49	50=	69	95	-	-
132	106	179	0.771	0.607	**0.004**	**0.001**	83	77	79	89	-	111=
133	116	216	0.837	0.759	**0.019**	**0.018**	53	48=	67	72	112=	94=
134	102	163	0.816	0.777	**0.04**	0.064	119=	93=	92	61	77	57
135	U	-	-	-	-	-	-	-	-	-	-	-
136	51	57	0.493	0.441	**0.013**	**0.003**	-	-	-	-	98=	-
137	120	201	0.815	0.5	**0.002**	**0.001**	50=	61	85	112	-	-
138	110	163	0.918	0.909	0.212	0.266	67	75=	40	26	29	25
139	112	210	0.756	0.505	**0.001**	**0.001**	93=	103=	102	127	-	-
140	117	198	0.932	0.921	0.184	0.277	46=	42=	32	20	24	21
141	26	26	0.422	0.423	0.133	0.117	-	-	-	117	41	40
142	107	203	0.734	0.527	**0.001**	**0.001**	93=	78	106	114	-	-
143	125	190	0.985	0.907	0.306	0.133	33	60	15	31	23	44
144	57	100	0.405	0.316	**0.001**	**0.001**	111=	104	118	122	-	-
145	118	212	0.854	0.684	**0.016**	**0.005**	65=	97=	96	111	98=	98=
146	113	201	0.849	0.801	**0.043**	0.065	56=	48=	66	67	86=	61
147	123	219	0.918	0.853	0.092	0.081	40=	35	47	56	65	59
148	U	-	-	-	-	-	-	-	-	-	-	-
149	135	301	0.91	0.79	**0.026**	**0.023**	15=	12	44	57	91=	73
150	91	106	0.877	0.809	0.183	0.148	84	90	52	75	34	39
151	109	184	0.846	0.835	0.061	0.103	75=	55	84	58	74=	50
152	39	39	0.52	0.51	**0.027**	**0.022**	-	-	133	131	86=	81
153	83	158	0.58	0.462	**0.001**	**0.001**	81	68=	88	92	-	-
154	105	188	0.745	0.593	**0.001**	**0.001**	101=	93=	-	142	-	-
155	119	198	0.864	0.688	**0.013**	**0.008**	46=	65	62	99	82	98=
156	12	12	0.237	0.237	**0.018**	**0.017**	-	-	-	138	98=	67=
157	99	133	0.835	0.762	0.074	0.056	108=	100	64	45	47	55
158	122	234	0.823	0.612	**0.004**	**0.002**	40=	34	73	88	112=	-
159	90	94	0.996	1.009	0.562	0.573	69=	91=	10	13	10	7
160	95	171	0.672	0.505	**0.001**	**0.001**	121=	123=	124	144	-	-
161	101	133	0.921	0.922	0.324	0.429	68	81	31	19	16	17
162	140	324	0.953	0.803	0.089	**0.02**	6=	7	23	48	45	80
163	136	305	0.932	0.782	**0.038**	**0.008**	13=	11	33	49	59	87=
164	96	114	0.897	0.824	0.245	0.202	75=	86	57	40	35	37
165	55	55	1.065	1.067	0.592	0.565	127=	130=	6	4	8	10
166	C	-	-	-	-	-	-	-	-	-	-	-
167	U	-	-	-	-	-	-	-	-	-	-	-
168	111	130	1.089	1.052	0.71	0.514	61	79=	5	5	3	12
169	78	100	0.677	0.573	**0.003**	**0.003**	117=	112=	117	116	95=	-
170	86	151	0.627	0.538	**0.001**	**0.003**	-	-	-	-	-	111=
171	126	271	0.892	0.856	**0.036**	**0.039**	34=	17	60	38	67	58
172	133	239	1.024	1.02	0.632	0.819	19	28	7	14	7	2
173	117	200	0.896	0.856	0.056	**0.035**	50=	54	61	42	58	62
174	108	168	0.85	0.767	**0.026**	**0.007**	73=	71=	74	65	78	87=
175	90	136	0.734	0.686	**0.009**	**0.009**	115=	114=	111	79	90	89=
176	51	58	0.519	0.491	**0.012**	**0.006**	-	-	128	104	84	94=

Table 6.7. Table showing the results of compatibility tests and the order of removal in subsequent boildowns of each character in the osteological data of Lee (2001) using CCSR, "Number" and LQP methods with and without the fuzzy compatibility. C = constant character and U = uninformative character. LQP results that are significant at the 5% level are highlighted in bold.

Removal of the 49 characters with initial LQP values greater than 0.1 resulted in two MPTs (L=385, CI=0.395, RI=0.726) under parsimony analysis. The strict consensus of these trees (Fig. 6.15a) still placed turtles within the Anapsida, as the sister-group of the pareiasaurs. The main differences between the strict consensus of these trees and the single MPT of all data concerned the internal resolution of the Diapsida. Applying a backbone constraint forcing turtles within the Diapsida resulted in four MPTs (L=390, CI=0.390, RI=0.720). Templeton tests showed that the difference in fit to the data between the shortest anapsid and diapsid turtle placements was not statistically significant (p-values ranged from 0.4458 to 0.5531).

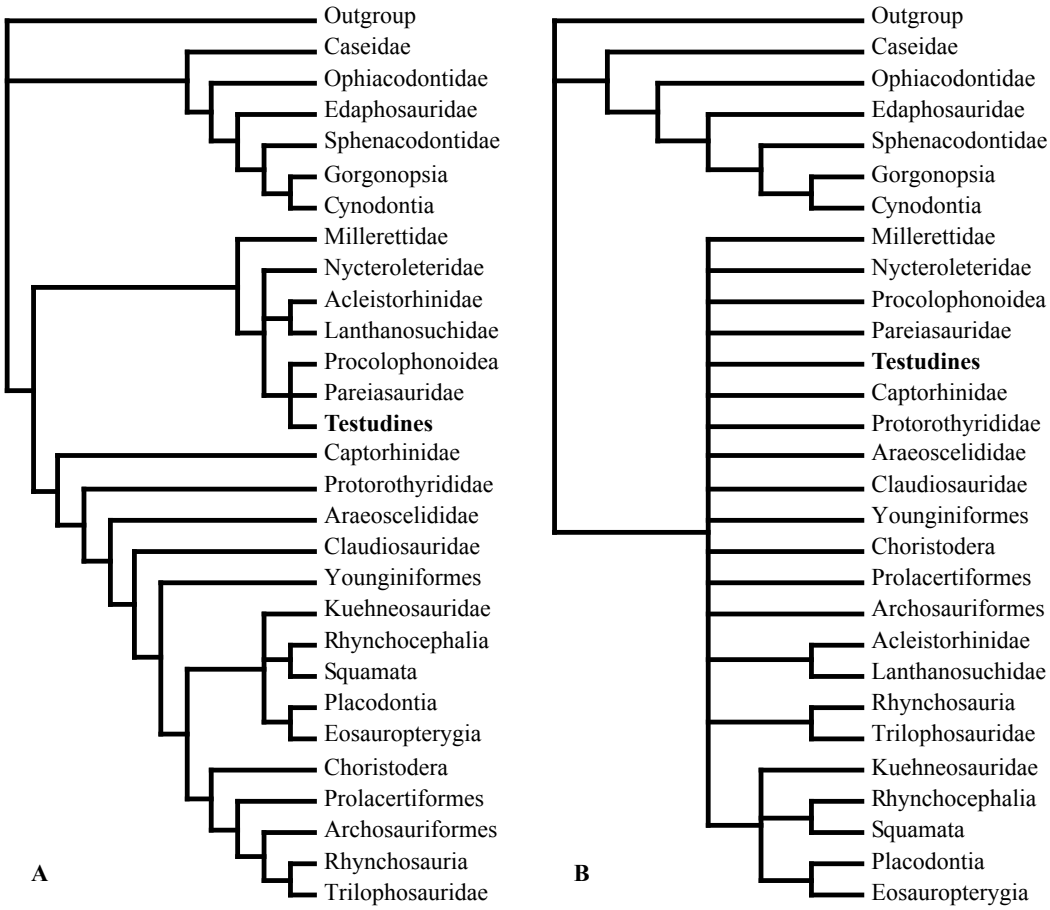


Figure 6.15. Strict consensus trees of analyses of the data of Lee (2001) with characters excluded that are not significantly better than random at the (A) 10% and (B) 5% level of LQP tests.

Removal of all 61 characters with LQP values greater than 0.05 yielded 12 MPTs (L=337, CI=0.412, RI=0.746) under parsimony analysis. The strict consensus of these trees (Fig. 6.15b) again showed turtles as the sister-group of the anapsid pareiasaurs. Much of the Diapsida was unresolved. Constraining turtles to be within the Diapsida produced nine MPTs (L=339, CI=0.410, RI=0.743). Templeton tests showed no significant difference in

length between the shortest trees with turtles within the anapsids and diapsids (p-values ranged from 0.7518 to 0.8112).

6.3.2.4.3 Boildown Analysis

With all three compatibility measures, Templeton tests at each stage of the boildowns (Fig. 6.16) showed a dramatic trend from initial significant support for the anapsid hypothesis towards support for the diapsid hypothesis. Both the CCSR and LQP analyses reached a Templeton p-value of zero when 56 characters had been removed, while the 'Number' analysis did so at 79 characters. In the CCSR boildown, support for the diapsid hypothesis reached significance at the 5% level when 103 characters had been removed.

6.3.2.4.4 Fuzzy Compatibility

Results of 'Number', CCSR and LQP fuzzy compatibility tests and subsequent boildowns of the osteological data of Lee (2001) are shown in table 6.7. The results of Templeton tests at each stage of the boildowns are shown in figure 6.17. Generally, as with the analysis of the data of Rieppel and Reisz (1999), the results using fuzzy compatibility are very similar to those using the normal compatibility method. However, the trend for the 'Number', CCSR and LQP analyses to move towards supporting the diapsid hypothesis occurred more quickly using the fuzzy method, reaching p-values of 0 when 88, 41 and 43 characters were removed, respectively. The 'Number' and CCSR analyses reached significant support for the diapsid hypothesis at the 5% level when 125 and 88 characters had been removed, respectively. Although the LQP boildown led to the diapsid hypothesis becoming more parsimonious, Templeton tests never showed a significant level of support for this placement over the shortest anapsid tree.

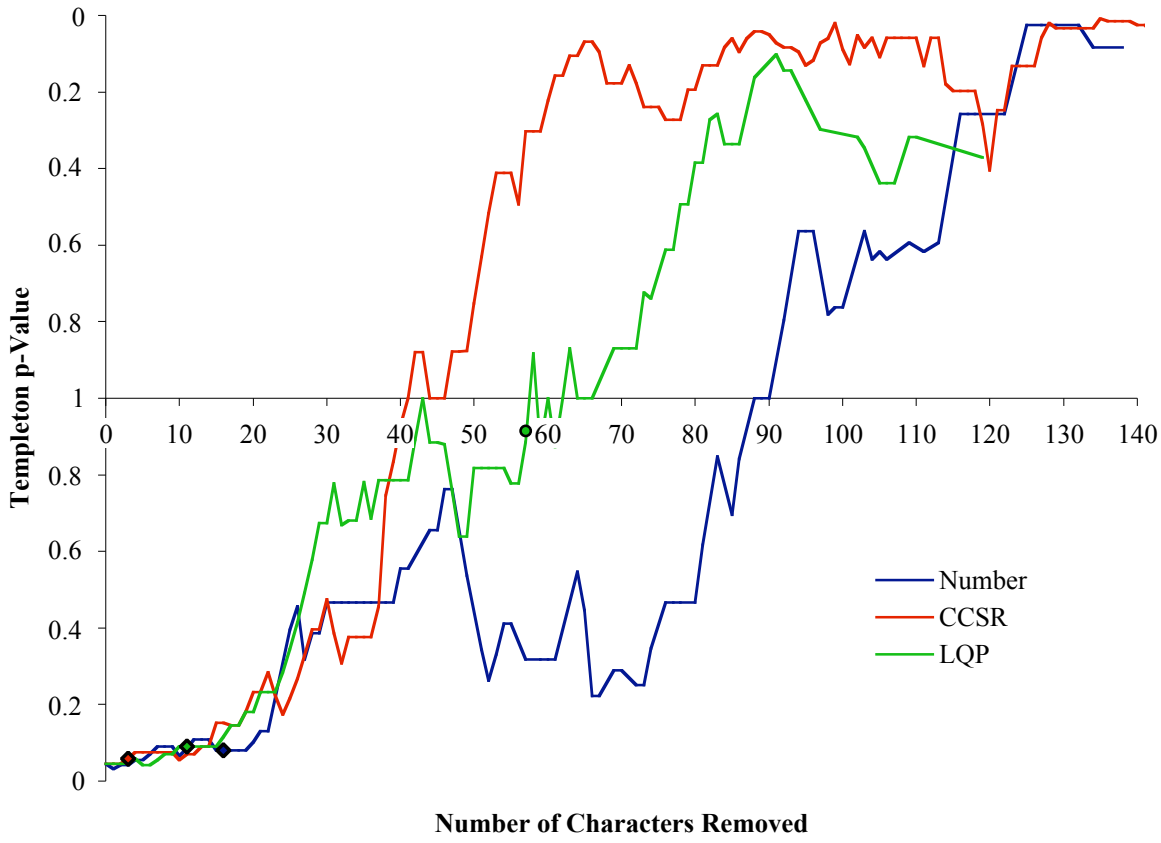


Figure 6.16. Line chart showing the Templeton test p-value when comparing the fit of the osteological data of Lee (2001) to trees with backbone constraints forcing turtles within the Anapsida and Diapsida at each stage of number, CCSR and LQP boildowns. The x-axis represents the number of characters removed by the boildown. The y-axis is the Templeton test p-value. Values above the x-axis indicate that the diapsid hypothesis is more parsimonious than the anapsid hypothesis, whereas those values below the x-axis indicate that the anapsid hypothesis is more parsimonious. Coloured diamonds indicate the point at which the boildown bootstrap identified maximum bootstrap values for each boildown. The green circle is the point at which characters removed in the LQP boildown have an LQP value lower than 0.05, and are therefore significantly more compatible than random characters.

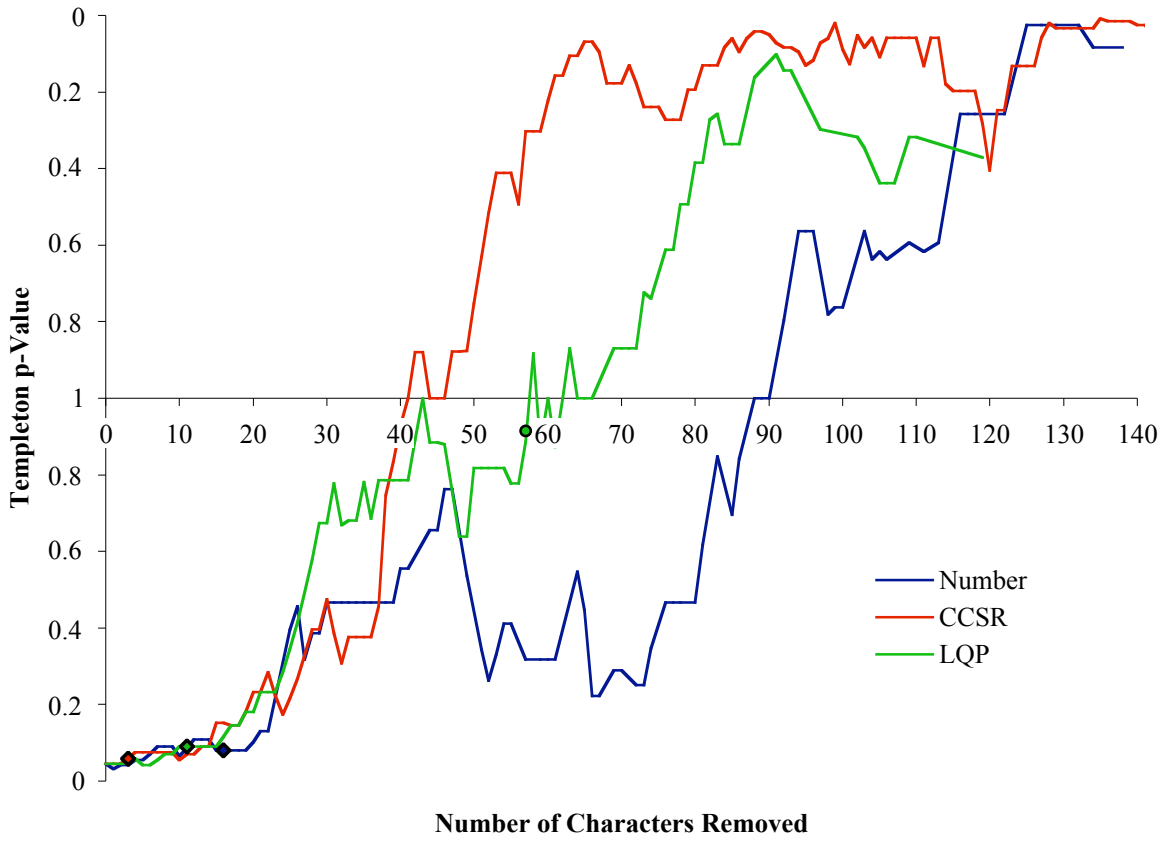


Figure 6.17. Line chart showing the Templeton test p-value when comparing the fit of the osteological data of Lee (2001) to trees with backbone constraints forcing turtles within the Anapsida and Diapsida at each stage of fuzzy number, CCSR and LQP boildowns. The x-axis represents the number of characters removed by the boildown. The y-axis is the Templeton test p-value. Values above the x-axis indicate that the diapsid hypothesis is more parsimonious than the anapsid hypothesis, whereas those values below the x-axis indicate that the anapsid hypothesis is more parsimonious. Coloured diamonds indicate the point at which the boildown bootstrap identified maximum bootstrap values for each boildown. The green circle is the point at which characters removed in the LQP boildown have an LQP value lower than 0.05, and are therefore significantly more compatible than random characters.

6.3.3 Consensus Matrix

Analysis of the strict consensus matrix produced three MPTs (TL = 749, CI = 0.538, RI = 0.654), of which two were the two MPTs found by analysis of Rieppel and Reisz (1999), and the third was the single MPT found by analysis of Lee (2001). Therefore, turtles plotted within the diapsids in two of the MPTs and within the anapsids in the third, so that a strict consensus of the three trees was highly unresolved (Fig. 6.18).

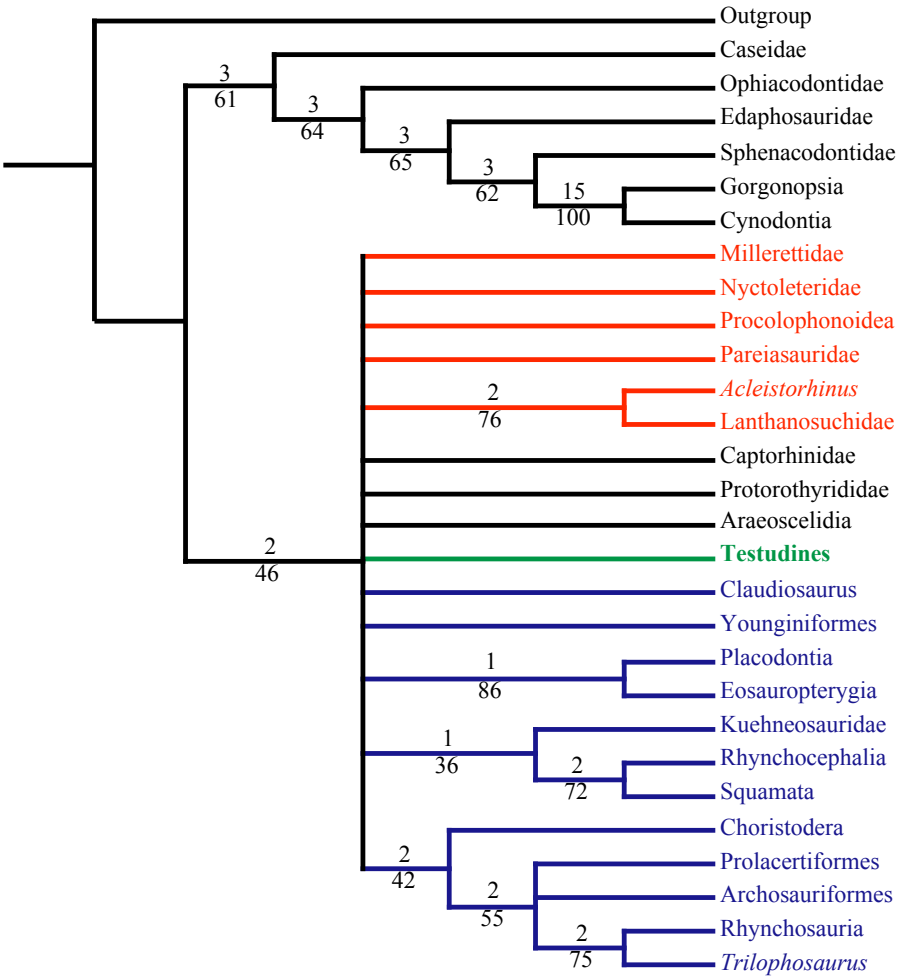


Figure 6.18. Strict consensus of the three MPTs from analysis of the consensus matrix. Numbers above nodes represent decay values, and those below nodes represent bootstrap proportions. The Anapsida are highlighted in red, the Diapsida in blue and turtles in green.

Results of the bootstrap leaf stability (Table 6.8) show that turtles were the least stable taxon, followed by the Cynodontia and Gorgonopsia. The average leaf stability of all of the taxa is lower than found with the analyses of Rieppel and Reisz (1999) and Lee (2001). However, given the variable position of the Testudines in the MPTs of the analysis of the consensus matrix, this instability is not surprising.

Leaf	Maximum	Difference	Entropy
Younginiformes	0.8899	0.7975	0.7289
Claudiosaurus	0.8890	0.7963	0.7255
Araeoscelidia	0.8862	0.7873	0.7215
Sphenacodontidae	0.8836	0.7875	0.7228
Edaphosauridae	0.8835	0.7874	0.7225
Ophiacodontidae	0.8832	0.7867	0.7211
Protorothyrididae	0.8819	0.7805	0.7085
Caseidae	0.8803	0.7821	0.7115
Acleistorhinus	0.8713	0.7695	0.6834
Lanthanosuchidae	0.8708	0.7690	0.6827
Placodontia	0.8695	0.7724	0.6827
Eosauropterygia	0.8695	0.7726	0.6830
Squamata	0.8690	0.7699	0.6812
Trilophosaurus	0.8684	0.7690	0.6770
Rhynchosauria	0.8676	0.7658	0.6782
Rhynchocephalia	0.8674	0.7684	0.6778
Archosauriformes	0.8654	0.7637	0.6683
Nyctoleterida	0.8652	0.7606	0.6689
Procolophonoidea	0.8605	0.7493	0.6565
Prolacertiformes	0.8596	0.7543	0.6599
Pareiasauridae	0.8570	0.7426	0.6508
Average	0.8556	0.7417	0.6576
Outgroup	0.8542	0.7356	0.6600
Captorhinidae	0.8507	0.7304	0.6437
Millerettidae	0.8498	0.7327	0.6358
Kuehneosauridae	0.8496	0.7402	0.6444
Choristodera	0.8452	0.7335	0.6355
Gorgonopsia	0.7827	0.6172	0.4916
Cynodontia	0.7824	0.6166	0.4910
Testudines	**0.6591**	**0.3705**	**0.3558**

Table 6.8. Maximum, difference and entropy leaf stability measures for the bootstrap trees from the analysis of the strict consensus matrix. Taxa are listed in order of decreasing stability. Turtles are highlighted in bold.

6.3.3.1 Sequential Sister-Group Removal

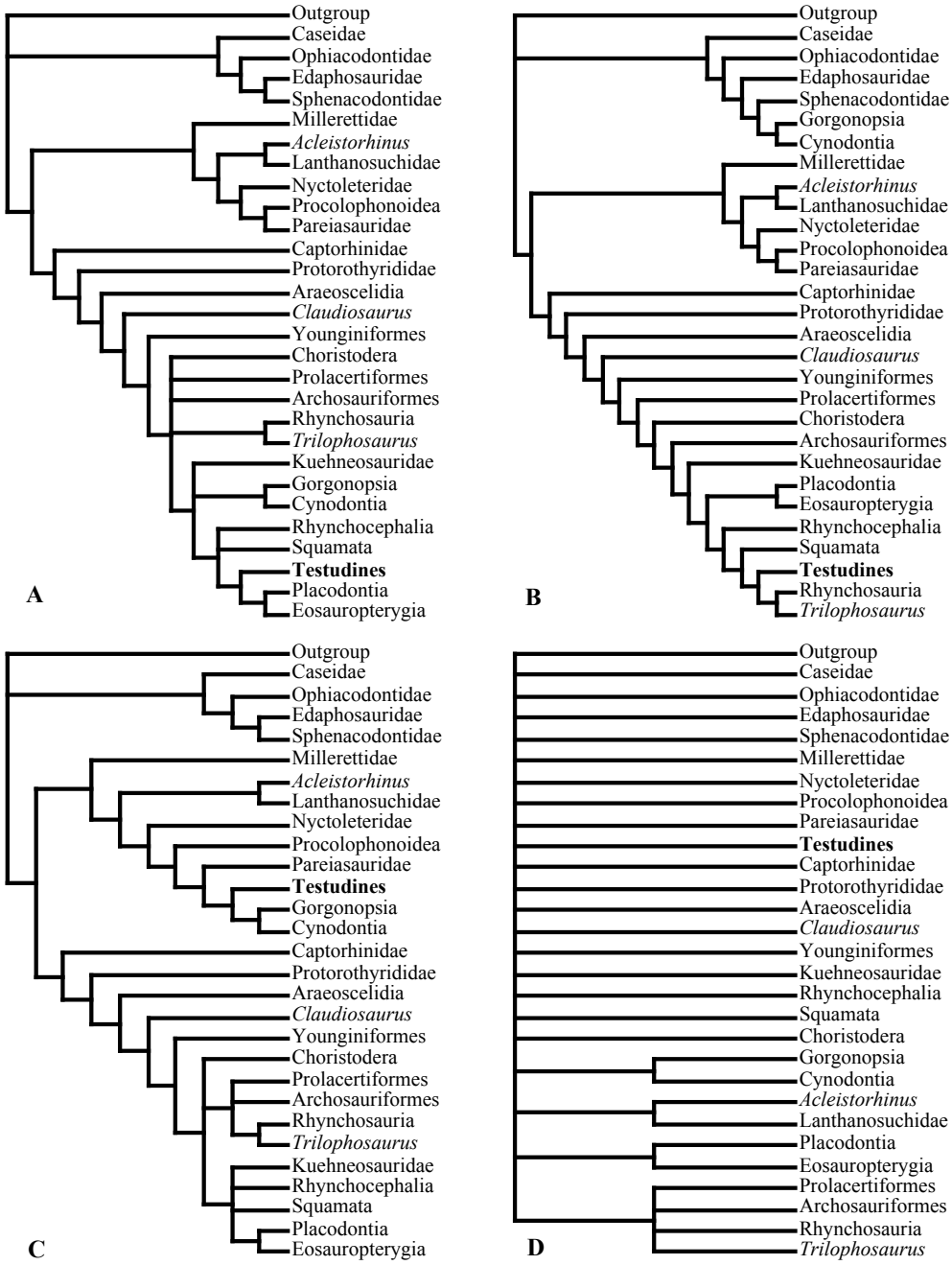
In the case of the consensus matrix, the analysis of all data places turtles with two separate sister-taxa, so both were removed individually to see how this affected the position of the turtles. Removal of the pareiasaurs, the sister-group of turtles in the MPT in which turtles were anapsids, led to two MPTs being found, in both of which turtles were the sister-group of the sauropterygians. Removal of the Eosauropterygia and Placodontia, which form the sister-group to turtles in the two MPTs where turtles were diapsids, caused turtles to move to become the sister-group of the pareiasaurs in the Anapsida. Removal of

both initial sister-groups led to the production of a single MPT with turtles within the Diapsida as the sister-group of the group comprising the Kuehneosauridae, Rhynchocephalia and Squamata. Removal of this group made turtles sister-group of the Rhynchosauria, and removal of these collapsed the strict consensus, producing two trees with turtles as sister-group to the anapsid procolophonids in one and the diapsid *Trilophosaurus* in the other.

6.3.3.2 RSACW

Figure 6.19 shows the results of RSACW of the consensus matrix. Two RSACW analyses were performed, one for the MPT with turtles in the Anapsida, and one for the consensus of the two MPTs with turtles within the Diapsida. Removing the 21 characters that fit the anapsid MPT perfectly resulted in 10 MPTs, placing turtles as the sister-group of the sauropterygians (Fig. 6.19a). The synapsid gorgonopsians and cynodontids also moved into the Diapsida. The next reweighting removed six characters, and led to the gorgonopsians and cynodontids moving back into the Synapsida (Fig. 6.19b). Turtles remained in the Diapsida, but became sister-group of the Rhynchosauria and *Trilophosaurus* (Fig. 6.19c). After the next reweighting, the turtles became the sister-group of the gorgonopsians and cynodonts within the Anapsida (Fig. 6.19d). Further reweighting also found this topology.

The first RSACW of the diapsid-turtle MPTs removed 33 characters, and led to 12 MPTs. The strict consensus of these (Fig. 6.19e) was very poorly resolved, because turtles plotted within the Anapsida in some of the MPTs and the Diapsida in others. Further reweightings did not change the topology further.


Outgroup
Caseidae
Ophiacodontidae
Edaphosauridae
Sphenacodontidae
Millerettidae
Acleistorhinus
Lanthanosuchidae
Nyctoleteridae
Procolophonoidea
Pareiasauridae
Captorhinidae
Protorothyrididae
Araeoscelidia
Claudiosaurus
Younginiformes
Choristodera
Prolacertiformes
Archosauriformes
Rhynchosauria
Trilophosaurus
Kuehneosauridae
Gorgonopsia
Cynodontia
Rhynchocephalia
Squamata
Testudines
Placodontia
Eosauropterygia
A
Outgroup
Caseidae
Ophiacodontidae
Edaphosauridae
Sphenacodontidae
Gorgonopsia
Cynodontia
Millerettidae
Acleistorhinus
Lanthanosuchidae
Nyctoleteridae
Procolophonoidea
Pareiasauridae
Captorhinidae
Protorothyrididae
Araeoscelidia
Claudiosaurus
Younginiformes
Prolacertiformes
Choristodera
Archosauriformes
Kuehneosauridae
Placodontia
Eosauropterygia
Rhynchocephalia
Squamata
Testudines
Rhynchosauria
Trilophosaurus
B
Outgroup
Caseidae
Ophiacodontidae
Edaphosauridae
Sphenacodontidae
Millerettidae
Acleistorhinus
Lanthanosuchidae
Nyctoleteridae
Procolophonoidea
Pareiasauridae
Testudines
Gorgonopsia
Cynodontia
Captorhinidae
Protorothyrididae
Araeoscelidia
Claudiosaurus
Younginiformes
Choristodera
Prolacertiformes
Archosauriformes
Rhynchosauria
Trilophosaurus
Kuehneosauridae
Rhynchocephalia
Squamata
Placodontia
Eosauropterygia
C
Outgroup
Caseidae
Ophiacodontidae
Edaphosauridae
Sphenacodontidae
Millerettidae
Nyctoleteridae
Procolophonoidea
Pareiasauridae
Testudines
Captorhinidae
Protorothyrididae
Araeoscelidia
Claudiosaurus
Younginiformes
Kuehneosauridae
Rhynchocephalia
Squamata
Choristodera
Gorgonopsia
Cynodontia
Acleistorhinus
Lanthanosuchidae
Placodontia
Eosauropterygia
Prolacertiformes
Archosauriformes
Rhynchosauria
Trilophosaurus
D

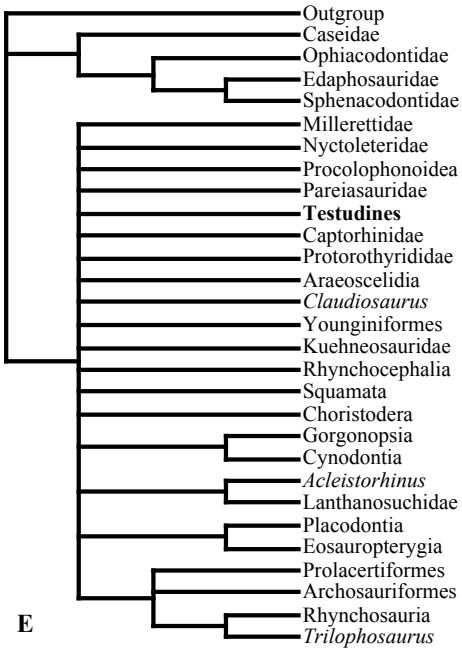


Figure 6.19. RSACW trees of the analysis of the consensus matrix. Trees A to D are the four stages of the RSACW of the anapsid MPT. Tree E is the only tree produced by RSACW of the diapsid MPTs.

6.3.3.3 Compatibility Analysis

6.3.3.3.1 Taxon Jacknife Analysis

The consensus matrix contained 6992 pairwise incompatibilities, and 13 characters were uninformative (in the *Boildown* program). Results of first- and second-order taxon jacknife analyses are shown in table 6.9. The taxa that were the cause of most incompatibilities were the Cynodontia, followed by the Testudines, Placodontia, Gorgonopsia and Rhynchosauria. Along with these taxa, the randomisation test showed that the Lanthanosuchidae, Kuehneosauridae, *Acleistorhinus* and *Trilophosaurus* were no less incompatible than would be expected by chance alone. Again, Cynodontia and Testudines were involved in significantly more incompatibilities than would be expected by chance alone.

Taxon Excluded	Uninf.	% Missing	1st Order	Exp	P	2nd Order
Cynodontia	16	4.2	733	410	0.996	1115
Testudines	**17**	**22.0**	**428**	**241**	**0.989**	**816**
Pareiasauridae	16	14.3	394	291	0.918	789
Placodontia	17	11.3	357	297	0.793	749
Gorgonopsia	15	4.2	351	377	0.418	760
Rhynchosauria	13	5.4	313	352	0.339	716
Trilophosaurus	13	15.5	180	297	0.058	592
Procolophonoidea	14	11.9	178	300	0.040	586
Rhynchocephalia	14	10.1	171	313	0.030	575
Claudiosaurus	14	11.3	169	297	0.039	571
Choristodera	14	14.3	169	283	0.042	570
Squamata	15	17.9	155	283	0.021	559
Kuehneosauridae	14	28.0	125	184	0.151	530
Edaphosauridae	13	13.7	107	289	0.004	513
Outgroup	13	4.8	103	368	0.001	508
Araeoscelidia	13	5.4	100	345	0.001	506
Eosauropterygia	14	28.0	99	192	0.036	507
Younginiformes	13	10.7	96	308	0.001	503
Nyctoleteridae	14	14.3	90	261	0.002	495
Prolacertiformes	13	16.7	88	272	0.001	496
Sphenacodontidae	13	3.0	77	362	0.001	488
Archosauriformes	13	16.7	69	255	0.001	477
Lanthanosuchidae	14	48.2	57	82	0.206	462
Millerettidae	13	16.7	45	245	0.001	453
Acleistorhinus	13	51.8	37	76	0.076	444
Ophiacodontidae	13	4.8	36	341	0.001	446
Caseidae	13	1.8	28	371	0.001	438
Captorhinidae	13	3.6	24	354	0.001	432
Protorothyrididae	13	7.1	6	305	0.001	414

Table 6.9. Results of first- and second-order taxon jacknife analyses of the consensus matrix. See caption of Table 6.3 for key. Taxa are listed in order of decreasing number of caused incompatibilities. Turtles are highlighted in bold.

6.3.3.3.2 Compatibility Analysis

Results of 'Number', CCSR and LQP compatibility and boildown analyses of the consensus matrix are shown in table 6.10. Average 'Number', CCSR and LQP compatibility values for the whole dataset were 83.24, 0.7144 and 0.1109, respectively. In the 'Number' analysis, character 14 was shown to be the most incompatible. However, using both the CCSR and LQP tests, character 22 was the worst character initially, with values of 1.257 and 1, respectively.

Char.	No.	Fuzzy No.	CCSR	Fuzzy CCSR	LQP	Fuzzy LQP	No. Boil	Fuzzy No. Boil	CCSR Boil	Fuzzy CCSR Boil	LQP Boil	Fuzzy LQP Boil
1	83	140	0.65	0.588	**0.002**	**0.002**	124	108=	130	135	-	-
2	U	-	-	-	-	-	-	-	-	-	-	-
3	58	60	0.745	0.743	0.149	0.156	113=	124	121	98	41	40
4	53	62	0.568	0.528	**0.049**	**0.028**	-	-	-	-	124=	102=
5	65	66	0.832	0.819	0.232	0.242	-	-	-	44	30	36
6	134	323	0.957	0.835	0.154	**0.038**	5=	4	21	33	59	70
7	52	54	0.668	0.671	0.104	0.098	108=	111=	96	96	52	48
8	117	158	1.071	1.007	0.73	0.509	33=	51	3	12	3	12
9	40	40	0.854	0.865	0.394	0.421	108=	111=	104	25	14	14=
10	118	246	0.935	1.051	0.179	0.625	32	17	38	4	32	4
11	101	190	0.767	0.665	**0.005**	**0.002**	53=	38=	51	60	62	71
12	U	-	-	-	-	-	-	-	-	-	-	-
13	67	89	0.56	0.446	**0.002**	**0.002**	117=	125=	123	112	110	105=
14	142	361	1.025	1.009	0.776	0.548	1	1	7	10	2	6
15	52	53	0.75	0.727	0.203	0.189	125=	122=	112	48	27	33
16	41	41	0.511	0.494	**0.048**	**0.045**	129=	129=	122	108	74	61
17	125	266	0.864	0.616	**0.002**	**0.001**	18=	18	54	82		132
18	66	78	0.59	0.444	**0.003**	**0.002**	98	102	97	110	112=	127
19	124	272	0.962	0.95	0.205	0.306	20=	13=	16	14	25	19
20	104	179	0.829	0.714	**0.014**	**0.015**	67=	44=	70	63	84	66
21	U	-	-	-	-	-	-	-	-	-	-	-
22	135	202	1.257	1.3	1	0.989	3=	27	1	1	1	1
23	119	197	1.009	1.032	0.467	0.514	27=	32=	9	8	13	9
24	78	97	0.709	0.623	**0.026**	**0.017**	94=	103=	89	75	65=	69
25	132	299	0.998	1.032	0.412	0.586	8=	7	10	6	10	5
26	91	109	0.919	0.892	0.294	0.286	72=	76	47	27	29	31
27	68	98	0.596	0.537	**0.005**	**0.002**	116	116=	118	126	-	120=
28	18	18	0.381	0.382	0.059	0.075	129=	129=	-	122	47	46
29	58	75	0.712	0.729	**0.01**	**0.049**	-	-	88	68	58	52
30	97	172	0.724	0.56	**0.001**	**0.002**	86=	62=	101	97	101	113=
31	108	194	0.819	0.659	**0.004**	**0.004**	62=	72	85	101	115	112
32	131	323	0.949	0.928	0.104	0.169	8=	3	18	17	28	32
33	96	117	0.985	0.98	0.432	0.439	49=	69	11	13	11	13
34	78	115	0.698	0.71	**0.029**	0.064	-	-	-	77	71	59
35	115	195	0.897	0.837	0.097	0.112	39	37	53	52	73	68
36	99	192	0.727	0.556	**0.001**	**0.001**	-	-	127	133	-	116=
37	115	164	0.977	0.851	0.353	0.151	29=	62=	14	26	16	37
38	95	181	0.671	0.439	**0.001**	**0.001**	91=	78	-	136	-	-
39	94	171	0.669	0.444	**0.001**	**0.001**	117=	-	-	-	-	-
40	15	15	0.343	0.344	**0.038**	**0.035**	-	-	-	130	64	57
41	105	206	0.794	0.625	**0.001**	**0.001**	70=	36	84	79	-	105=
42	139	307	0.96	0.718	0.08	**0.003**	2	6	19	59	39	83
43	21	21	0.446	0.448	0.1	0.105	125=	125=	116	118	48	43
44	52	57	0.531	0.464	**0.014**	**0.011**	-	-	-	-	70	75
45	89	134	0.692	0.47	**0.001**	**0.001**	-	-	-	-	116=	-
46	86	118	0.668	0.378	**0.001**	**0.001**	94=	103=	95	125	-	-
47	88	90	1.164	1.159	0.707	0.685	76	84=	2	2	4	2
48	107	199	0.758	0.49	**0.001**	**0.001**	72=	94=	91	120	-	-
49	122	238	0.868	0.614	**0.002**	**0.001**	24=	25	55	85	103	113=
50	104	210	0.737	0.518	**0.001**	**0.001**	113=	113=	113	129	-	-
51	131	317	0.913	0.78	**0.015**	**0.014**	11	5	37	43	90	72
52	110	214	0.758	0.453	**0.001**	**0.001**	44=	62=	105	132	-	-
53	101	226	0.726	0.556	**0.001**	**0.001**	67=	38=	83	92	-	125=
54	C	-	-	-	-	-	-	-	-	-	-	-
55	101	174	0.719	0.493	**0.001**	**0.001**	93	79=	-	-	119=	120=
56	105	174	0.881	0.901	0.139	0.25	53=	40=	57	31	55	38
57	85	148	0.671	0.609	**0.004**	**0.003**	-	-	-	-	97	97
58	80	129	0.663	0.624	**0.011**	**0.015**	-	-	-	-	87=	80=
59	85	144	0.678	0.599	**0.005**	**0.01**	-	125=	-	-	98	92=
60	59	73	0.544	0.46	**0.006**	**0.001**	-	-	-	-	-	-
61	118	228	0.93	0.961	0.191	0.315	29=	23=	31	16	35	21

62	121	238	0.904	0.755	**0.044**	**0.021**	26	20	44	42	54	62
63	87	111	0.765	0.648	**0.007**	**0.006**	81	82=	60	62	53	64
64	108	154	1.02	1.023	0.536	0.561	53=	56=	13	11	8	8
65	124	247	0.893	0.68	**0.016**	**0.001**	20=	21	52	78	116=	102=
66	120	269	0.862	0.716	**0.002**	**0.01**	27=	15	40	53	68=	77
67	113	180	0.931	0.895	0.123	0.15	37	48	30	23	37	35
68	63	77	0.532	0.414	**0.001**	**0.001**	102=	96=	99	99	122=	116=
69	U	-	-	-	-	-	-	-	-	-	-	-
70	87	132	0.669	0.469	**0.001**	**0.001**	89=	79=	98	103	-	128=
71	23	23	0.491	0.484	0.207	0.21	-	-	-	100	33	30
72	105	149	0.946	0.921	0.312	0.308	53=	56=	25	20	23	27
73	38	40	0.488	0.436	**0.034**	**0.029**	-	-	-	-	85=	78
74	122	243	0.868	0.674	**0.002**	**0.001**	24=	23=	45	74	91	92=
75	101	210	0.721	0.544	**0.001**	**0.001**	53=	42	65	83	107=	130=
76	134	346	0.964	0.922	0.196	0.162	5=	2	23	18	77	29
77	118	197	1.046	1.072	0.706	0.679	29=	34	6	3	5	3
78	127	264	0.909	0.738	**0.023**	**0.007**	15=	19	33	64	78	90
79	135	295	0.975	0.822	0.268	**0.014**	3=	9	17	39	31	89
80	91	136	0.778	0.727	**0.027**	**0.03**	94=	82=	111	105	92=	92=
81	U	-	-	-	-	-	-	-	-	-	-	-
82	127	223	0.997	0.847	0.48	0.099	15=	28	12	30	12	47
83	110	187	0.847	0.64	**0.013**	**0.005**	41=	49	58	70	76	87=
84	29	42	0.286	0.323	**0.001**	**0.002**	123	118=	119	127	111	100
85	102	160	0.853	0.804	**0.049**	0.052	62=	52=	68	50	67	54=
86	123	226	0.909	0.667	**0.04**	**0.003**	22=	30=	35	71	61	99
87	89	107	0.916	0.892	0.268	0.273	70=	77	42	32	26	26
88	129	222	0.957	0.694	0.137	**0.001**	12	26	22	73	65=	98
89	72	93	0.625	0.506	**0.004**	**0.001**	108=	113=	115	119	102	116=
90	80	129	0.663	0.624	**0.011**	**0.015**	-	-	-	-	87=	80=
91	110	204	0.826	0.632	**0.003**	**0.001**	44=	50	74	87	92=	105=
92	74	118	0.59	0.497	**0.002**	**0.002**	86=	81	87	84	82	91
93	C	-	-	-	-	-	-	-	-	-	-	-
94	30	33	0.397	0.365	**0.006**	**0.006**	-	-	-	-	105=	92=
95	105	193	0.777	0.533	**0.001**	**0.001**	72=	75	93	109	122=	-
96	74	109	0.731	0.821	0.114	0.314	-	-	128	40	50	34
97	30	30	0.623	0.622	0.187	0.188	117=	116=	79	46	24	23
98	U	-	-	-	-	-	-	-	-	-	-	-
99	104	175	0.785	0.565	**0.002**	**0.001**	67=	59=	81	95	-	133
100	U	-	-	-	-	-	-	-	-	-	-	-
101	75	78	0.975	0.988	0.368	0.406	82=	91=	41	15	20	17
102	81	113	0.69	0.534	**0.005**	**0.002**	-	-	129	-	114	128=
103	47	72	0.387	0.345	**0.001**	**0.001**	108=	100	109	104	112=	124
104	87	96	0.933	0.84	0.202	0.129	82=	86	56	61	40	49
105	87	143	0.698	0.633	**0.003**	**0.005**	99=	98=	106	113	116=	125=
106	54	66	0.477	0.388	**0.002**	**0.002**	117=	118=	114	111	95	101
107	107	185	0.83	0.719	**0.023**	**0.012**	51	46	61	69	87=	82
108	57	57	0.773	0.747	0.127	0.11	99=	103=	92	86	45	41=
109	127	272	0.937	0.861	0.124	0.087	14	11=	29	29	75	51
110	94	166	0.711	0.524	**0.002**	**0.001**	72=	62=	73	89	119=	105=
111	80	142	0.576	0.385	**0.001**	**0.001**	94=	87	103	114	-	-
112	29	29	0.612	0.621	0.224	0.209	125=	118=	78	45	22	25
113	88	129	0.658	0.397	**0.001**	**0.001**	89=	101	94	123	-	130=
114	105	177	0.785	0.546	**0.003**	**0.001**	77=	94=	90	116	-	-
115	96	170	0.695	0.464	**0.002**	**0.001**	61	55	67	93	104	120=
116	114	203	0.894	0.782	0.078	**0.03**	38	32=	46	49	56	63
117	125	241	0.939	0.813	0.089	**0.041**	18=	22	34	37	57	53
118	55	55	1.079	1.069	0.567	0.586	99=	103=	4	5	7	7
119	69	115	0.529	0.372	**0.001**	**0.001**	-	-	-	-	-	-
120	91	175	0.678	0.504	**0.001**	**0.001**	113=	113=	120	128	-	-
121	14	14	0.336	0.337	**0.039**	**0.027**	-	-	-	-	94	76
122	41	41	0.841	0.827	0.298	0.345	117=	122=	80	35	19	18
123	100	124	0.919	0.808	0.229	0.101	53=	73=	43	66	43	50
124	110	259	0.815	0.805	**0.001**	**0.02**	40	16	64	34	96	54=
125	98	164	0.743	0.536	**0.002**	**0.001**	125=	125=	124	137	-	-

126	134	269	0.96	0.749	0.2	**0.021**	5=	13=	20	57	46	65
127	111	156	0.992	0.925	0.441	0.279	43	59=	15	22	17	22
128	90	154	0.68	0.526	**0.001**	**0.001**	107	108=	126	134	-	-
129	49	51	0.658	0.661	0.07	0.079	104=	108=	102	90	42	45
130	123	232	0.902	0.658	**0.018**	**0.001**	22=	29	36	76	63	110=
131	110	190	0.805	0.549	**0.001**	**0.001**	44=	47	71	91	107=	110=
132	100	168	0.768	0.599	**0.002**	**0.002**	48	67	75	81	119=	104
133	107	195	0.82	0.736	**0.006**	**0.008**	49=	43	62	67	105=	87=
134	98	150	0.835	0.773	0.059	0.052	91=	70=	86	56	60	58
135	U	-	-	-	-	-	-	-	-	-	-	-
136	46	53	0.468	0.43	**0.007**	**0.004**	-	-	-	-	-	120=
137	111	183	0.796	0.483	**0.001**	**0.001**	44=	61	82	102	-	-
138	103	150	0.91	0.898	0.21	0.257	53=	68	32	21	34	24
139	92	157	0.671	0.435	**0.001**	**0.001**	102=	107	107	124	-	-
140	90	129	0.831	0.818	0.125	0.181	77=	66	59	36	38	28
141	26	26	0.445	0.442	0.14	0.139	-	-	-	106	36	39
142	82	147	0.604	0.422	**0.001**	**0.001**	104=	93	108	115	-	-
143	114	170	0.961	0.877	0.197	0.074	21=	52=	27	28	51	56
144	54	94	0.408	0.318	**0.001**	**0.001**	104=	96=	110	117	-	-
145	110	190	0.843	0.651	**0.01**	**0.003**	62=	89=	100	107	100	116=
146	105	181	0.838	0.773	**0.035**	**0.025**	52	44=	63	65	83	79
147	115	197	0.908	0.82	0.108	0.054	21=	35	50	58	79	73
148	U	-	-	-	-	-	-	-	-	-	-	-
149	126	277	0.898	0.777	**0.021**	**0.018**	17	11=	49	54	85=	67
150	80	92	0.818	0.757	0.146	0.113	88	89=	72	72	49	44
151	104	172	0.852	0.821	0.069	0.088	62=	40=	76	55	80	60
152	36	36	0.531	0.516	**0.022**	**0.03**	-	-	125	121	99	84
153	80	155	0.591	0.478	**0.001**	**0.001**	79=	58	-	88	-	105=
154	97	166	0.73	0.557	**0.002**	**0.001**	108=	98=	-	-	-	-
155	110	175	0.845	0.65	**0.009**	**0.003**	41=	52=	69	94	-	113=
156	11	11	0.235	0.234	**0.016**	**0.012**	-	-	-	131	81	74
157	96	129	0.855	0.783	0.091	0.063	82=	84=	48	38	44	41=
158	116	212	0.828	0.594	**0.004**	**0.002**	33=	30=	66	80	107=	92=
159	83	87	1.011	1.027	0.523	0.519	79=	88	8	9	9	10
160	89	162	0.663	0.506	**0.001**	**0.001**	117=	118=	117	138	124=	-
161	95	118	0.927	0.898	0.312	0.343	62=	73=	24	19	21	20
162	132	301	0.95	0.788	0.083	**0.005**	8=	8	26	41	68=	85=
163	129	284	0.936	0.773	0.059	**0.008**	13	10	28	47	72	85=
164	80	80	0.961	0.926	0.459	0.407	85	91=	39	24	18	16
165	38	38	0.878	0.875	0.436	0.452	-	-	77	51	15	14=
166	C	-	-	-	-	-	-	-	-	-	-	-
167	C	-	-	-	-	-	-	-	-	-	-	-
168	104	122	1.08	1.043	0.664	0.523	53=	70=	5	7	6	11

Table 6.10. Table showing the results of compatibility tests and the order of removal in subsequent boildowns of each character in the consensus matrix using CCSR, "Number" and LQP methods with and without the fuzzy compatibility. C = constant character and U = uninformative character. LQP results that are significant at the 5% level are highlighted in bold.

Analysis of the consensus matrix with 54 characters that fail the LQP test at the 10% level removed yielded three MPTs (L=449, CI=0.561, RI=0.740). The strict consensus of these trees (Fig. 6.20a) placed turtles within the Diapsida, as the sister-group of the eosauropterygians and placodonts, the hypothesis championed by Rieppel and Reisz (1999). Analysis of the same data using backbone constraints to find the shortest tree with turtles within the Anapsida produced four MPTs (L=454, CI=0.555, RI=0.734), the strict consensus of which placed turtles as the sister-group of the procolophonoids. Templeton tests showed that the diapsid hypothesis is not a significantly better fit to the data than the anapsid hypothesis (Templeton p-values range from 0.3841 to 0.4458).

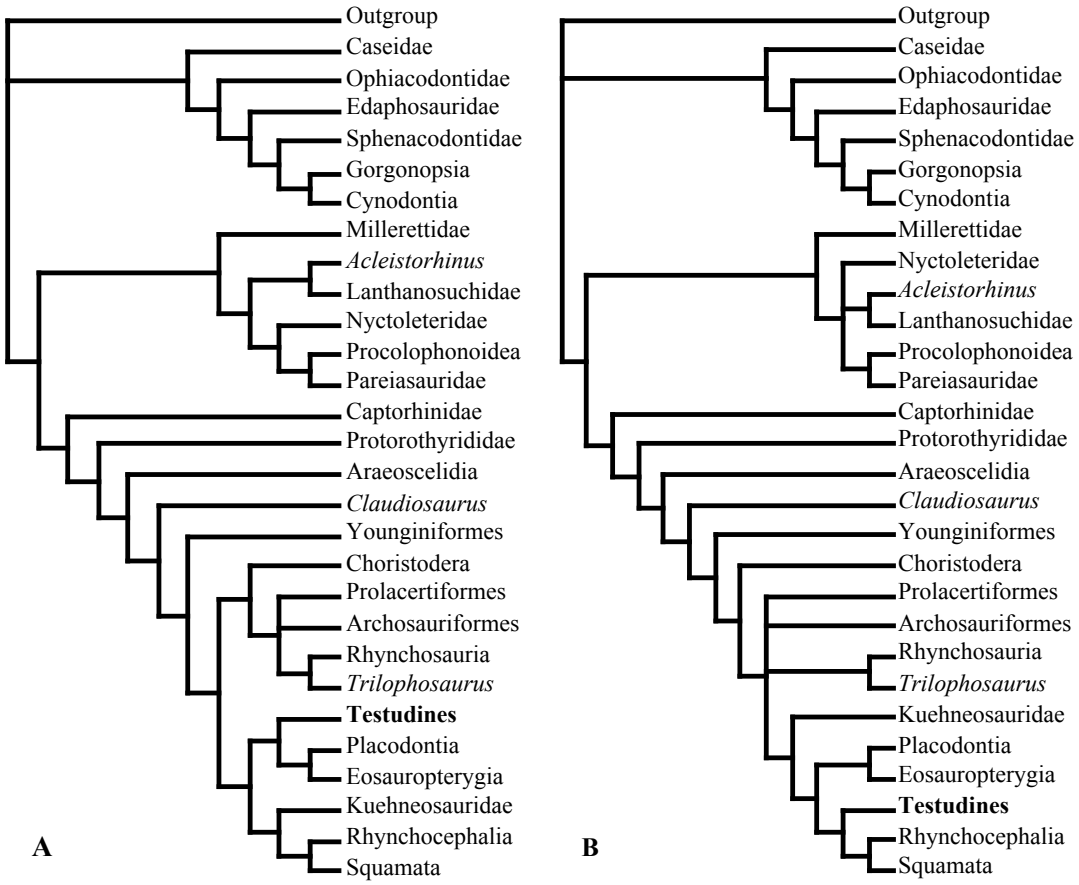


Figure 6.20. Strict consensus trees of the MPTs for the consensus matrix with characters excluded that are not significantly better than random at the (A) 10% and (B) 5% level during LQP tests.

Removal of the 64 characters that failed the LQP test at the 5% level produced 12 MPTs (L=395, CI=0.577, RI=0.758) in parsimony analysis. The strict consensus of these (Fig. 6.20b) trees placed turtles within the Diapsida as the sister-group of the rhynchocephalians and squamates. Analysis employing a backbone constraint forcing turtles into the Anapsida returned 12 MPTs (L=404, CI=0.564, RI=0.745), the strict consensus of which again placed turtles as the sister-group of the procolophonoids.

Templeton tests showed that the diapsid hypothesis is not significantly better than the anapsid hypothesis, although the p-values approached significance (p-values range from 0.0947 to 0.1282).

6.3.3.3.3 Boildown Analysis

With all three compatibility measures, Templeton tests at each stage of the boildowns (Fig. 6.21) showed a trend from initially providing no support for either the anapsid or diapsid hypothesis towards support for the diapsid hypothesis. Using all three compatibility measures, significant support for the diapsid hypothesis was reached. The 'Number' analysis reached 5% significance when 85 characters had been removed, the CCSR analysis when 69 characters had been removed, and the LQP analysis when 83 characters had been removed. Both the 'Number' and LQP analyses lost statistical significance when only a few extra characters were removed, whereas the CCSR analysis significantly supported the diapsid hypothesis until 128 characters had been removed.

6.3.3.3.4 Fuzzy Compatibility

Results of 'Number', CCSR and LQP fuzzy compatibility tests and subsequent boildowns of the consensus matrix are shown in table 6.10. The results of Templeton tests at each stage of the boildowns are shown in figure 6.22. Generally the results using fuzzy compatibility are very similar to those using the normal compatibility method. Again, the trend towards support for the diapsid hypothesis is slightly greater at first than when using non-fuzzy compatibility. However, the 'Number' analysis only reached significance when 120 characters had been removed, far later than when not using the fuzzy method. Both the CCSR and LQP reached significant support for the diapsid hypothesis at the 5% level quicker when using fuzzy compatibility, when 50 and 68 characters had been removed, respectively. As with the non-fuzzy analysis, the CCSR method retained significant support longer than either of the other compatibility measures.

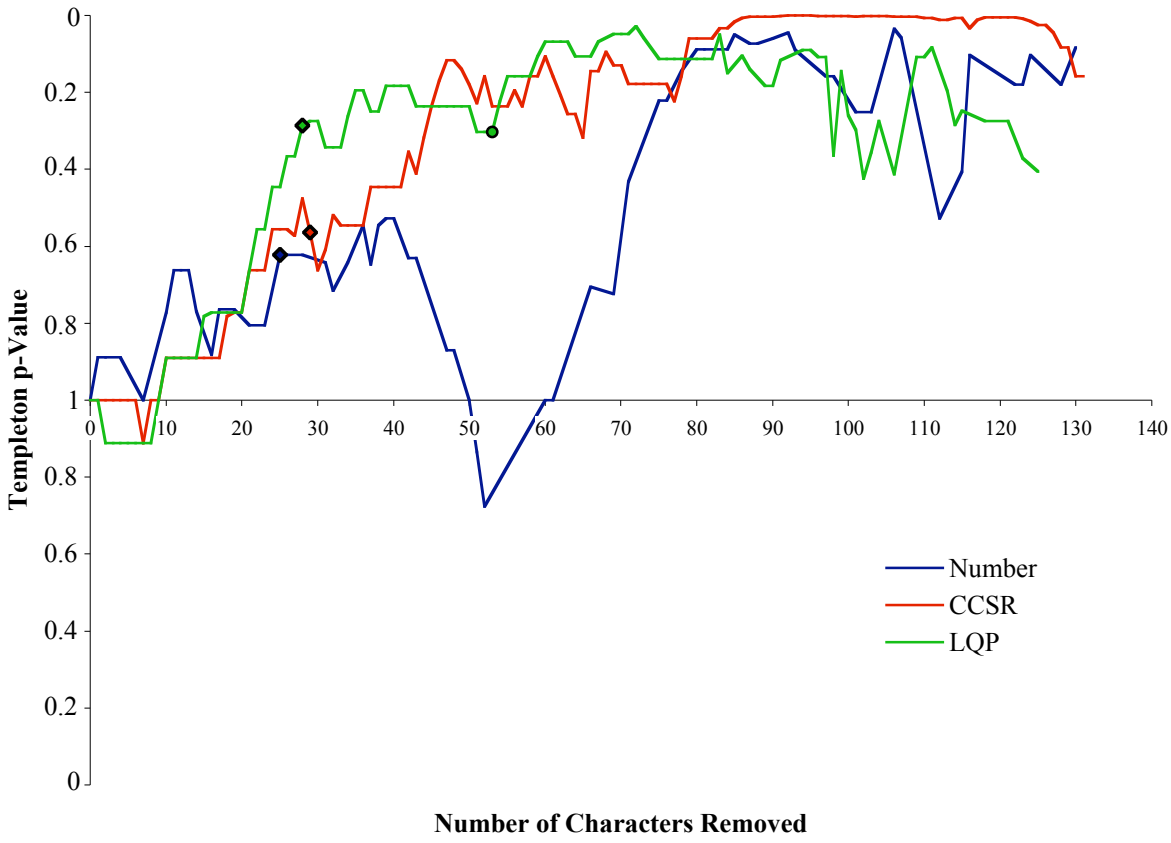


Figure 6.21. Line chart showing the Templeton test p-value when comparing the fit of the consensus data matrix to trees with backbone constraints forcing turtles within the Anapsida and Diapsida at each stage of number, CCSR and LQP boildowns. The x-axis represents the number of characters removed by the boildown. The y-axis is the Templeton test p-value. Values above the x-axis indicate that the diapsid hypothesis is more parsimonious than the anapsid hypothesis, whereas those values below the x-axis indicate that the anapsid hypothesis is more parsimonious. Coloured diamonds indicate the point at which the boildown bootstrap identified maximum bootstrap values for each boildown. The green circle is the point at which characters removed in the LQP boildown have an LQP value lower than 0.05, and are therefore significantly more compatible than random characters.

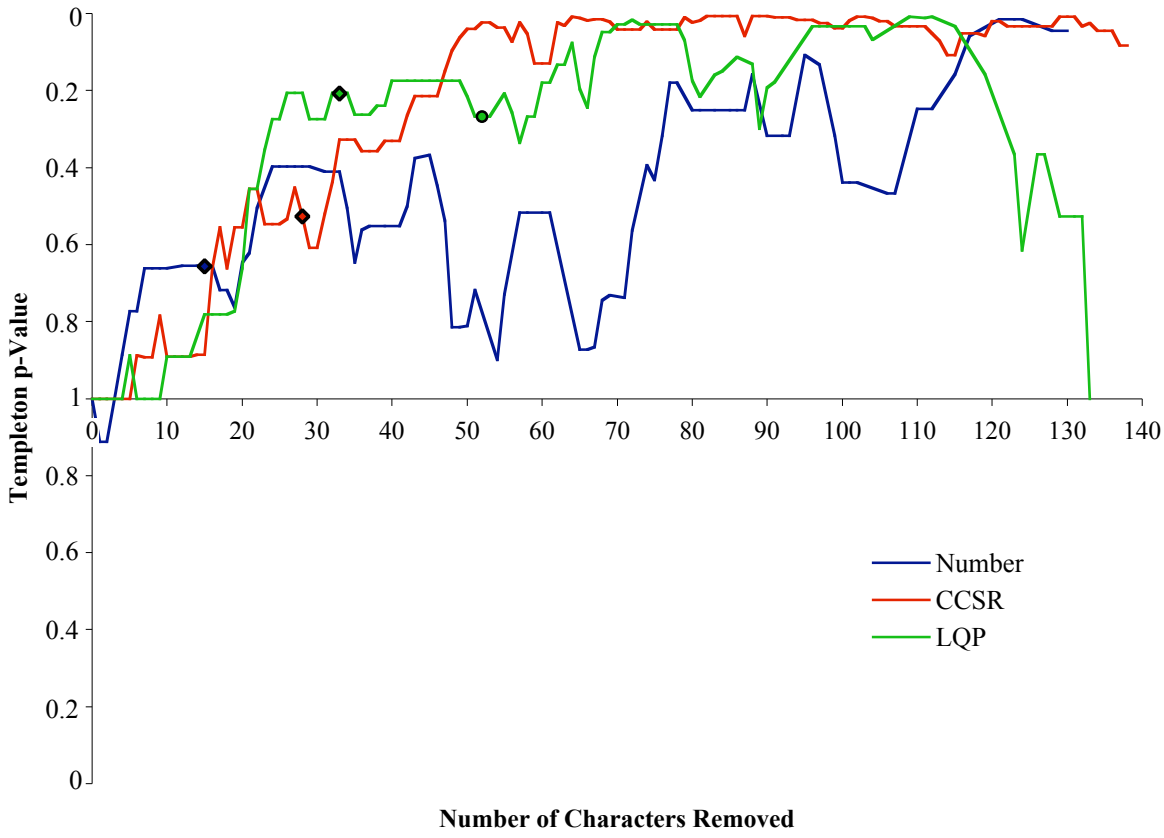


Figure 6.22. Line chart showing the Templeton test p-value when comparing the fit of the consensus data matrix to trees with backbone constraints forcing turtles within the Anapsida and Diapsida at each stage of fuzzy number, CCSR and LQP boildowns. The x-axis represents the number of characters removed by the boildown. The y-axis is the Templeton test p-value. Values above the x-axis indicate that the diapsid hypothesis is more parsimonious than the anapsid hypothesis, whereas those values below the x-axis indicate that the anapsid hypothesis is more parsimonious. Coloured diamonds indicate the point at which the boildown bootstrap identified maximum bootstrap values for each boildown. The green circle is the point at which characters removed in the LQP boildown have an LQP value lower than 0.05, and are therefore significantly more compatible than random characters.

6.4 Discussion

The affinities of Testudines remains one of the most intractable problems in phylogenetic systematics despite recent claims that strongly supported solutions have been found (Rieppel and Reisz, 1999; Lee, 2001). Reanalysis of the datasets of Rieppel and Reisz (1999) and Lee (2001) produced trees with moderately high bootstrap support and decay indices (see Fig. 6.9), suggesting that their results are indeed relatively robust. Lee (2001) even showed that using his data the best tree placing turtles within the Anapsida is significantly better in Templeton tests than the best tree supporting a diapsid-turtle link. However, it is obviously the case that both hypotheses cannot be true.

As recognised by many authors of the past (see history of the debate section in this chapter), most of the problems in placing the turtles within the Amniota probably stem from their highly divergent morphology. This single adaptation seems to have drastically altered the morphology of much of the body within a geologically very short timescale. In fact, the earliest described turtle (an undescribed turtle shell fragment from the Middle Triassic, Ladinian, has been rumoured (see Lucas *et al.*, 2000)), *Proganochelys quenstedti*, from the Upper Triassic (Norian) appears far more morphologically similar to modern-day chelonians than to any of its contemporary potential relatives.

6.4.1 Locating the Disagreement: The Consensus Matrix

Rieppel and Kearney (2002) claimed that stagnated phylogenetic debates are usually the result of errors in morphological assessment, character coding and character scoring, and that solutions can only be found by reassessing the construction and scoring of controversial characters. If this were true, it could be predicted that by identifying areas of conflict within the data and working to resolve that conflict, the true position of turtles could be attained. Therefore, one initial aim of this study was to identify the differences in character construction and scoring between the data of Rieppel and Reisz (1999) and Lee (2001). A second aim was to look at the signal produced by the data that was agreed upon by the two authors.

By creating a consensus matrix, all disagreements between the two major opponents in the debate were identified and removed. However, during creation of this matrix it was discovered that only 77 of the 4872 scorings differed between the two analyses. The differences in scoring between the matrices of Rieppel and Reisz (1999) and Lee (2001) are shown in table 6.11. The fact that seemingly strong support for two hypotheses as diametrically opposed as the anapsid- and diapsid-turtle hypotheses can be attained by

rescoring less than 2% of data suggests that, whichever, if either, scorings are correct, caution should be used when drawing conclusions from the results. When analysed using parsimony, the consensus matrix produced another unexpected result. The three MPTs found were identical to the two MPTs yielded by analysis of the data of Rieppel and Reisz (1999) and the single MPT from the analysis of the data of Lee (2001). The most simple explanation for the stagnation of the debate over turtle affinities would be that it is a result of subjectivity of interpretations of morphology. However, that results of analysis of the undisputed 98% of the data contained in the consensus matrix could not separate the alternative hypotheses suggests that perhaps the problems in phylogeny reconstruction are reflecting reality, and that either there is a real problem of convergence in turtle morphology, or the disparity between turtles and other known taxa is so great that homology assessment is impossible. Even so, the data for which there are differences in character scorings between the authors need to be re-examined. Randomisation tests, in which 77 character scorings (equal to the number of actual conflicts) marked as 'scoring conflicts' were randomly distributed 10,000 times within a matrix with identical dimensions to the consensus matrix, were used to test the null hypothesis that the scorings over which Rieppel and Reisz (1999) and Lee (2001) disagree were distributed randomly throughout the matrix. Results showed that five characters (41, 63, 69, 74 and 140) and three taxa (Eosauropterygia, Pareiasauridae and Testudines) contained significantly more ($p<0.05$) of the conflicting scorings than would be expected by chance alone, and, therefore, re-examination of these characters and taxa should be prioritised. Interestingly, the taxon containing most scoring conflicts (14) was the Testudines, followed by the Pareiasauridae (10) and Eosauropterygia (8), the two taxa suggested by Lee (2001) and Rieppel and Reisz (1999) respectively as the sister-taxon to turtles. This is consistent with at least two explanations: that these taxa are the most difficult to score, or the authors have put more effort into checking the scorings of the taxa that they consider most important in attempting to resolve the affinities of turtles.

Char	Taxon	Rieppel & Reisz (1999)	Lee (2001)
6	Eosauropterygia	1	0
19	Lanthanosuchidae	1	0
29	Keuhneosauridae	0	?
31	Eosauropterygia	0&1	1
32	Acleistorhinidae	1	0
32	Lanthanosuchidae	1	0
41	Acleistorhinidae	1	?
41	Placodontia	1	?
41	Eosauropterygia	1	?
46	Pareiasauridae	?	1
59	Eosauropterygia	1&2	2
63	Ophiacodontidae	?	0
63	Millerettidae	?	0
63	Nycteroleteridae	?	0
63	Younginiformes	?	0
63	Rhynchocephalia	?	1
63	Edaphosauridae	1	0
63	Sphenacodontidae	1	0
63	Squamata	0	1
64	Pareiasauridae	0&1	1
64	Placodontia	?	1
65	Pareiasauridae	2	1&2
69	Outgroups	1	0
69	Caseidae	1	0
69	Ophiacodontidae	1	0
69	Edaphosauridae	1	0
69	Sphenacodontidae	1	0
69	Gorganopsia	1	0
69	Cynodontia	1	0
69	Captorhinidae	1	0
69	Protorothyrididae	1	0
69	Millerettidae	1	0
69	Acleistorhinidae	1	0
69	Lanthanosuchidae	1	0
69	Nycteroleteridae	1	0
69	Araeoscelididae	1	0
69	Claudiosauridae	1	0
69	Younginiformes	1	0
72	Acleistorhinidae	1	0
72	Lanthanosuchidae	1	0
73	Testudines	0&1	0
74	Captorhinidae	1	0
74	Millerettidae	1	0
74	Procolophonoidea	1	0
74	Nycteroleteridae	?	0
77	Testudines	0&1	1
77	Placodontia	1	0
80	Pareiasauridae	0	1
82	Testudines	0&2	2
83	Testudines	1	0
83	Placodontia	1&2	1
87	Pareiasauridae	0	1
87	Eosauropterygia	1	0
89	Placodontia	0	1
97	Lanthanosuchidae	?	0
103	Pareiasauridae	0	1
120	Cynodontia	0	1
120	Testudines	0	?
121	Testudines	1	?
124	Testudines	1	0
125	Millerettidae	0	?
125	Pareiasauridae	0&1	?
126	Testudines	1	2
127	Testudines	0&2	2
128	Lanthanosuchidae	?	0
139	Protorothyrididae	1	0
139	Testudines	1	0
140	Pareiasauridae	0	1
140	Placodontia	1	0
140	Eosauropterygia	1	0&1
142	Testudines	1	0
142	Eosauropterygia	1&2	0&1
144	Pareiasaurs	0&?	1
150	Pareiasaurs	1&?	?
152	Testudines	0&1	0
164	Testudines	0&1	0
167	Testudines	0&1	1

Table 6.11. Scoring differences found between the data matrices of Rieppel and Reisz (1999) and Lee (2001) that were replaced with missing data in the consensus matrix.

6.4.2 Identification of Problematic Taxa

Problems associated with the divergent morphology of turtles are also revealed by the results of the taxon jacknife analyses on the three data matrices studied here (tables 6.3, 6.6 and 6.9). In all cases, turtles were the cause of many incompatibilities, in both first and second order jacknife analyses. They also always exhibited extremely high p-values (0.995, 0.996 and 0.989 in the matrices of Rieppel and Reisz, 1999, Lee, 2001, and the consensus matrix, respectively) in first order jacknife randomisation tests, so that the null hypothesis, that they caused no fewer incompatibilities than a randomly permuted taxon, could not be rejected. In fact, their p-value was so high that if the test were two-tailed it would be concluded that turtles are more incompatible with the rest of the data than a randomly permuted taxon. What such a result means in practice is difficult to determine. It should also be noted that turtles were not the only taxon in the matrices to exhibit such high levels of incompatibility. The Cynodontia had even higher p-values of 0.999, 0.999 and 0.996, and the Pareiasauridae had values of 0.878 and 0.918 in the data of Lee (2001) and the consensus matrix, respectively. As previously mentioned, results of jacknife randomisation tests are not always easy to interpret. If two taxa are extremely similar, they will appear to be a good fit to the data and exhibit low p-values even if they are highly incompatible with all other taxa. This is simply because removal of one of the two similar taxa does not remove many incompatibilities, because the second taxon, that exhibits many of the same incompatibilities, will still be present. For this reason, using the matrix of Rieppel and Reisz (1999), where the Pareiasauridae is split into three exemplar taxa (*Anthodon*, *Bradysaurus* and *Scutosaurus*), none exhibits high p-values during taxon jacknife randomisation tests, even though as a group they are probably the cause of many incompatibilities. Although the taxon jacknife suffers from this problem of interpretation, so that some problematic taxa may erroneously be identified as highly compatible, any taxa that exhibit high p-values, such as the turtles, cynodontids and pareiasaurs, can confidently be assessed to be especially problematic within the data. This may be part of the reason that using the same analysis techniques, and essentially the same character constructions, Rieppel and Reisz (1999) and Lee (2001) have independently produced persuasive evidence supporting dramatically conflicting hypotheses.

It is not only taxon jacknife compatibility tests that can highlight the problematic nature of the Testudines in these datasets. Leaf stability analyses on the bootstrap trees from the three analyses indicated that turtles were the least stable taxon in each case, and

were especially unstable in the analysis of the consensus matrix. This is not a surprising result given the two highly distinct positions of turtles found in the MPTs using this matrix, but could also result from the fact that turtles are coded as unknown for many characters because of the high number of disagreements between the authors over the scorings of this taxon. Again, the Cynodontia are also shown by leaf stability tests to be highly unstable in the bootstrap trees from all three matrices, along with a second set of synapsids, the Gorgonopsia. This high degree of agreement using *a priori* and *a posteriori* methods for identifying problematic taxa must allow us to afford extra confidence in their results.

Further evidence that the results obtained by Rieppel and Reisz (1999) and Lee (2001) are not as strong as they first appear come from the results of sequential sister-group removal. This method especially highlighted a potential weakness in the results of the analysis of Lee (2001). Lee initially found turtles to be sister-group to the anapsid pareiasaurs, which nests the Testudines deep within the Anapsida. If strong evidence were present for an anapsid-turtle link, it would thus be expected that if the pareiasaurs were removed from the analysis the turtles would stay within the Anapsida, probably as the sister-group to the procolophonoids (the sister-group to Testudines + Pareiasauridae in the original analysis). However, this is not the case. When the pareiasaurs are removed, turtles move to become the sister-group of the Rhynchosauria within the Diapsida. This shows that the nesting of turtles within the Anapsida may not be particularly strong. Using the data of Rieppel and Reisz (1999), two successive sister-groups had to be removed before turtles left the Diapsida and joined the Anapsida. However, these two sister-groups comprised five taxa, so a large number of diapsid taxa were removed. With the consensus matrix, removal of the two sister-groups in the MPTs from the analysis of all data (the pareiasaurs and sauropterygians) led to turtles nesting within the Diapsida. A further two sister-groups (4 taxa) had to be removed before any MPTs showed an anapsid positioning again. Together, the results of successive sister-group removal on the three datasets suggest that with the data of Rieppel and Reisz (1999) and the consensus matrix the turtles show stronger nesting within the Diapsida than they show nesting within the Anapsida in the data of Lee (2001). This appears counterintuitive when it is remembered that only the data of Lee (2001) achieved significant support for its most parsimonious hypothesis in Templeton tests. The lack of support of nesting within the Anapsida suggests that this Templeton test result may not be as strong as it first appears.

6.4.3 Identification of Competing Signals

It is clear from the result obtained here that there are at least two major signals (diapsid and anapsid placements for turtles) included in the data of Rieppel and Reisz (1999) and Lee (2001), and it appears that identifying the true signal from the subsignals is not simple. For that reason, the RSACW method was applied in order to try to identify exactly what subsignals were present within the dataset. Within the three datasets a number of major competing signals were identified, most of which appear to be linked to the problematic nature of the Testudines. The first is the position of the clade comprising the Gorgonopsia and Cynodontia. In the initial analysis of all three datasets this clade was positioned as derived synapsids, as would be expected from previous literature (e.g. see Rubidge and Sidor, 2001), a position that has not been disputed by either Rieppel and Reisz (1999) or Lee (2001). However, RSACW of all three datasets produced trees with this clade positioned outside of the Synapsida. Using the data of Lee (2001) they moved into the Anapsida, and with the data of Rieppel and Reisz (1999) they moved into the Diapsida. The constant factor in each case was that they moved to become the sister-group of the Testudines. This clearly shows that homoplastic signal exists in the data that supports a relationship between turtles and derived synapsids, a grouping that no recent morphologist has championed. Strangely, RSACW did not identify any subsignals in the data of Rieppel and Reisz (1999) that supported an anapsid position for turtles or a link to the Pareiasauridae, and similarly, using the data of Lee (2001) no subsignals were found that placed turtles within the Diapsida or supported a sauropterygians-turtle link. This suggests, given all of the other evidence indicating the presence of these two major subsignals in the data, that the RSACW method may not be a very effective method. RSACW of the consensus matrix did identify the signals positioning turtles within the Anapsida and Diapsida, but this is not surprising given that these two signals can be seen in the three MPTs from initial analysis of this data.

6.4.4 Compatibility as an Alternative to Parsimony

It is clear that the debate over turtle origins runs deeper than simply disagreements over morphological assessment between authors, and that only extreme rescoring of characters or identification of new characters can robustly resolve the conflict if only the currently popular methods of character analysis are employed. However, the data used here have been drawn up by leaders in the field, and it is likely that any problems with homology assessment are due in part to the divergent morphology of turtles rather than

poor character identification, coding and scoring alone. It is therefore likely that the debate can only be solved using parsimony if new, intermediate fossils are discovered.

However, other methods, such as compatibility, exist that can be used to investigate the current data further than simple parsimony-based analysis. Here these methods were employed in an attempt to provide evidence in support of one of the hypothesised turtle affinities.

Removing characters that are not significantly more compatible than random should strengthen the true signal by removing noise from the dataset. Therefore, removing characters that fail the LQP test at the 10% or 5% levels should increase support for any relationships that have a strong founding in the data. It was hoped that this procedure would, in the case of the three turtle datasets, strengthen the hypothesis that represented the true phylogeny of the Amniota. The results obtained were interesting. Using the data of Rieppel and Reisz (1999) and Lee (2001), removing characters that failed the LQP test did not lead to the initial most parsimonious hypothesis in either case being overthrown. However, looking at the support for the two hypotheses did show an interesting trend. Using the data of Rieppel and Reisz (1999), Templeton tests between the diapsid and anapsid hypotheses on the complete data produced a Templeton value of 0.4697 to 0.4880. When characters failing the LQP test were removed, this value decreased (0.3376 to 0.3429 with removal at the 10% level, and 0.3961 to 0.4227 at the 5% level), indicating an increase in the support for the diapsid hypothesis over the anapsid hypothesis. Although this change is very small, and the diapsid hypothesis does not become statistically significantly better than the anapsid hypothesis, it does show that, at the very least, removal of characters with high LQP values does not weaken the diapsid hypothesis using the data of Rieppel and Reisz (1999). When characters failing the LQP were removed from the data of Lee (2001), a different result was observed. As already mentioned, analysis of the entire data showed that the anapsid hypothesis was a significantly better fit to the data than the diapsid hypothesis (Templeton value = 0.045). When incompatible characters were removed from the data, this significance was lost, and the more characters that were removed, the lower the support for the anapsid hypothesis became (Templeton value ranged from 0.4458 to 0.5531 with removal at the 10% level, and 0.7518 to 0.8112 at the 5% level). Therefore, despite the anapsid hypothesis being the more parsimonious in analysis of all of these versions of the data, there are indications that much of its initial significant support stems from characters that are no more compatible with the rest of the data than a randomly permuted character. LQP analysis of the consensus matrix also

provided support for the diapsid hypothesis. Using these data, however, initial parsimony analysis did not provide support for either hypothesis, so it was a level playing field on which to run the LQP test. When characters failing the LQP test were removed, the diapsid hypothesis became the more parsimonious, and Templeton tests showed that although the difference between the two hypotheses did not reach significance (Templeton value ranged from 0.3841 to 0.4458 with removal at the 10% level, and 0.0947 to 0.1282 at the 5% level), it did approach it.

Similar results were obtained when boildown analyses were carried out on the three datasets. If anything, the results of these tests were even more marked, although, with a technique as little used as the boildown procedure, it is difficult to know how exactly to interpret the results. It is not known, for example, what trends would be expected with the methods used here if it were known that one of the hypotheses was true and strongly supported. A number of trends are possible, that could be produced by different circumstances. For example, imagine two opposing hypotheses, A and B, and that hypothesis A is supported by the data with a Templeton value of 0.5 when the two hypotheses are tested against each other. Therefore, the two hypotheses are not significantly different in their fit to the data. If a boildown is run on the data, a number of trends could be envisaged, depending on the underlying signals in the data.

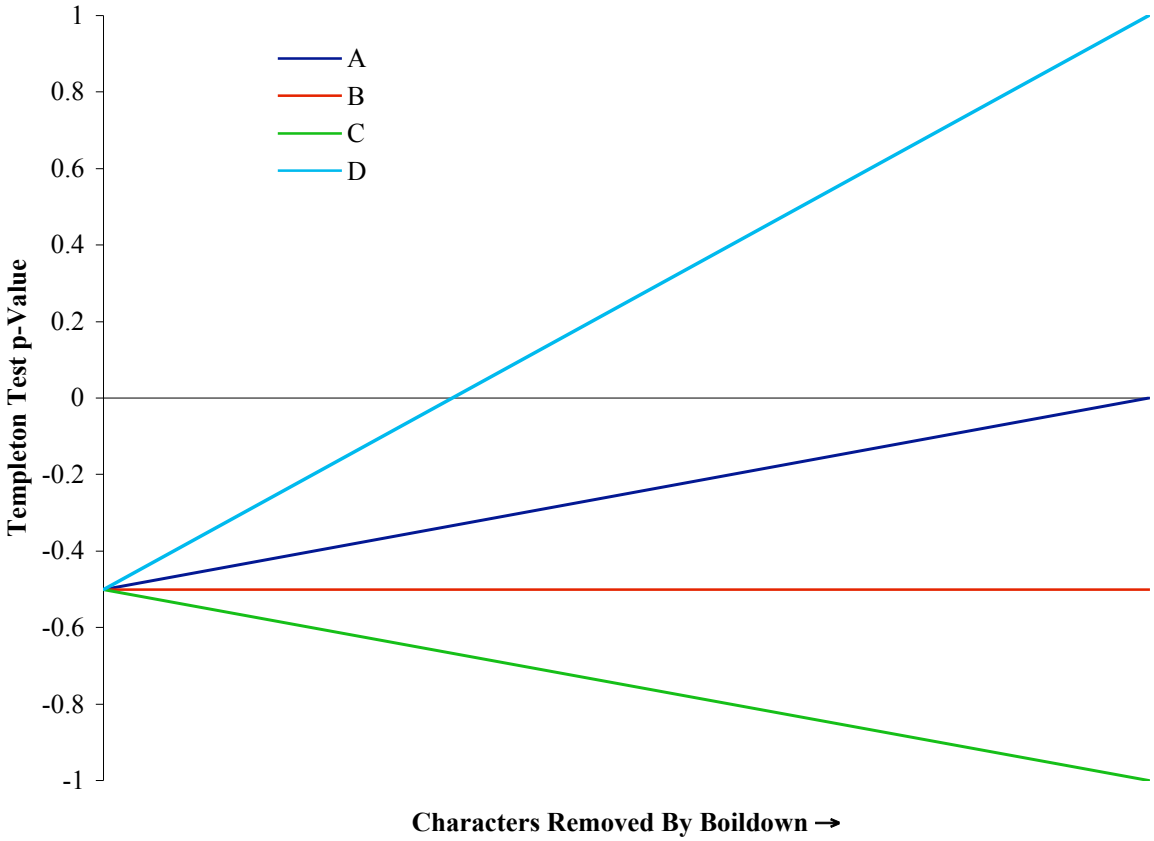


Figure 6.23. Hypothetical boildown results when two phylogenetic hypotheses are tested against each other. See text for possible explanations for recovering trends A to D in a boildown analysis.

- First, if characters are removed from the data by the boildown at random relative to the two hypotheses, then it is likely that more characters supporting hypothesis A will be removed, leading to the boildown initially trending towards a Templeton value of zero (Fig. 6.23a). This is simply because it is likely that hypothesis A will be supported by more characters in the data than hypothesis B, as it is the more parsimonious hypothesis. Therefore, when characters are chosen to be removed at random it is more likely that they will be characters supporting hypothesis A. This process is frequency dependent, in that all the time hypothesis A is supported by the larger set of characters it is more likely that one of those characters will be removed, whereas if this removal leads to there being more characters supporting hypothesis B, then it will become more likely that one of those characters will be removed. Therefore, when the trend reaches a Templeton value of zero it is likely to level out. If such a trend is produced, it suggests that there is no strong evidence in the data supporting either hypothesis over the other
- A second possibility is that the probability of removing a character supporting either hypothesis is equal. In this case, the trend is likely to be a horizontal line, so that the

Templeton test p-value will remain at around 0.5 (Fig. 6.23b). This situation is only possible if the characters supporting hypothesis B are slightly more compatible with the characters supporting neither hypothesis, so that the mutual compatibility of the higher number of characters that support hypothesis A is offset.

- A third possibility is that a strong signal supporting hypothesis A is present in the data, which is not unlikely given that this hypothesis is more parsimonious. In such a case, a large set of characters supporting the hypothesis is likely to be present in the data, leading to the boildown process identifying characters incompatible with this set as poor and removing them. In this situation the trend produced by the boildown is likely to be an iterative increase in the support for hypothesis A over hypothesis B, so that the p-value produced by Templeton tests between the two hypotheses is likely to increase (Fig. 6.23c).
- The final trend that is possible is a shift from support for hypothesis A towards support for hypothesis B (Fig 6.23d). This situation can occur if, despite parsimony supporting hypothesis A, the characters supporting this hypothesis are not mutually compatible, and the largest set of characters in the data supports hypothesis B. It has to be assumed that such a situation, and therefore such a trend would be rare.

It must be noted here that late on in the boildown procedure trends are likely to halt or reverse. There are a number of reasons for this. First, when the boildown has removed a large number of characters from the data, it starts to remove characters that are not particularly weak, so that before the process finds a group of completely pairwise-compatible characters it will have removed a number of characters that are highly, but not completely compatible with the data. Effectively, therefore, as the boildown nears its completion it begins randomly deleting useful information, which can lead to an alteration in the trend. Second, when a small number of characters remain, removal of a single character can cause a larger effect on the signal given by the data, because the relative number of steps each character contributes to the total tree length increases as the number of characters reduces. This means that near the end of the boildown, removal of a character supporting one hypothesis can lead to an apparently large trend on the graphs of Templeton values. Finally, near the end of the boildown it is possible that all characters supporting one of the hypotheses under investigation have been removed, so that only characters supporting the other hypothesis can be removed to affect the trend on the graph. By chance alone, some of these characters are likely to be removed, leading to the suggestion of evidence in support of the hypothesis that has no characters supporting it. In such a case, what is often observed on the Templeton graph is a strong trend towards one

hypothesis at the start of the boildown followed by a trend back towards a Templeton value of zero near to the end of the process. It is therefore an important point that the end of a boildown procedure is not the important result of the method. The trend towards the beginning of the procedure is more interesting, because at this point the data contains far more characters and the characters being removed are more incompatible that those removed later.

6.4.5 Ending the Boildown

The boildown procedures run in this study were run until either no incompatibility remained in the data (in the cases of the 'No' and CCSR boildowns), or until no permuted versions of any characters were more incompatible than the original character (in the case of the LQP boildown). However, one of the most problematic aspects of the boildown is judging how many characters should be removed before it can be considered that useful signal is being lost from the data rather than just noise. Until this point no methods for deciding when to stop the boildown have been suggested, so two were tested here. First, in the case of the LQP boildown it is simply possible to halt the process when all characters remaining in the data are significantly more compatible with the rest of the data than would be expected from a random permutation of the same character (here at the 5% level). The point at which this occurred in the analyses of turtle data is shown as a green circle in figures 6.12, 6.13, 6.16, 6.17, 6.21 and 6.22. The second method employed here is the boildown bootstrap, which was introduced in Chapter 4 for the first time. Two versions of this method were introduced, and the results obtained were extremely similar. The boildown bootstrap results presented in this chapter are of type 1 boildown bootstraps, which run bootstrap analyses at each stage of the boildown process. The average bootstrap value for the nodes of each of these bootstrap trees is then calculated, and the point during the boildown with the highest average bootstrap is considered the point at which the process should be halted. This is simply based on the idea that while the bootstrap value is increasing the data is being improved by the character removal process of the boildown. When the bootstrap value begins to drop the signal in the data is being reduced, and so the boildown should be stopped. The type 1 boildown was used in this study, because it is implemented in the *Boildown* (see Chapter 5) program and is much quicker to run than the type 2 method, which is a very user intensive and time consuming process. The results of these methods are discussed below.

6.4.6 Boildowns and the turtle debate

In boildowns of all three datasets analysed here there is evidence that characters supporting the anapsid-turtle hypothesis are more incompatible than those supporting the diapsid hypothesis. Both the data of Rieppel and Reisz (1999) and the consensus matrix showed a trend similar to that shown in figure 6.23c in CCSR and LQP boildowns. In the case of the data of Rieppel and Reisz (1999), the trend is from initial insignificant support for the diapsid hypothesis, whereas with the consensus matrix neither hypothesis is more parsimonious. In both cases, there are trends towards increasing support for the diapsid hypothesis, suggesting that this hypothesis is the more compatible with the rest of the matrix, and that the anapsid hypothesis is supported by relatively incompatible characters. The "number" boildown showed a different trend. With the data of Rieppel and Reisz it initially wavers around the original Templeton value, but does not trend strongly towards either hypothesis. Later, when almost 120 characters have been removed, it suddenly trends towards the anapsid hypothesis, which becomes a significantly better fit. The fact that this trend occurs so late in the boildown suggests that it is probably, as described above, an idiosyncrasy of the method rather than any true support for the anapsid hypothesis. Similarly, with the consensus matrix, the "number" boildown originally wavers around the original Templeton value of zero before trending towards the diapsid hypothesis. Again, this trend may not have much meaning. All three boildowns of the osteological data of Lee (2001) showed trends similar to that in figure 6.23d, from significant support for the anapsid hypothesis towards significant support for the diapsid hypothesis. Such a trend, as stated above, is unexpected. It suggests that even with very strong support in parsimony analysis, the anapsid hypothesis is supported by characters that are highly incompatible with the rest of the data, whereas the characters supporting the diapsid hypothesis are more compatible, and are therefore removed later in the boildown, or not at all.

In many cases, the trend for preferentially removing characters supporting the anapsid hypothesis early in the boildown can be shown statistically. Table 6.12 shows the results of Kruskal-Wallis tests on the results of the boildown analysis. This test is very similar to the Mann-Whitney test described in Chapter 2, but can be applied to data in which there are more than two independent samples. In this case, the data for each boildown was ranked on the basis of the removal order of characters in the boildown. The characters were then separated into three categories, those which supported the hypothesis

of Rieppel and Reisz (1999) over that of Lee (2001), those that supported the hypothesis of Lee over Rieppel and Reisz, and those that supported neither hypothesis over the other. The Kruskal-Wallis test simply tests the null hypothesis that there is no difference in the distributions of the three categories of characters through the boildown.

	CCSR	Number	LQP
Rieppel and Reisz (1999)	0.198	0.110	**0.012**
Lee (2001)	**0.003**	**0.016**	**0.014**
Consensus Matrix	**0.026**	0.183	**0.009**

Table 6.12. Results of Kruskal-Wallis tests on the removal order of CCSR, "Number" and LQP boildowns of the three datasets under study. The method was used to test for a significant difference between the removal orders of characters supporting the anapsid and diapsid hypotheses and those supporting nether hypothesis. Significant results at the 5% level are highlighted in bold.

		CCSR	Number	LQP
Rieppel and Reisz (1999)	N vs D	0.7153 N	0.0908 D	**0.0042 N**
	N vs A	0.1275 A	0.1169 A	0.7279 N
	D vs A	0.0575 A	0.9163 D	**0.0210 A**
Lee (2001)	N vs D	0.9253 N	0.3215 D	0.1875 N
	N vs A	**0.0009 A**	**0.0042 A**	**0.0273 A**
	D vs A	**0.0117 A**	0.2533 A	**0.0065 A**
Consensus Matrix	N vs D	0.6071 N	0.5535 D	**0.0290 N**
	N vs A	**0.0161 A**	0.0598 A	0.1101 A
	D vs A	**0.0107 A**	0.4708 A	**0.0022 A**

Table 6.13. Results of Mann-Whitney tests on the removal order of CCSR, "Number" and LQP boildowns of the three datasets under study. Characters were split into those supporting the diapsid hypothesis (D), those supporting the anapsid hypothesis (A) and those supporting neither hypothesis (N). Each character type was tested against one another to see if one type was removed significantly earlier in the boildowns. P-values indicate probability of significant difference in a one-tailed test, and the letter following the p-value indicates which set of characters was removed earlier on average. Results showing significance at the 5% level are highlighted in bold.

As shown in table 6.12, significant values were obtained for all but three boildowns (CCSR and "Number" boildowns of the data of Rieppel and Reisz, and the "Number" boildown of the consensus matrix), showing that in six cases the three categories of characters are not being removed during the boildown at the same rate. Unfortunately, the Kruskal-Wallis test does not show which characters are being removed earlier and which

later. Therefore, Mann-Whitney tests (see Chapter 2) were used to test for differences in the time of removal between each pair of character categories in each boildown. The results are shown in table 6.13.

It can be seen that in many cases characters supporting the hypothesis of Lee (2001) are removed significantly earlier in the boildown than those supporting the hypothesis of Rieppel and Reisz (1999) or supporting neither hypothesis. This provides evidence that the characters supporting the hypothesis of Lee are weak. In many cases, characters supporting the hypothesis of Rieppel and Reisz (1999) are removed significantly later than characters that support neither hypothesis, providing support for this hypothesis being more compatible with the data as a whole. In the only boildown in which there appears to be a trend towards support for the anapsid hypothesis (the "number" boildown of the data of Rieppel and Reisz), characters supporting the anapsid hypothesis are still removed earlier on average, but there is no significant difference between the removal orders of the two groups ($p=0.9163$).

The results of the methods for identifying a stopping point for the boildown were confusing. Generally it was found that the boildown bootstrap method suggested the boildown should be halted very early into the procedure, whereas the LQP based method allowed the process to run much further. This means that in the case of the LQP method the trends described earlier are obvious before the suggested stopping point, whereas in many cases the boildown bootstrap halted the procedure before any obvious trend could be seen. This difference may, however, be due to the properties of the two procedures, and may suggest that they are useful for different purposes. The LQP method is searching for the point where all remaining characters have an LQP value of less than 5%, which is essentially looking for continuity between all characters, whereas the boildown bootstrap is looking for an improvement in tree resolution. Therefore, the boildown bootstrap is affected more by the number of characters remaining, because resolution is likely to be lost from a parsimony-based tree when the number of characters remaining drops too low, whereas compatibility is easier to achieve when the number of characters is smaller. In the case of the turtle data, it is not surprising that the boildown bootstrap method halts the process early. The turtle data appears to contain two signals, and the characters removed early in the boildown tend to be those supporting the anapsid positioning of turtles. However, in the case of the data of Rieppel and Reisz (1999) and Lee (2001), the tree produced from the entire data has high support, and the tree is very well resolved. This means that the bootstrap values are high at the start of the analysis. It has been shown

above that the characters removed near the start of the boildowns tended to be those supporting the anapsid hypothesis. Even if these characters are misleading regarding the position of turtles, they still often contain useful evidence about relationships elsewhere in the tree. Therefore, removal of these characters is likely to reduce the high initial bootstrap value and cause the boildown bootstrap to stop. In the case of the consensus data, the initial data does not show such high support for a single hypothesis, and so the boildown procedure increases bootstrap support much longer through the process. The utility of the boildown bootstrap may therefore be in identifying characters that are simply noise, and helping to find extra resolution in analyses that produce initially poorly resolved trees.

One problematic feature of the boildown bootstrap, and possibly all bootstrap analyses, was detected during this study. Unless a large number of bootstrap replicates are carried out in each bootstrap, the random nature of the process means that the random variation in average bootstrap is high enough to mask the true results. Unfortunately, due to time restrictions it was only possible to use 100 bootstrap replicates in the analyses above. To test how great this effect was, a small study was carried out in which 20 bootstrap analyses were carried out on the consensus matrix (including all characters). This was done using 100, 1,000 and 10,000 bootstrap replicates. The average bootstrap for each replicate was recorded to see how much variation was present between the replicates. The results, along with an approximate time for the analysis is shown in table 6.14 below.

No. Replicates	Time Taken (mins)	Minimum Av. Bootstrap	Maximum Av. Bootstrap	Difference
100	30	48.72	55.62	6.90
1000	300	51.80	53.89	2.09
10000	3000	52.51	53.20	0.69

Table 6.14. Results of a study into the amount of variation in average bootstrap between 20 repeats of the same bootstrap analysis with different numbers of bootstrap replicates.

It can be seen that a large amount of variation in average bootstrap value was found in only 20 replicates of the same bootstrap analysis, especially when 100 bootstrap replicates were used. The values of difference shown in table 6.14 represent the average difference per node on the tree, and is a worrying value. These results suggest that the number of bootstrap replicates used in a boildown bootstrap procedure, and in general, is of paramount importance. I would suggest that all bootstrap analyses should be carried out with at least 10,000 bootstrap replicates where possible.

Although none of the results presented here provide compelling evidence in support of either of the hypotheses for the affinities of turtles when taken separately, the wider

picture indicates that a recurring signal through the results of many of the analyses. Using many methods of data exploration based on both parsimony and compatibility there is evidence that support for the anapsid placement of turtles is provided by characters that are highly incompatible and homoplastic, and that the underlying signal within the data under investigation is in support of a diapsid positioning. Such strong evidence for one hypothesis in a debate that has historically been so difficult to separate must be considered important, especially given that the same signal supporting the diapsid hypothesis has been identified in data that has previously been used to provide statistically significant support for the anapsid hypothesis. The evidence either indicates that the true phylogenetic position of turtles lies somewhere within the Diapsida, possibly as the sister-group to sauropterygians, or that the method by which the datasets under study have been produced has led to flawed results. It is possible that because Lee (2001) based his analysis on characters and scoring originally produced by Rieppel and deBraga (1996) in a dataset that supported a diapsid positioning, an underlying signal remained in support of this hypothesis despite Lee's (2001) rescorings. It has already been shown that statistically more of these rescorings involved the turtles and their two prospective sister-groups (the eosauropterygians and pareiasaurs) than would be expected by chance. This could indicate that more effort was put into checking the scorings of these taxa, because they were considered more important in resolving the origins of turtles, and therefore, underlying signal supporting the diapsid-turtle hypothesis may have remained unchecked.

6.4.7 Why All the Confusion?

Long branch, or long edge attraction (LBA), as defined by Hendy and Penny (1989) is a well-known phenomenon in phylogenetic analysis, especially of nucleotide data. It is one of the biggest problems of modern phylogenetic reconstruction, and, unlike many other problems in phylogenetics, can be made worse by the addition of more characters to the data (e.g. see Hillis *et al.*, 1994) and taxa (Swofford and Poe, 1999). Felsenstein (1978) showed with a simple, four-taxon example, that two unrelated taxa can appear to be sister-groups if they both have long branch lengths relative to the rest of the branches in a phylogeny (See also Huelsenbeck and Hillis, 1993; Huelsenbeck, 1995). A simplified explanation of this result is as follows. When a branch length is very long, many state changes must have occurred on that branch, and because the number of states for each character is finite, many reversals are likely to occur, leading to the data for that taxon becoming saturated, and effectively random. If two taxa have long branches, by chance

alone they are often drawn together, because both have relatively random character state data. The same phenomenon can be explained in a different way, which may be termed short branch attraction (SBA). In this explanation, when two or more taxa have extremely long branches they become very dissimilar to their respective sister-groups. Because the lengths of branches between the other groups are small, they appear to be more closely related than they are to the taxa with long branches, leaving the long branch taxa as outcasts. This may be why taxa with long branches are often expelled from the rest of the ingroup to become placed next to the outgroup of the analysis (Pisani pers. comm. 2004).

LBA attracts much attention in analyses of nucleotide data, because of the nature of this type of data. Excluding gaps, there are only four possible states for a nucleotide character, and reversals are possible, meaning that saturation is highly probable on long branches. There are far fewer reported cases of LBA in analyses of morphological data, although the original example given by Felsenstein (1978) was not specific to nucleotide data, and there is no reason that the phenomenon should not occur using morphological data. The reason morphological data tends to be assumed less susceptible to LBA is that the number of states possible for each character could theoretically be almost infinite, meaning that saturation of states is not so much of a problem on long branches. In practice, however, most morphological characters possess only a few states that incorporate all known morphologies of the structure for which they code, and the subjective process of scoring means that taxa are assigned characters states based on the fact they are more like state 0 than state one, or vice versa. An extreme case is that of presence-absence characters which only have two possible states. In such cases, reversals and saturation are not impossible. In addition, SBA, which is effectively the same phenomenon, is not reliant on state saturation. The problems in resolving the position of turtles may be such a case of LBA in morphological data. If branch lengths are plotted onto the most parsimonious trees of each analysis (Fig. 6.24), it can be seen that the some of longest branches are within and surrounding the Gorgonopsia and Cynodontia, the Pareiasauridae, the sauropterygians and the Testudines. These are the taxa that have been found to be unstable using the data exploration methods utilised here, suggesting that LBA is present in the data, leading to a number of taxa, including the Testudines, becoming unsettled. Further evidence that LBA may be at work comes from the results of successive sister-group removal. In the analysis of Lee's (2001) data, only a single sister-group needed to be removed before the turtles switched to become part of the Diapsida. This shows a lack of support within this data for the nesting of turtles within the Anapsida, and suggests that the reason turtles plot in the

Anapsida in this analysis might be convergence with the Pareiasauridae. Such a lack of support for nesting has been used as a method for identifying groups that have been subject to LBA (Lyons-Weiler and Hoelzer, 1997), and strongly suggests that the anapsid positioning of turtles could be attributed to LBA, a result that fits with the identification of anapsid-turtle supporting characters as weak by compatibility tests. Furthermore, the branch lengths surrounding the pareiasaurs are considerably greater than those around the sauropterygians in all analyses (see Fig.6.24), providing more evidence that the anapsid link may be the result of LBA.

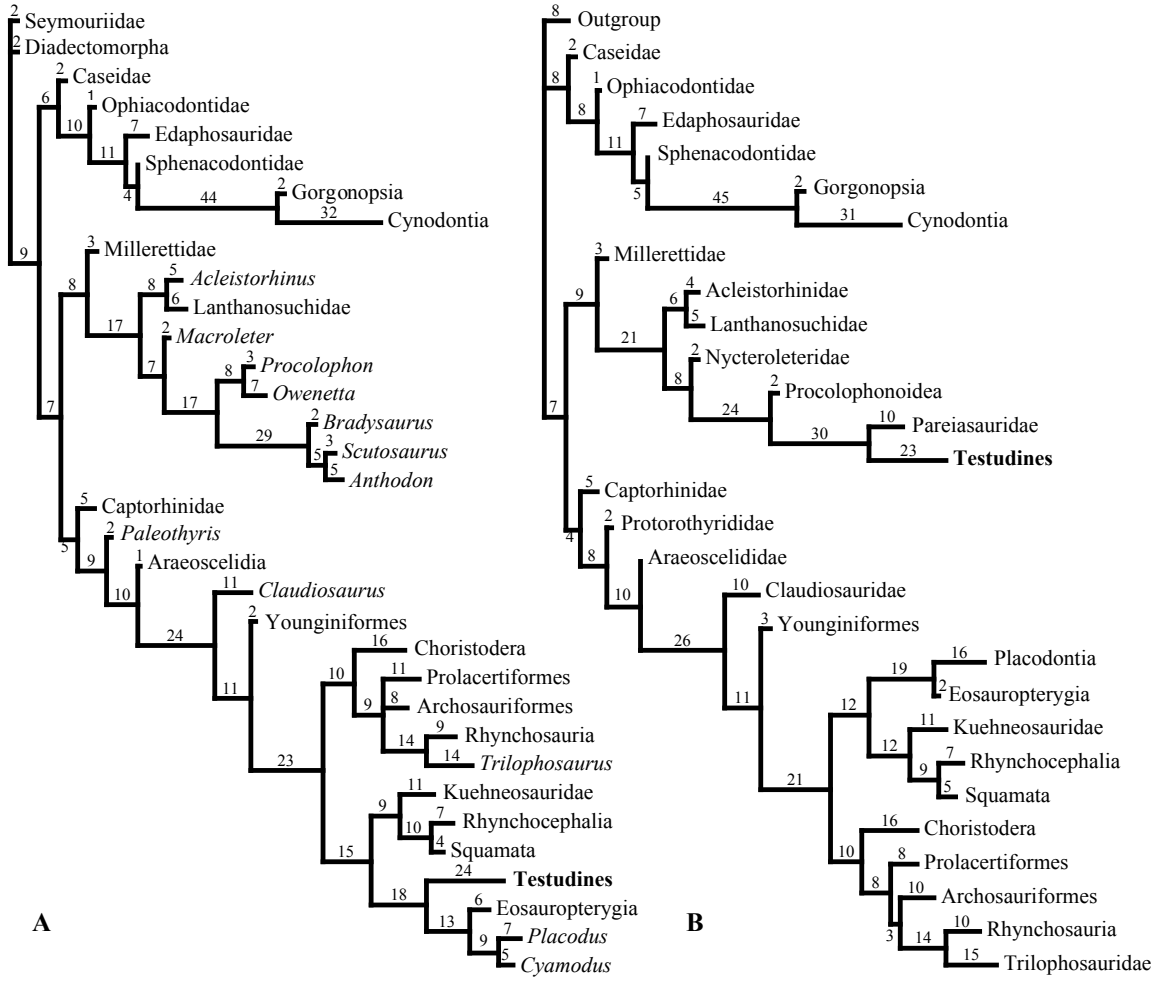


Figure 6.24. Phylograms depicting branch lengths on the strict consensus tree of a parsimony analysis of the data of (A) Rieppel and Reisz (1999) and (B) Lee (2001). Branch lengths on the MPTs of analyses of the consensus matrix are extremely similar to those shown in A and B. Numbers on branches indicate branch lengths. Turtles are highlighted in bold.

The possibility that the confusion in identifying the origins of turtles is due to LBA is worrying, because in this case the long branches are not considerably longer than other branches in the tree, yet it is enough to cause taxa to become unstable and change their

position drastically. The potential of LBA to affect the results of morphological phylogenetic analyses are rarely discussed, and most morphological workers tend to assume it is a phenomenon associated only with molecular data, and as such do not take into account the possible effects it may have. However, the results here suggest that only small differences in branch length can cause LBA in morphological analyses, so that the phenomenon may not be uncommon, and should certainly be afforded more attention by morphological phylogeneticists.

6.4.8 The Utility of Alternatives to Parsimony

The most important message that can be drawn by the results of this case study is that parsimony should not be considered the one and only step in searching for a phylogeny. There are a plethora of data exploration and support methods available to the phylogeneticist that can provide an insight into the data far beyond the information given by parsimony analysis. Even though it could be argued that the results obtained here are not compelling, it cannot be denied that questions about the strength of the characters supporting the anapsid-turtle link have been raised that simple parsimony analysis has missed. This shows that although stagnated debates may indeed only be solved by reassessing character homologies (Rieppel and Kearney, 2002) or finding new, intermediate fossils (Heckert and Lucas, 2003), there is often more information held within the currently available data that can provide evidence in favour of one conclusion (see also Harris *et al.*, 2003a). It is also seems likely that in many seemingly strong published phylogenetic analyses the true tree may be obscured by secondary signal, perhaps as a result of LBA. The use of more detailed data exploration methods in these cases could highlight potential convergence and possibly reduce the number of stagnated debates in the future. Workers should certainly not dismiss phylogeny reconstruction methods other than parsimony. Although it is not suggested here that compatibility should be used alone for producing phylogenies (its weaknesses have been highlighted in the literature e.g. see Felsenstein, 1978), the use of a number of different methods can be used a measure of support for a hypothesis. If parsimony and compatibility agree on a set of relationships, then it can be more confidently concluded that this relationship is strong. If they disagree, then further investigation may be advisable in order to understand the reasons why.

6.5 Conclusions

The most recent datasets in the debate over the origins of turtles continue to produce inconclusive results. Reanalyses of the two most recent morphological analyses of the position of turtles within the Amniota confirmed that they provide evidence for two alternative positions for the turtles. Bootstrap analyses showed high levels of instability in the Testudines, Pareiasauridae and the clade comprising the Gorgonopsia and Cynodontia, suggesting that there is a great deal of convergence between these groups. The same groups were also shown up by compatibility tests as being the most incompatible with the rest of the data, again highlighting their instability. It is concluded that the convergence associated with these groups is likely to be the result of LBA, because they are the taxa that exhibit the longest branches in the most parsimonious trees of analyses of all three datasets.

The data exploration methods provide considerable evidence that the characters supporting the anapsid-turtle hypothesis are highly incompatible with the rest of the data, suggesting that this hypothesis may be the one that is being supported by LBA, and leading to the suggestion that the diapsid hypothesis may be the stronger of the two. The evidence for this conclusion is not highly robust, but, in the light of the stagnated nature of the debate in recent years, any evidence supporting one hypothesis over the other is important, and the repeated occurrence of this same conclusion from different analysis techniques suggests that the signal is real.

This case study highlights the utility of data exploration techniques in cases where a debate between two or more phylogenetic hypotheses has reached a position of stalemate. It is also suggested that such techniques should be employed in even seemingly strong phylogenies to evaluate support and the potential of convergences in the data, including LBA.

Chapter 7: A Closer Look at Turtle Morphology

7.1 Aims

The aim of this chapter is to look more closely at the osteological characters used in the analyses of Rieppel and Reisz (1999) and Lee (2001), to try to resolve the scorings over which these authors disagreed, and to compare the results of a re-evaluation of the characters with those of the analytical methods employed in Chapter 6.

7.2 Methods

Each character employed in the analyses of Rieppel and Reisz (1999) and Lee (2001) was scrutinised in detail, and problems and errors in character constructions and scorings were identified and improved. The scoring differences between the authors that were identified in production of the consensus matrix (see Chapter 6) were resolved as far as possible. Each character was also placed into the coding categories identified in Chapter 2, and this information was used along with the results of compatibility methods from Chapter 6 to test whether certain types of character construction were causing more conflict in the data than would be expected by chance alone. An extra type of character, called shape, was also added. This type contains all characters that describe the shape of a bone or part of a bone, and therefore includes all ratio and extent characters. The addition of the shape character type was to assess whether such characters are more homoplastic than those based on seemingly more objective attributes, such as topological position or presence and absence.

7.3 Character Examination

A number of problems were identified when the character constructions and scorings were examined in detail. Below is a list of the 168 characters common to the analyses of Rieppel and Reisz (1999) and Lee (2001). Coding and scoring problems are highlighted and any changes made to the matrix explained. All scorings were made by consulting the relevant literature (Broom, 1903a; 1903b; 1912; 1913; Boonstra, 1929a; 1929b; 1932a; 1932b; 1934a; 1934b; 1934c; Haughton, 1929; Haughton and Boonstra, 1929; 1930a; 1930b; Watson, 1941; Romer, 1956; 1964; Walker, 1961; Robinson, 1962; Ewer, 1965; Colbert, 1966; 1970; Carroll, 1969a; 1975; Cruickshank, 1972; Gow, 1972; 1975; Clark and Carroll, 1973; Chatterjee, 1974; Gaffney, 1975a; 1975b; 1979b; 1990; Carroll and Galton, 1977; Kemp, 1978; 1979; Heaton and Reisz, 1980; Reisz *et al.*, 1984; Carroll and

Lindsay, 1985; Gaffney *et al.*, 1987; Fraser, 1988; Reisz and Laurin, 1991; Storrs, 1991; Rieppel, 1994; 1995; 1998; 2000a; Sues *et al.*, 1994; deBraga and Reisz, 1995; Dilkes, 1995; 1998; Dodick and Modesto, 1995; Gaffney and Kitching, 1995; Lee, 1995; Rougier *et al.*, 1995; de la Fuente *et al.*, 1996; deBraga and Reisz, 1996; Rougier *et al.*, 1996; Jalil, 1997; Wu *et al.*, 1997; Gower and Weber, 1998; Reisz *et al.*, 1998; Reynoso, 1998; Joyce, 2000; Evans *et al.*, 2001; Modesto and Smith, 2001; Reisz and Laurin, 2001; Rubidge and Sidor, 2001; Gower and Walker, 2002; Reisz and Scott, 2002; Tong *et al.*, 2002; deBraga, 2003; Sen, 2003; Senter, 2003; Modesto and Sues, 2004; Sidor and Smith, 2004). After each original character description, a list of any coding categories into which that character fits is included. For explanation of these categories see Chapter 2.

1) **Premaxilla exposure: exposure anterolateral to external nares small, restricted to low posterolateral process, forming less than one-half the height of the premaxilla (0), posterolateral process, tall reaching dorsal process (1).** SINGLE STRUCTURE RATIO, SHAPE

2) **Premaxilla/prefrontal contact: absent (0), present (1).** SHAPE

This character is uninformative, because contact is only present in rhynchosaurs.

3) **Premaxilla dentition: present (0), absent (1).** LOSS

Rescored as 0&1 for Testudines, because *Proganochelys* possesses teeth on its premaxilla (see Kordikova, 2002).

4) **Premaxilla/external nares relationship: excluded from posterior border of nares (0), contributes to posterior border (1).** SHAPE

This character is linked to character 1 (DeBraga, Rieppel and Reisz pers. comms. 2003), and has therefore been removed from the new matrix.

5) **Septomaxilla facial process: absent (0), present (1).** SHAPE

Archosauriformes, sauropterygians, *Trilophosaurus*, and Testudines lack septomaxillae, but these are scored as state 0 along with taxa with septomaxillae but no facial process. The character was split into two characters to remove this potential false homology. The new characters are:

5a) **Septomaxilla: present (0), absent (1)**

5b) **Septomaxilla facial process: absent (0), present (1)**

All taxa scored as lacking a septomaxilla in character 5a are scored as unknown (?) for 5b.

Edaphosauridae rescored as 0 for 5a and 0 for 5b.

6) **External nares exposure: dorsal process of premaxilla broad, restricting nares to a lateral exposure (0), dorsal process narrow resulting in dorsal exposure of nares (1).** POSITIONAL, SHAPE

At first sight, this character appears to be linked to character 7, because taxa with confluent external nares cannot have a broad dorsal process of the premaxilla that restricts the external nares to a lateral exposure. However, instead of rescoring these taxa as unknown, the character description was changed to:

6) **External nares: not dorsally exposed (0), dorsally exposed (1).**

Testudines rescored as 0&1, because *Proganochelys* does not show dorsal exposure of the external nares.

Choristodera rescored as 0, because external nares are dorsally exposed.

7) **External nares: separated by intranarial bar of premaxilla (0), confluent (1).** POSITIONAL, SHAPE

8) **Choana palatal exposure: parallel medial border of maxilla (0), deflected posteromedially (1), hidden in palatal view (2).** INAPPLICABLE DATA (multistate), SHAPE

Cynodontia rescored as 1 for this character (e.g. see Kemp, 1978; Carroll, 1988).

9) **Nasals: paired (0), fused (1), lost (2).** INAPPLICABLE DATA (multistate), LOSS

This character was split so that taxa with paired or fused nasals are united because of their possession of nasals. The new characters are as follows:

9a) **Nasals: present (0), absent (1)**

9b) **Nasals: paired (0), fused (1)**

All taxa scored as lacking nasals for character 9a are scored as unknown for 9b.

State two in the original character, and therefore character 9b in the new construction are uninformative, because nasals are lacked only by some squamates.

10) **Nasal/frontal ratio: nasal equal to or shorter than frontal (0), nasal at least one-third longer or better (1).** RATIO, SHAPE

11) **Maxilla ascending process: absent (0), present between orbit and external nares (1).** SHAPE

12) **Maxillary horn: absent (0), present directly behind external nares (1).** SHAPE

13) **Anterolateral maxillary foramen: absent or if present equal in size to all other foramina (0), present, at least twice the diameter of all other foramina (1).** UNIFYING, RATIO, SHAPE

14) **Maxilla length: extends to posterior orbital margin (0), does not reach posterior orbital margin (1).** EXTENT, SHAPE

15) **Maxilla orbital exposure: absent (0), present (1).** POSITIONAL

Archosauriformes are rescored as 0, because they do exhibit maxilla orbital exposure.

16) **Maxilla/quadratojugal relationship: not in contact (0), in contact (1).** POSITIONAL, SHAPE

Taxa scored as lacking a quadratojugal in character 42 are rescored as unknown for this character.

17) **Lacrimal morphology: present and contributing to external nares (0), present, at least as long as tall, but excluded from external nares (1), if present small, restricted to orbital margin or absent entirely (2).** POSITIONAL, SHAPE, INAPPLICABLE DATA (multistate), UNIFYING, LOSS
This character homologises a small lacrimal on the orbital margin with lack of a lacrimal, and that the presence or absence of the lacrimal is not being utilised as a phylogenetic character. To change this, the character was split into the following two characters:
17a) **Lacrimal: present (0), absent (1);**
17b) **Lacrimal morphology: contributing to external nares (0), at least as long as tall, but excluded from external nares (1), small, restricted to orbital margin (2).**
Cynodontia rescored as having a lacrimal that is at least as long as tall (e.g. see Kemp, 1978; Carroll, 1988).

18) **Lacrimal duct: enclosed by lacrimal only (0), lateral border formed by maxilla (1).** POSITIONAL
All taxa lacking a lacrimal are rescored as unknown for this character.

19) **Skull proportions: preorbital skull length equal to postorbital length (0), preorbital length exceeds postorbital skull length (1), postorbital length exceeds preorbital skull length (2).** RATIO, SHAPE
Cynodontia rescored as 1 for this character, because early forms have a longer preorbital than postorbital region (e.g. see Kemp, 1978; Carroll, 1988).

20) **Prefrontal/palatine antorbital contact: narrow, forming less than one-third the transverse distance between the orbits (0), contact broad, forming at least one-half the distance between the orbits (1).** RATIO, SHAPE

21) **Bulbous medial process of prefrontal: absent (0), present (1).** SHAPE

22) **Frontal orbital contribution: present (0), absent (1).** SHAPE

23) **Frontal anterior margins: frontal suture with nasal transverse (0), oblique, forming an angle of at least 30 degrees with long axis of the skull (1).** SHAPE
Any taxa scored as lacking nasals in character 9 would also be scored as unknown. However, as only some squamates lack nasals, no taxa needed to be rescored in this way. Gorgonopsia rescored as having a transverse nasal-frontal suture (Carroll, 1988).

24) **Frontal lateral lappet: absent (0), present (1).** SHAPE
This character is linked to character 22. The frontal lateral lappet is a projection of the frontal between the prefrontal and postfrontals that contacts the orbit (DeBraga and Reisz pers. comm. 2003). Therefore, all taxa with a lateral lappet must have a frontal contribution to the orbit. For this reason, all taxa scored lacking a frontal orbital contribution are rescored as unknown for this character.

Gorgonopsia rescored as polymorphic for this character (see Romer, 1956; Carroll, 1988).

25) **Frontal posterolateral processes: absent (0), present (1).** SHAPE
Kuehneosauridae rescored as lacking a posterolateral process (see Robinson, 1962).

26) **Frontal proportions: length exceeds width by at least four times (0), length no greater than twice the width (1).** SINGLE STRUCTURE RATIO, SHAPE

27) **Frontal morphology: parallelogram shaped (0), hour-glass shaped (1).** SHAPE

28) **Orbit shape: generally circular (0), anteroposteriorly elongate so that the length exceeds the height by at least 30% (1).** SHAPE, SINGLE STRUCTURE RATIO

29) **Postfrontal contribution to upper temporal fenestra: postfrontal excluded (0), postfrontal included (1).** POSITIONAL
The outgroups (Diadectomorpha and Seymouriidae) were rescored as unknown for this character, because they do not have upper temporal fenestrae.

30) **Postorbital/supratemporal relationship: in contact (0), contact absent (1), supratemporal absent (2).** INAPPLICABLE DATA (multistate), POSITIONAL
This character has been removed from the new matrix, because it is linked to characters 50 (the presence of an upper temporal fenestra) and 53 (the shape of the supratemporal). Both this character and character 53 have a state for absence of a supratemporal, in itself a poor character construction, but worse is the fact that the taxa scored as lacking supratemporals are not the same in the two characters. Kuehneosauridae, Choristodera and *Trilophosaurus* are scored as unknown for this character, but as lacking a supratemporal in character 53, Testudines, Rhynchocephalia, Squamata, Prolacertiformes and Archosauriformes are scored as possessing supratemporal in this character, but are polymorphic (present and absent) for character 53, and Rhynchosauria, Placodontia and Eosauropterygia are scored as possessing a supratemporal in this character, but lacking one in character 53. These discrepancies are dealt with in the discussion of character 53. This character is also linked to the possession of an upper temporal fenestra. In taxa with an upper temporal fenestra it is situated between the supratemporal (where present) and postorbital. These things together make this character highly problematic and support its removal from the matrix.

31) **Postorbital/parietal relationship: in contact (0), contact absent (1).** POSITIONAL

32) **Postorbital posterior extent: terminates prior to reaching posterior limit of parietal (0), extends to at least the posterior limit of the parietal (1).** EXTENT, SHAPE
Gorgonopsia rescored as 1 and Cynodontia rescored as 0&1 for this character (see Romer, 1956; Kemp, 1978; Carroll, 1988).

33) **Jugal posterior process: extends posteriorly only to the middle of the cheek (0), reaches nearly the posterior limit of the skull (1).** EXTENT, SHAPE

Edaphosauridae rescored as 0 for this character (see Romer, 1956; and compare with cynodonts in Kemp, 1978; Carroll, 1988).

34) **Zygomatic arch configuration: squamosal excluded (0), squamosal included (1).** POSITIONAL

35) **Squamosal lateral exposure: ventral process long, descends level to limit of orbital margin (0), ventral process short, terminates prior to reaching ventral orbital margin (1), ventral process absent or restricted to region above dorsal limit of orbit (2).** EXTENT, SHAPE, INAPPLICABLE DATA (multistate)

Most taxa do not have an obvious ventral process on their squamosal, so here this character definition is changed to:

35) **Squamosal lateral exposure: descends level to limit of orbital margin (0), terminates prior to reaching ventral orbital margin (1), restricted to region above dorsal limit of orbit (2)**

36) **Squamosal contribution to posttemporal fenestra: absent (0), present (1).** POSITIONAL

All taxa that are scored as lacking a posttemporal fenestra in character 59 are rescored as unknown for this character.

37) **Squamosal occipital flange: absent or poorly developed, forming only a thin ridge (0), flange well developed, forming a broadly exposed lappet (1).** SHAPE

38) **Quadrate excavation: absent along posterior edge (0), posterior edge deeply excavated forming a concave region (1), quadrate greatly reduced (2).** INAPPLICABLE DATA (multistate)

To remove the unexplained homology of a reduced quadrate with quadrate excavation in large quadrates, this character was split into the following two characters:

38a) **Quadrate: large (0), greatly reduced (1)**

38b) **Quadrate: posterior edge not excavated (0), posterior edge deeply excavated forming a concave region (1)**

Taxa with greatly reduced quadrates are scored as unknown for 38b.

39) **Quadrate exposure laterally: absent (0), present (1).** POSITIONAL

Unsurprisingly, taxa with a greatly reduced quadrate in character 38 are all scored as lacking lateral quadrate exposure. These taxa are all recoded as unknown for this character to remove this logical linkage.

40) **Quadrate lateral conch: absent (0), present (1).**

Again, taxa with a greatly reduced quadrate in character 38 are all scored as lacking a quadrate lateral conch. These taxa are all recoded as unknown for this character to remove this logical linkage.

41) **Quadrate anterior process: long, extending forward along its sutural contact with the quadrate process of the pterygoid to nearly reach the level of the transverse flange (0), short, not extending anteriorly beyond 55% the length of the quadrate process of the pterygoid (1).** EXTENT, RATIO, SHAPE

Taxa coded as having a greatly reduced quadrate in character 38 are recoded as unknown for this character.

42) **Quadratojugal morphology: present and horizontal dimension exceeds vertical dimension by a factor of at least three (0), present but vertical dimension exceeds horizontal by a factor of at least two (1), present but greatly reduced and restricted to condylar region (2), absent (3).** INAPPLICABLE DATA (multistate), SINGLE STRUCTURE RATIO, SHAPE, POSITIONAL, LOSS

This character has been split into two in the new matrix to unite taxa that possess a quadratojugal. States 0, 1 and 2 of the original character are kept as a single character, but states 1 and 2 have been reworded to account for the fact that many taxa seem to fall between the original state definitions. The new characters are as follows:

42a) **Quadratojugal: present (0), absent (1)**

42b) **Quadratojugal shape: wider than tall (0), taller than wide (1), greatly reduced and restricted to condylar region (2)**

All taxa lacking a quadratojugal are scored as unknown for 42b.

Archosauriformes are rescored as 1&2 for 42b, because some forms, such as *Proterosuchus* have a large, tall quadratojugal (see Cruickshank, 1972), while in others, such as *Euparkeria* it is reduced (see Ewer, 1965). *Acleistorhinus* and *Lanthanosuchus* are rescored as possessing a quadratojugal that is wider than tall (see deBraga and Reisz, 1996). *Trilophosaurus* is rescored as possessing a quadratojugal that is taller than wide.

43) **Quadratojugal ornamentation: absent (0), present (1).**

Taxa scored as lacking a quadratojugal in character 42 are recoded as unknown for this character.

44) **Stapedial shaft: rod-like in cross section (0), blade-like in cross section (1).** SHAPE

45) **Stapes morphology: robust with its greatest depth exceeding one-third of its total length (0), slender with the length at least four times the depth (1).** SINGLE STRUCTURE RATIO, SHAPE

46) **Stapedial dorsal process: present as ossified process (0), absent (1).** SHAPE

47) **Parietal skull table: broad with the mid-line transverse width not less than half of the length measured along the element's midline (0), constricted with the length exceeding the width by at least three times (1), forming sagittal crest (2).** SINGLE STRUCTURE RATIO, SHAPE

The presence of a sagittal crest does not seem homologous to the other two states. It seems plausible to be able to have a sagittal crest on a broad or constricted parietal skull table. For this reason, and because the presence or absence of a sagittal crest appears linked to the presence or absence of an upper temporal fenestra and a parietal shelf, this state is removed from the new matrix. Taxa previously scored as possessing a sagittal crest (Rhynchosauria, *Trilophosaurus* and Cynodontia) are all rescored as state 1 for this character.

48) **Parietal shelf for adductor musculature: absent (0), present as shallow excavations on the lateral margins of the parietal (1).** INAPPLICABLE DATA (multistate), SHAPE

This character is biologically linked to the possession of an upper or lower temporal fenestra that contacts the parietal, because although it is logically possible for taxa without these fenestra to possess a parietal shelf, in practice, it is only found on the border of these fenestrae. Therefore, this character has been removed from the new matrix.

49) **Pineal foramen position: located in the middle of the body of the parietal (0), displaced posteriorly (1), displaced anteriorly (2), absent (3).** POSITIONAL, INAPPLICABLE DATA (multistate), LOSS

To unite taxa possessing a pineal foramen, this character was split into the following two characters:

49a) **Pineal foramen: present (0), absent (1)**

49b) **Pineal foramen position: middle of parietal (0), displaced posteriorly (1), displaced anteriorly (2).**

Gorgonopsia rescored as possessing a pineal foramen that is displaced anteriorly, and Ophiacodontidae, Edaphosauridae and Spenacodontidae rescored as possessing a centrally-placed pineal foramen (see Romer, 1956; Carroll, 1969b).

50) **Upper temporal fenestra: absent (0), present (1).**

51) **Lower temporal fenestra: absent (0), present, quadratojugal included (1), present, quadratojugal excluded (2), open ventrally (3).** INAPPLICABLE DATA (multistate)

Lee (2001) noted that this character construction included three binary characters, the presence of a lower temporal fenestra, the ventral opening of the fenestra and the contribution of the quadratojugal to the fenestra. Lee (2001) split the character into two (51 and 170) and disregarded the contribution of the quadratojugal to the fenestra on the ground that it is correlated to the presence of the quadratojugal. Here the character is split into the following three separate characters:

51a) **Lower temporal fenestra: absent (0), present (1)**

51b) **Lower temporal fenestra: closed (0), open (1)**

51c) **Quadratojugal contribution to lower temporal fenestra: included (0), excluded (1)**

All taxa scored as lacking a lower temporal fenestra are scored as unknown for 51b. All taxa lacking a quadratojugal in character 42 are recoded as unknown for character 51c. This negates the linkage highlighted by Lee (2001).

52) **Postparietal: Present and paired (0), present but fused (1), absent (2).** INAPPLICABLE DATA (multistate), LOSS

This character construction does not unite taxa that possess postparietals, so the character was split into the following two characters:

52a) **Postpareital: present (0), absent (1)**

52b) **Postpareital: paired (0), fused (1)**

Taxa scored as lacking postpareitals are scored as unknown for 52b.

53) **Supratemporal: present and large, with its transverse dimension nearly equal to its parasagittal dimension (0), present but reduced so that its transverse dimension is less than half of its parasagittal dimension (1), absent (2).** INAPPLICABLE DATA (multistate), SINGLE STRUCTURE RATIO, SHAPE, LOSS

This character is split into the following two characters to make phylogenetic use of the (presumably homologous) presence of a supratemporal:

53a) **Supratemporal: Present (0), absent (1)**

53b) **Supratemporal: Large (0), small (1)**

All taxa lacking a supratemporal are scored as unknown for character 53b.

This character is linked to character 30, which has been removed from the new matrix. Both characters contained a state for absence of the supratemporal, and the taxon scorings for these two identical characters were not the same (see character 30 for more information). Here, Kuehneosauridae, Choristodera, *Trilophosaurus*, Eosauropterygia and Placodontia are scored as lacking a supratemporal (0,?), as they were in the original character 53. Rhynchosauria are rescored as possessing a supratemporal, a bone found in early forms such as *Mesosuchus* (Dilkes, 1998) and *Rhynchosaurus* (Benton, 1990), although it is absent in later Triassic forms (Benton pers. comm., 2003 and see Chatterjee, 1974). The scorings of Testudines, Rhynchocephalia, Squamata, Prolacertiformes, Archosauriformes depends upon the taxon sampling, and the inconsistencies in the original data matrices are therefore difficult to resolve. Therefore, the polymorphic scorings are retained for these taxa (but see taxon sampling section below).

54) **Intertemporal: present (0), absent (1).** LOSS

The intertemporal is only present in the outgroup, so this character is uninformative when a single outgroup is used, as in Lee (2001). The character is informative when two separate outgroups (Diadectomorpha and Seymouriidae) are included as in Rieppel and Reisz (1999).

55) **Tabular: present but restricted to dorsal region of occiput (0), present but ventrally elongate descending to the level of occipital condyle (1), absent (2).** INAPPLICABLE DATA (multistate), POSITIONAL, LOSS

To unite taxa possessing a tabular, this character was split into the following two characters:

55a) **Tabular: present (0), absent (1)**

55b) **Tabular: restricted to dorsal region of occiput (0), ventrally elongate descending to the level of the occipital condyle (1).**

Taxa scored as lacking a tabular are scored as unknown for 55b.

56) **Supraoccipital: plate-like with no sagittal crest (0), body of supraoccipital constricted at midline forming sagittal crest (1).** SHAPE

57) **Occiput configuration: broad and plate-like, forming broad sutural contact with the tabular dorsolaterally (0), open with only slight contact, if any, with tabular (1).** SHAPE

This character is logically linked to the presence or absence of a tabular (character 55). Taxa which lack a tabular are rescored as unknown for this character.

58) **Angle of occiput: oriented primarily vertically (0), tilted or sloping anteriorly at an angle of about 45 degrees (1).** SHAPE

59) **Posttemporal fenestra: absent (0), present but diameter less than half of the diameter of the foramen magnum (1), large posttemporal fenestra with a diameter at least equal to that of the foramen magnum (2).** INAPPLICABLE DATA (multistate), RATIO, SHAPE

This character was split into the following two characters so that the presumably homologous presence of a posttemporal fenestra is utilised for phylogeny reconstruction:

59a) **Posttemporal fenestra: absent (0), present (1)**

59b) **Posttemporal fenestra diameter: less than half that of the foramen magnum (0), at least equal to that of foramen magnum (1)**

Taxa lacking a posttemporal fenestra are scored as unknown for 59b.

60) **Orientation of paraoccipital process: extends laterally forming 90 degrees with parasagittal plane (0), paraoccipital process deflected posterolaterally at an angle of about 20 degrees from the transverse width of the skull (1), paroccipital process deflected dorsolaterally at a angle of nearly 45 degrees (2).** SHAPE

61) **Paroccipital process morphology: slender, with anteroposterior dimension not exceeding dorsoventral dimension (0), heavy, with anteroposterior dimension at least one-third greater than dorsoventral dimension (1).** SINGLE STRUCTURE RATIO, SHAPE

62) **Exoccipital bones: meet below foramen magnum (0) do not meet (1).** POSITIONAL, SHAPE

63) **Basioccipital / basisphenoid relationship: floor of braincase with gap between both elements (0), elements fused to floor of brain cavity (1).** POSITIONAL

64) **Basi/parasphenoid ratio: narrowest transverse width no more than 60% of the maximum length measured from basipterygoid process to posterior most limit (0), narrowest part (waist) exceeds 80% of the length (1).** SINGLE STRUCTURE RATIO, SHAPE

65) **Ventral braincase tubera: absent (0), present and restricted to basioccipital (1), present, very large and restricted to basisphenoid (2).** INAPPLICABLE DATA (multistate), POSITIONAL

66) **Opisthotic/cheek contact: not in contact (0), in contact and tightly sutured (1).** EXTENT, SHAPE

67) **Prootic/parietal contact: absent (0), present (1).** POSITIONAL, EXTENT, SHAPE

68) **Medial wall of inner ear: unossified (0), ossified (1).**

Testudines rescored as 0&1, because *Proganochelys* has a "hiatus acusticus" (Gaffney 1990: Figs. 43-45). Archosauriformes is also rescored as 0&1, because many basal Archosauriformes, such as *Proterosuchus* and *Euparkeria*, have an unossified medial wall of their inner ears (Gower pers. comm. 2003 and see Gower and Weber, 1998: 395).

69) **Occipital flange: absent (0), present (1).** SHAPE

In the analysis of DeBraga and Rieppel (1997), the presence of an occipital flange on the parietal bone was identified as a synapomorphy of the Procolophonoidea. Rieppel and Reisz (1999) considered that many more taxa possessed an occipital flange (Reisz pers. comm. 2003), and so changed many scorings in this character. However, because they did not explain the reason for changing these scorings, Lee (2001) reverted to the original scoring of DeBraga and Rieppel (1997). Here we consider there to be occipital flanges in all taxa coded as possessing them in Rieppel and Reisz (1999).

70) **Sphenethmoid: present (0), absent (1).** LOSS

71) **Pleurosphenoid: absent (0), present (1).**

72) **Palate: kinetic (0), akinetic (1).**

73) **Interpterygoid vacuity: anterior end tapers sharply (0), anterior border crescentric (1), absent (2).** INAPPLICABLE DATA (multistate), SHAPE, LOSS

To unite taxa possessing an interpterygoid vacuity, this character was split into the following two characters:

73a) **Interpterygoid vacuity: present (0), absent (1)**

73b) **Interpterygoid vacuity: anterior end tapers sharply (0), anterior border crescentric (1).**
All taxa without an interpterygoid vacuity are scored as unknown for 73b.

74) **Suborbital fenestra: absent (0), present but with contribution from either maxilla or jugal along lateral border (1), present but with both maxilla and jugal excluded from lateral border (2).**
INAPPLICABLE DATA (multistate), POSITIONAL
Lee (2001) considered the small suborbital foramen in millerettids, captorhinids, nycteroleterids and procolophonids as non-homologous to the larger fenestra in some other taxa. For this reason he created two separate characters coding presence and absence of suborbital fenestrae (his character 74) and suborbital foramina (his character 170). These two characters are obviously logically linked, because it is not possible for a taxon to possess both a suborbital foramen and fenestra. It is considered here that, in the absence of contradictory evidence, all suborbital fenestrae and foramina are homologous. The original character is split into the following two characters to unite taxa that possess suborbital fenestrae:
74a) **Suborbital fenestra: absent (0), present (1)**
74b) **Suborbital fenestra: with contribution from either maxilla or jugal along lateral border (0), with both maxilla and jugal excluded from lateral border (1)**

75) **Cultriform process: long, exceeding length of parasphenoid body and reaching forward to the level of the posterior limit of the internal nares (0), short, not reaching the level of the internal nares (1).**
EXTENT, SHAPE

76) **Palatal process of pterygoid: extends anterior to the anterior limit of the palatine (0), forms oblique suture with palatine but process ends before reaching anterior limit of the palatine (1) forms transverse suture with palatine (2).** INAPPLICABLE DATA (multistate), EXTENT, SHAPE

77) **Orientation of transverse flange of pterygoid: directed predominantly laterally (0), oriented in an anterolateral direction (1).** SHAPE

78) **Dentition on transverse flange: present as a shagreen of teeth (0), present but with one large distinct row of teeth along the posterior edge of the transverse flange (1), edentulous (2).**
INAPPLICABLE DATA (multistate), LOSS
In order to distinguish taxa possessing teeth on their transverse flange, this character was split into the following two characters:
78a) **Dentition on transverse flange: present (0), absent (1)**
78b) **Dentition on transverse flange: shagreen of teeth (0), single row along posterior edge of transverse flange (1)**
All taxa scored as lacking teeth on their transverse flange are scored as unknown for 78b.
Choristodera are rescored as possessing a shagreen of teeth on their transverse flange.

79) **Ventral extent of transverse flange: extends below maxillary tooth row (0), terminates at level of or above maxillary tooth (1).** EXTENT, SHAPE

80) **Transverse flange lateral margin: posterolateral margin forms sharp edge with anteromedial margin (0), posterolateral margin merges smoothly into anteromedial margin forming a smoothly convex lateral outline (1).** SHAPE

This character is linked to the orientation of the transverse flange of the pterygoid (character 77). In their original description of this character, DeBraga and Reisz (1997, page 302) state that:

"Primitively the transverse flange is directed laterally at an angle of nearly 90º and forms an acute angle where it turns sharply anteriorly (0). In some derived states the transverse flange (char. # 77 this analysis) is directed anterolaterally (Reisz & Laurin, 1991) but the angle formed between the lateral and forward directed portions of the transverse flange remain sharp. In the derived state the transverse flange is directed anteriorly at an angle of less than 45º to the parasagittal axis and the lateral and forward portions of the transverse flange merge smoothly forming a curved anterolateral margin (1)."

Therefore, taxa scored as having a transverse flange that is directed laterally in character 77 must all be state 0 for this character. For this reason, all taxa scored as having a laterally oriented transverse flange are rescored as unknown for this character.

81) **Ectopterygoid: present and edentulous (0), present and dentigerous (1), absent replaced by medial process of jugal (2), absent replaced by lateral process of pterygoid (3).** INAPPLICABLE DATA (multistate), LOSS

This confusing character combines three attributes of the ectopterygoid, its presence or absence, the presence of teeth upon it if it is present, and the structure it is replaced by if it is absent. For this reason this character was split into the following three characters in the new matrix:

81a) **Ectopterygoid: present (0), absent (1)**

81b) **Ectopterygoid: edentulous (0), dentigerous (1).**

81c) **Ectopterygoid replaced by: medial process of jugal (0), lateral process of pterygoid (1)**

All taxa lacking an ectopterygoid are scored as unknown for character 81b, and all taxa possessing and ectopterygoid are scored as unknown for character 81c.

82) **Mandibular joint: even with occiput (0), behind occiput (1), anterior to occiput (2).** POSITIONAL

83) **Coronoid process: absent (0), present, formed by coronoid (1), present, formed by dentary (2).**

This character is rightly unordered, because the coronoid processes formed by the coronoid and dentary may not be homologous. The outgroups and Pareiasauridae are rescored as possessing a coronoid process formed by the coronoid.

84) **Coronoid number: more than one (0), only one (1).** LOSS

The state descriptions of this character are altered to take into account the fact that some taxa lack coronoids altogether. The new character definition is:

84) **Coronoid number: more than one (0), one or zero (1).**

85) **Meckelian fossa: faces mediodorsally (0), faces dorsally due to greatly expanded prearticular (1).** SHAPE

86) **Surangular length: extends anterior to coronoid eminence (0), terminates prior to reaching level of coronoid eminence (1).** EXTENT, SHAPE

Taxa that lack a coronoid eminence in character 83 are rescored as unknown for this character.

87) **Surangular lateral shelf: absent (0), present (1).** SHAPE

88) **Splenial: enters mandibular symphysis (0), present but excluded from mandibular symphysis (1) absent (2).** INAPPLICABLE DATA (multistate), POSITIONAL, LOSS

To unite taxa possessing a splenial, this character was split into the following two characters:

88a) **Splenial: present (0), absent (1)**

88b) **Splenial: enters mandibular symphysis (0), excluded from mandibular symphysis (1)**

Taxa lacking a splenial are scored as unknown for 88b.

Acleistorhinus is rescored as possessing a splenial, although it is impossible to tell if it enters the mandibular symphysis (deBraga and Reisz, 1996).

89) **Angular lateral exposure: exposed along one-third of the lateral face of the mandible (0), exposed only as a small sliver along the lateral face (1), absent from lateral aspect (2).** POSITIONAL

90) **Ventral edge of angular: smooth no ventral projection (0), keeled (reflected laminar) (1).** SHAPE

91) **Prearticular: extends anterior to coronoid eminence (0), terminates prior to reaching coronoid eminence (1).** EXTENT, SHAPE

Taxa that lack a coronoid eminence in character 83 are rescored as unknown for this character.

92) **Retroarticular process: absent (0), present (1).**

Captorhinidae rescored as possessing a retroarticular process (see Romer, 1956; Carroll, 1969a).

93) **Labyrinthine infolding: present (0), absent (1).**

This dentition character is rescored as unknown in taxa lacking teeth.

94) **Tooth implantations: set in deep sockets (0), loosely attached to medial surface of jaw (1), ankylosed to jaw (2).**

This character is rescored as unknown in taxa lacking teeth.

95) Caniniform teeth: present (0), absent (1).

This character is rescored as unknown in taxa lacking teeth.

96) **Single canine tooth: absent (0), present (1).**
This character is rescored as unknown in taxa lacking teeth.

97) **Presacral vertebral number: more than twenty (0), twenty or less (1).**
Eosauropterygia rescored as having more than 20 presacral vertebrae (e.g. see Storrs, 1991; Rieppel, 1994; 1998).

98) **Number of caudal vertebrae: twenty or more usually twenty-five (0), less than twenty (1).**

99) **Vertebral centra: notochordal (0), non-notochordal (1).** SHAPE

100) **Vertebral central articulations: amphicoelous (0), platycoelous (1), other (2).** UNSPECIFIED HOMOLOGUE, SHAPE
This character is linked to character 99. All notochordal centra must be amphicoelous. This character is also problematic because the distinction between amphicoelous and platycoelous centra is highly subjective, and all other types of centra are lumped together in a single state. For these reasons this character has been removed from the new matrix.

101) **Accessory vertebral articulations: absent (0), present (1).**

102) **Atlantal ribs: ossified (0), not ossified (1).**

103) **Cervical centra: ventrally smooth or rounded (0), ventrally keeled (1).**

104) **Cervical intercentra: present (0), absent (1).** LOSS
Biologically this character seems to be linked to character 107 (presence of dorsal intercentra). Although there is not not 100% equivalence in the taxonomic distributions of these two characters, there also appears to be high levels of correspondence, due to intraorganismal homology (see Chapter 3). For this reason, characters 104 and 107 are replaced with a single character describing the presence or absence of presacral intercentra (see character 107).

105) **Cervical ribs: without anterior process (0), anterior process present (1).**

106) **Trunk neural arches: swollen with heavy zygapophyseal buttress (0), narrow, strongly excavated neural arch with no heavy buttress (1), swollen but with narrow tall zygapophyseal buttress (2).** INAPPLICABLE DATA (multistate), SHAPE

107) **Dorsal intercentra: present (0), absent (1).** LOSS

This character has been modified to:

107) **Presacral intercentra: present (0), absent (1)**

See character 104 for discussion

108) **Dorsal transverse processes: short no more than the total transverse width of the neural arch (0), long, exceeding the transverse width of the neural arch (1).** RATIO, SHAPE

109) **Number of sacral vertebrae: two (0), three or more (1).**

110) **Caudal lateral projections (transverse processes): absent beyond fifth caudal (0), present beyond fifth caudal (1).**

111) **Caudal rib shape: L-shaped, curved (0), straight (1).** SHAPE

112) **Chevron position: intercentral (0), located on anterior pedicel (1).** POSITIONAL

113) **Cleithrum: present (0), absent (1).** LOSS

114) **Clavicle: interclavicular process of clavicle broad and plate-like with the maximum anteroposterior length at least one-third of its transverse dimension (0), slender with its anteroposterior length less than one-fifth of the transverse dimension (1).** SINGLE STRUCTURE RATIO, SHAPE

115) **Interclavicle: anterior end rhomboidal (0), T-shaped but with broad transverse bar with its anteroposterior dimension at least one-fourth the transverse width of the bar (1), T-shaped but transverse bar slender with its anteroposterior dimension much less than one fourth the transverse width (2).** SHAPE INAPPLICABLE DATA (multistate)

116) **Mineralized sternum: absent (0), present (1).**

117) **Scapula: short and broad with its height not exceeding its width (measured at the level of the glenoid) by more than three times (0), tall and blade-like with its height exceeding the width by at least a factor of four (1), tall and slender nearly cylindrical in cross-section (2).** INAPPLICABLE DATA (multistate), SINGLE STRUCTURE RATIO, SHAPE

118) **Acromion process: absent (0), present, blade-like, parallelogram in lateral aspect and arising from the lateral edge of the scapular (1), present, triangular in lateral aspect and arising from ventromedial border of scapular (2).** INAPPLICABLE DATA (multistate), SHAPE

The homology of the 'acromion processes' in turtles and pareiasaurs has been the topic of a great deal of debate in recent years. It is not uncommon for reptiles to have small bumps or processes towards the base

of the anterior surface of the scapula for articulation with the clavicle (see Fig. 7.1). Such processes have come to be collectively known as acromion processes after the similar, but non-homologous structure in mammals and their ancestors (e.g. *Ophiacodon*, see Romer, 1956). The recent debates have all centred on the homology of these small acromion processes with the more elongate, strut-like structure that protrudes from the scapula-coracoid complex of turtles (see Fig. 7.1 C and D). The debate hinges on the identity of the bones in the shoulder girdle of turtles. Unlike other anapsids, turtles possess only a single coracoid ossification. This could have occurred by the anterior coracoid fusing with either the scapula or the posterior coracoid. Knowing which is the case is instrumental in identifying the homology of the turtle acromion process. The historical interpretation is that the acromion is a process on the scapula (e.g. see Romer, 1956).

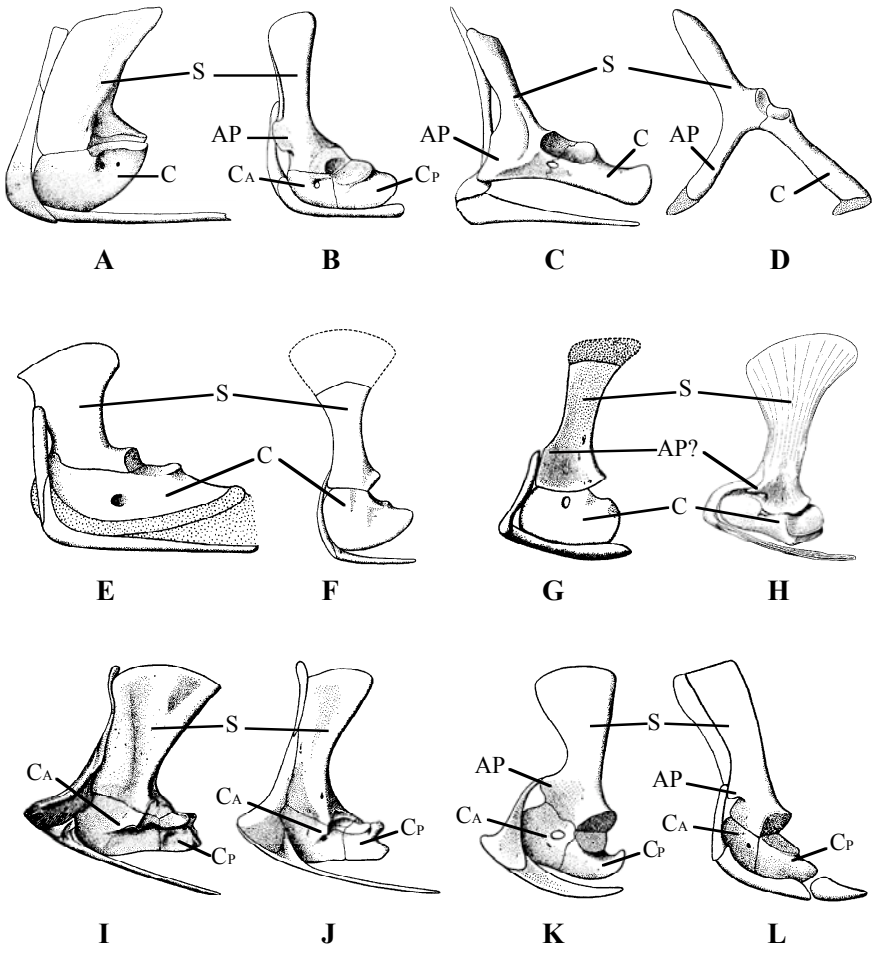


Figure 7.1. Pectoral girdles of amniotes. A & B = anapsids, C & D = Testudines, E – H = diapsids and I – L = synapsids. (A) *Seymouria* (B) *Bradysaurus* (C) *Proganochelys* (D) *Chelone* (E) The extant rhynchocephalian, *Sphenodon* (F) The rhynchosaur, *Howesia* (G) *Euparkaria* (H) *Stagonolepis* (I) *Ophiacodon* (J) The sphenacodontid, *Dimetrodon* (K & L) The cynodonts, *Leavachia* and *Kannemeyeria*. A – G & I – L from Romer (1956). H from Walker (1961). AP = 'acromion process', C = coracoid, CA = anterior coracoid, CP = posterior coracoid, S = scapula.

However, Gaffney (1990) proposed that it was actually homologous with the anterior coracoid of primitive amniotes, which had fused with the scapula. Topologically this hypothesis makes sense,

because the acromion of turtles projects anteroventrally from the glenoid and articulates with the entoplastron, which is homologous to the interclavicle. This position is homologous to that of the anterior coracoid in primitive amniotes. Further support comes from developmental studies that have shown that the acromion ossifies later, and from a separate centre than the rest of the scapula (Rieppel, 1993), at the same time as the posterior coracoid. Lee (1996b) argued that these similarities were simply convergence. He came to this conclusion because in the phylogenetic studies of Lee (1995; 1997a) and Laurin and Reisz (1995) "presence of an acromion process diagnoses a pareiasaur-turtle clade" (Lee, 1996b: 114), so he therefore inferred that the loss of the anterior coracoid was a subsequent modification in turtles alone. He continues "Pareiasaurs are the critical taxon that possess an acromion but retain an anterior coracoid, thus demonstrating that these changes were not concurrent …because an acromion and anterior coracoid co-occur in the same taxon (pareiasaurs) they cannot be homologous." (Lee, 1996b: 114). However, Rieppel (1996) pointed out the element of circularity in Lee's argument. The presence or absence of the acromion, and therefore a conjecture about the homology of the structure, was being used as a character in the phylogenetic analyses from the results of which Lee was drawing conclusions about the homology of the acromion. Rieppel (1996) showed that when the alternative hypothesis, of the acromion of turtles being the homologue of the anterior coracoid, was coded in the dataset used by Lee (1995) the same tree and tree length was yielded as was found by Lee (1995). So, both hypotheses of homology of the acromion were equally parsimonious even when turtles were found to be the sister-group to pareiasaurs. Lee's (1996b) citation of the test of conjunction (Patterson, 1982) to support his hypothesis is also based on circularity (Rieppel, 1996). By saying that because the acromion and anterior coracoid co-exist in pareiasaurs he is making the a priori assumption that the acromion of turtles is homologous to the acromion of pareiasaurs, exactly the question he is trying to answer. Furthermore, as Lee's (1995) analysis included only parareptiles and a single 'diapsid' taxon, turtles were effectively being forced into the anapsid clade, and hence the tree supported by Lee's (1995) may not represent the true position of turtles. Claims about homology based on such an analysis are clearly flawed, and the homology of various pectoral girdle elements may have to await a more robust phylogeny.

The scoring of this character in the data matrices of Rieppel and Reisz (1999) and Lee (2001) is confusing. Turtles and cynodonts are scored as possessing homologous acromion processes, while pareiasaurs are scored as possessing a non-homologous acromion. Such a coding is very strange given the debate described above. On no occasion has there been a suggestion that the turtle acromion is homologous to that of synapsids, making this choice of homology very strange. Also, given his stance on the acromion debate, it seems odd that Lee (2001) retained the scoring of Rieppel and Reisz (1999), which effectively discounts homology of the 'acromions' of pareiasaurs and turtles. Without a resolution to the correct homology of the acromion processes of various groups, this character is impossible to score confidently. For this reason this character has been removed from the new matrix.

119) **Supraglenoid buttress: present (0), absent (1).**

120) **Coracoid ossifications: one (0), two (1).**

The scoring of this character is linked to the identity of the turtle acromion process (see discussion in character 118). If the acromion process of turtles is homologous to the anterior coracoid, this character should be scored 1 for turtles. If the turtle acromion process is homologous to a scapula projection, then this character should be scored 0. However, the debate over the homology of the acromion process is yet to be resolved, and so turtles are scored as unknown in the new matrix.

121) **Coracoid foramen: enclosed by coracoid only (0), enclosed by coracoid and scapular (1).** POSITIONAL

The scoring of this character is linked to the identity of the turtle acromion process (see discussion in character 118). If the acromion process of turtles is homologous to the anterior coracoid, the coracoid foramen would be between the two coracoids and would presumably be scored 0 for turtles. If the turtle acromion process is homologous to a scapula projection, then this character should be scored 1. Without knowing the identity of the turtle acromion it is not possible to score this character, so turtles are scored as unknown in the new matrix.

122) **Humeral epicondyles: large, forming distinct processes (0), reduced so that distal end of humerus appears only marginally broader than shaft (1).** RATIO, SHAPE

123) **Humeral torsion: proximal and distal ends of humerus set off at 45 degree angle (0), angle between opposing ends reduced to no more than 20 degrees (1).** SHAPE

124) **Humeral shaft/distal end ratio: shaft length less than one-third the maximum width of the distal end of the humerus (0), shaft long at least four times the width of the distal end (1).** RATIO, SHAPE

The original discussion of this character (deBraga and Rieppel, 1997: 290) explains that state 0 of this character should be "shaft length less than or equal to three times the maximum width of the distal end of the humerus". This character is linked to the size of the humeral epicondyles (character 122) and to character 162, which describes the general shape of the limbs (short and stout versus long and thin). The original discussion of character 162 (deBraga and Rieppel, 1997: 297-298) defines short and stout limbs as "the width of the distal ends of the humerus are greater than one-third the total length of the entire bone". This is obviously logically linked to the current character even though the distribution of states among taxa is not identical. For this reason, character 124 has been removed from the new matrix.

125) **Humeral distal articulations: distinct trochlea and capitellum (0), low double condyle (1).** SHAPE

126) **Supinator process: large angled away from humeral shaft (0), large confluent with shaft (1), small or absent (2).** INAPPLICABLE DATA (multistate), SHAPE

This character is split into the following two characters to unite taxa with a large supinator process:

126a) **Supinator process: large (0), small or absent (1)**

126b) **Supinator process: angled away from humeral shaft (0), confluent with shaft (1).**

Taxa scored as having a small supinator process, or lacking the process altogether are scored as unknown for 126b. In the case of taxa with a small process, 126b is scored as unknown because the angle is difficult to identify.

127) **Ectepicondylar groove/foramen: foramen absent, but a deep groove present along anterior edge of humerus (0), foramen and groove absent, but a small notch present anterodistally (1), completely enclosed foramen present, but no deep groove (2).** INAPPLICABLE DATA (multistate)

128) **Entepicondylar foramen: present (0), absent (1).** LOSS

129) **Radius/ulna ratio: radius shorter than ulna (0), radius longer than ulna (1), radius and ulna sub-equal (2).** RATIO, SHAPE

130) **Olecranon: large and set off from proximal end of ulna (0), small or entirely absent (1).** UNIFYING

131) **Perforating foramen of manus: present (0), absent (1).** LOSS

132) **Metacarpal IV/III ratio: fourth longer than third (0), fourth equal to or shorter than third (1).** RATIO, SHAPE

133) **Thyroid fenestra: absent (0), present (1).**

134) **Posterior process of iliac blade: long, extending posteriorly well past level of acetabulum (0), posterior process reduced, distal end of ilium fan-shaped (1).** EXTENT, SHAPE

135) **Anterior process of iliac blade: blade not expanded anteriorly with at most only a very small anterior process (0), anterior process large often exceeding dimension of posterior process (1).** RATIO, SHAPE

136) **Pubic tubercle: if present small and directed anteroventrally (0), large and strongly turned ventrally (1).** UNIFYING, SHAPE

137) **Acetabulum: oval (0), circular (1).** SHAPE

138) **Acetabular process: weakly developed (0), large, overhangs femoral head, appears as triangular lateral extension when viewed from below (1).**

139) **Femoral shaft: short and stout (0), sigmoidally curved and slender (1).** UNIFYING, SHAPE

This character is logically linked to character 162, which defines limbs as short and stout or long and thin. However, the sigmoidal curvature of some taxa does constitute a justifiable character, and so this character definition is changed to:

139) **Femoral shaft: Straight (0), sigmoidally curved (1).**

All taxa were rescored for this new character construction (see matrix for new scorings).

140) **Femoral fourth trochanter: present (0), absent (1).** SHAPE

141) **Femoral trochanter major: absent (0), present and deflected distally from the proximal head of the femur (1), pyramidal in shape and nearly in line with the head of the femur (2), similar in shape to state 1, but positioned at mid-shaft length (3).** INAPPLICABLE DATA (multistate), SHAPE, POSITIONAL

This character is slightly problematic because of the uncertainty over the homology of the trochanter major between groups (e.g. see Romer, 1956: 364). It has, however been retained in the new matrix.

142) **Intertrochanteric fossa: well defined (0), reduced (1), absent (2).** INAPPLICABLE DATA (multistate), SHAPE

143) **Distal femoral condyles: large, projecting from distal end of shaft (0), reduced, not projecting beyond distal end of femur (1).**

144) **Anterior femoral condyle: larger, extends distal to posterior condyle (0), anterior condyle reduced and sub-equal or smaller than posterior condyle (1).** RATIO, SHAPE

145) **Fibula: bowed away from tibia (0), straight not bowed away (1).** SHAPE

146) **Perforating artery of pes: located between astragalus and calcaneum (0), located between distal ends of tibia and fibula (1).** POSITIONAL

147) **Tibia/astragalus articulation: loose fitting (0), tightly fitting with well developed articulations (1).**

Taxa lacking an astragalus in character 148 are rescored as unknown for this character.

148) **Discrete astragalus: absent (0), present (1).**

149) **Astragalus/calcaneum relationship in adult: never fused (0), fused (1), hinge present (2).**

Taxa lacking an astragalus in character 148 are rescored as unknown for this character. In some cases a separate intermedium is present between the ‘astragalus’ and calcaneum. In these cases the relationship between the intermedium and calcaneum is scored.

150) **Astragalus/ distal tarsal IV articulation: articulation poorly defined (0), articulation well defined (1), articulation absent (2).** INAPPLICABLE DATA (multistate)
Taxa lacking an astragalus in character 148 are rescored as unknown for this character.

151) **Calcaneal tuber: absent (0), present (1).**

152) **Distal tarsal I: present (0), absent (1).** LOSS

153) **Distal tarsal V: present (0), absent (1).** LOSS

154) **Metatarsal V: long and slender with length exceeding the width of the base by at least three times (0), short and broad with base width equivalent to at least two times the length of the element measured along its midline (1).** SINGLE STRUCTURE RATIO, SHAPE

155) **Metatarsal V shape: straight (0), hooked (1).** SHAPE

156) **Metatarsal V plantar tubercle: absent (0), present (1).** SHAPE

157) **Metatarsal I/IV ratio: metatarsal I greater than 50% the length of metatarsal IV (0), metatarsal I less than 50% the length of metatarsal IV (1).** RATIO, SHAPE

158) **Number of pedal centrali: both lateral and medial centrali present (0), medial centrali lost (1), both centrali lost (2).**

159) **Fifth pedal digit: longer than first digit (0), shorter and more lightly built than first (1).** RATIO, SHAPE

160) **Metapodials: not overlapping proximally (0), overlapping (1).** SHAPE

161) **Pedal phalangeal formula: 2,3,4,5(4),4 (0), 2,3,4,4,3 (1), 2,3,3,4,3 or less (2).** LOSS

162) **Limbs: short and stout (0), long and slender (1).** SHAPE

163) **Manus and pes: short and broad (0), long and slender (1).** SHAPE, UNIFYING

164) **Ungual size: unguals shorter than phalanges (0), unguals at least 50% longer than penultimate phalanges (1).** RATIO, SHAPE

165) **Body osteoderms: absent (0), present but few restricted to mid-line (1), present but spread all over back (2).** INAPPLICABLE DATA (multistate), POSITIONAL

This character is split into the following two characters to unite taxa possessing osteoderms:

165a) **Osteoderms:absent (0), present (1)**

165b) **Osteoderm position: restricted to midline (0), all over back (1)**

Taxa lacking osteoderms are scored as unknown for 165b.

166) **Osteodermal ridges: absent (0), fine regular spaced ridges (1), heavy irregular ridges (2).**

INAPPLICABLE DATA (multistate)

This character is not split, because only a single taxon possesses each of the two derived states, so that the character is uninformative. Taxa lacking osteoderms are scored as unknown for this character.

167) **Osteodermal limb studs: absent (0), present as conical studs (1).**

Taxa lacking osteoderms are scored as unknown for this character.

168) **Gastralia: present (0), lost (1).** LOSS

7.3.1 Debated Scorings

The new matrix is shown in Appendix 3. Table 7.1 shows the resolutions used in the new matrix for scorings over which Rieppel and Reisz (1999) and Lee (2001) disagree. The eight osteological characters added to the matrix of Rieppel and Reisz (1999) by Lee (2001) are not included in the new matrix, because they have not been scrutinised by Rieppel and Reisz, and many are impossible to score from literature, so that Lee's scorings cannot be checked here. Of these eight characters, two (179 and 170) are an attempt to improve some poor character constructions in existing characters, and these issues are dealt with above, and five of the remaining six unite turtles and pareiasaurs (and other parareptiles) to the exclusion of the sauropterygians and many other diapsids. The final character is scored as unknown for turtles, but is a synapomorphy of some of the more derived anapsids.

Character	Taxon	Rieppel & Reisz (1999)	Lee (2001)	Resolution	Reference Used
6	Eosauropterygia	1	0	1	(Rieppel, 1994)
19	Lanthanosuchidae	1	0	1	(deBraga and Reisz, 1996)
29	Keuhneosauridae	0	?	0	(Robinson, 1962; Colbert, 1966; 1970)
31	Eosauropterygia	0&1	1	*	
32	Acleistorhinidae	1	0	0	(deBraga and Reisz, 1996)
32	Lanthanosuchidae	1	0	1	(deBraga and Reisz, 1996)
41	Acleistorhinidae	1	?	?	
41	Placodontia	1	?	1	(Romer, 1956; Rieppel, 2000a)
41	Eosauropterygia	1	?	1	(Rieppel, 1994)
46	Pareiasauridae	?	1	1	(Lee, 1995)
59	Eosauropterygia	1&2	2	*	
63	Ophiacodontidae	?	0	?	
63	Millerettidae	?	0	?	
63	Nycteroleteridae	?	0	?	
63	Younginiformes	?	0	?	
63	Rhynchocephalia	?	1	?	
63	Edaphosauridae	1	0	?	
63	Sphenacodontidae	1	0	?	
63	Squamata	0	1	?	
64	Pareiasauridae	0&1	1	1	(Boonstra, 1934a)
64	Placodontia	?	1	1	(Romer, 1956; Rieppel, 2000a)
65	Pareiasauridae	2	1&2	2	(Boonstra, 1934a)
69	Outgroups	1	0	1	See discussion of character 69 in Chapter 7
69	Caseidae	1	0	1	
69	Ophiacodontidae	1	0	1	
69	Edaphosauridae	1	0	1	
69	Sphenacodontidae	1	0	1	
69	Gorganopsia	1	0	1	
69	Cynodontia	1	0	1	
69	Captorhinidae	1	0	1	
69	Protorothyrididae	1	0	1	
69	Millerettidae	1	0	1	
69	Acleistorhinidae	1	0	1	
69	Lanthanosuchidae	1	0	1	
69	Nycteroleteridae	1	0	1	
69	Araeoscelididae	1	0	1	
69	Claudiosauridae	1	0	1	
69	Younginiformes	1	0	1	
72	Acleistorhinidae	1	0	?	
72	Lanthanosuchidae	1	0	?	
73	Testudines	0&1	0	*	
74	Captorhinidae	1	0	1	
74	Millerettidae	1	0	1	(Gow, 1972)
74	Procolophonoidea	1	0	1	(Carroll and Lindsay, 1985)
74	Nycteroleteridae	?	0	?	
77	Testudines	0&1	1	*	
77	Placodontia	1	0	?	(Romer, 1956; Rieppel, 2000a)
80	Pareiasauridae	0	1	1	(Boonstra, 1934a; Romer, 1956)
82	Testudines	0&2	2	*	
83	Testudines	1	0	1	(Gaffney, 1990)
83	Placodontia	1&2	1	2	(Romer, 1956; Rieppel, 2000a)
87	Pareiasauridae	0	1	1	(Haughton and Boonstra, 1930a)
87	Eosauropterygia	1	0	1	(Rieppel, 1994)
89	Placodontia	0	1	0	(Romer, 1956; Rieppel, 2000a)

97	Lanthanosuchidae	?	0	0	(Lee, 1995)
103	Pareiasauridae	0	1	1	(Boonstra, 1934c)
120	Cynodontia	0	1	1	(Kemp, 1979)
120	Testudines	0	?	?	
121	Testudines	1	?	?	
124	Testudines	1	0	0	(Gaffney, 1990)
125	Millerettidae	0	?	1	(Gow, 1972)
125	Pareiasauridae	0&1	?	0&1	(Boonstra, 1932b)
126	Testudines	1	2	2	(Gaffney, 1990)
127	Testudines	0&2	2	*	
128	Lanthanosuchidae	?	0	0	(Lee, 1995)
139	Protorothyrididae	1	0	0	(Carroll, 1964)
139	Testudines	1	0	1	(Gaffney, 1990)
140	Pareiasauridae	0	1	1	(Haughton and Boonstra, 1930b; Boonstra, 1932a)
140	Placodontia	1	0	1	(Romer, 1956; Rieppel, 2000a)
140	Eosauropterygia	1	0&1	*	
142	Testudines	1	0	0	(Gaffney, 1990)
142	Eosauropterygia	1&2	0&1	0&1	(Rieppel, 1994)
144	Pareiasaurs	0&?	1	1	(Haughton and Boonstra, 1930b; Boonstra, 1932a)
150	Pareiasaurs	1&?	?	1	(Broom, 1913; Boonstra, 1929b)
152	Testudines	0&1	0	*	
164	Testudines	0&1	0	*	
167	Testudines	0&1	1	*	

Table 7.1. Scoring differences found between the data matrices of Rieppel and Reisz (1999) and Lee (2001) that were replaced with missing data in the consensus matrix in chapter 6. The final two columns show the resolutions of the scorings used in the rescored matrix along with the references used to resolve the conflict. ? = resolution not possible from the literature, X = character excluded from matrix, * = resolution not possible without knowledge of the exemplars used to represent the taxon.

7.3.2 Taxon Sampling

During phylogenetic analyses of large groups, such as the amniotes, it is not possible to include every known species as a separate taxon. Instead, strongly supported groups of species must be combined into more inclusive taxa (supraspecific taxa). There are two main methods by which this can be achieved, as outlined by Yeates (1995). First, a hypothetical groundplan for the supraspecific taxon can be produced. This can be a subjective process, which depends on having prior knowledge of at least part of the tree, and mistakes or poor judgement in deciding the plesiomorphic condition for the taxon can lead to inaccuracies in the tree produced. The second method is the use of a number of exemplars of the taxon. In this case, a number of preferably basal members of the clade are scored as separate taxa for the analysis. This second method should produce more repeatable results, and, as long as the exemplars chosen are relatively basal members of the clade, should give accurate results (Yeates, 1995). Yeates (1995) recommended replacing

each supraspecific taxon with three exemplars for best results. Bininda-Emonds *et al.* (1998) introduced a third method, which they called the 'democratic method'. This method involves scoring the supraspecific taxon by taking the most common state in a number of exemplars of the group.

Using simple hypothetical examples and complex real data Bininda-Emonds *et al.* (1998) found that including supraspecific taxa is fraught with dangers, and that including more than one supraspecific taxon increased the likelihood of erroneous results. They found the groundplan method to be the most accurate method when reconstructing phylogeny. However, the success of this method relies on correct assessment of the primitive states of characters for the taxon. Also, in their examples using the exemplar method, only a single exemplar was used for each supraspecific taxon. The democratic method appeared to give the least accurate results.

In their analyses of the amniotes, Rieppel and Reisz (1999) and Lee (2001) used many supraspecific taxa, some of which contain huge numbers of known species, such as the Archosauriformes and Squamata. On many occasions during examination of the characters and scorings used by Rieppel and Reisz (1999) and Lee (2001), and when attempting to resolve their differences of opinion, a problem was encountered due to uncertainty over the exact method that had been used for scoring some of these supraspecific taxa. Many of the debated scorings were in supraspecific taxa that were scored as polymorphic for a large number of characters, suggesting that the scoring of the taxon was not based on a single exemplar, or a hypothetical groundplan. Instead, it appeared that these authors scored their supraspecific taxa using a method similar to the democratic method. However, instead of using the most common scoring of a number of exemplars, they combined the scorings of all of the exemplars into polymorphic scorings. So, if for the taxon 'Testudines', three exemplar species are examined, two of which exhibit state 0 for a character and the other state 1 for the same character, that character would be scored as polymorphic (0&1) for Testudines. Such an approach is counterproductive, because it leads to the production of many polymorphic scorings in the data matrix, which is essentially equivalent to having a large amount of missing data. This in turn can lead to loss of information about relationships and therefore a loss of resolution in the results of the analysis. This problem is especially obvious in some of the taxa that are most important in trying to resolve the debate over the origin of turtles. For example, many characters are coded as polymorphic for Testudines themselves, and due to apparent inconsistencies between characters, it is very difficult to understand which turtle species

have been used to score the Testudines. For example, if we consider the scorings of Rieppel and Reisz (1999) and Lee (2001) for characters 17 (lacrimal morphology) and 19 (skull proportions), such an inconsistency can be seen. The two authors agree on the scorings for these characters. For character 17, turtles are scored as possessing a lacrimal that is at least as long as tall, but excluded from the external nares. At present, *Proganochelys* is the only turtle known to possess a lacrimal (Gaffney, 1990). Although Rougier *et al.* (1995) scored *Palaeochersis* and *Australochelys* as plesiomorphic for their character 4, in which the derived state was "lacrimal bone and duct absent", the preorbital region of *Palaeochersis* is poorly preserved and there is no obvious lacrimal (Rougier *et al.*, 1995: Fig. 1), and *Australochelys* has a large lacrimal duct, but the sutures separating individual bones are impossible to discern (Gaffney and Kitching, 1995). Therefore, for character 17 it appears that *Proganochelys* is the only turtle scored in the analyses. However, a different picture is produced when looking at a number of other characters. In character 19, Testudines are scored as having skulls that are both equal in length preorbitally and postorbitally, and that are larger postorbitally. Obviously, this scoring suggests at least two species are being used to score the taxon and precludes the possibility that only *Proganochelys* has been scored. Furthermore, all of the most primitive turtles, including *Proganochelys*, *Palaeochersis*, *Australochelys* and *Kayentachelys,* have a longer postorbital than preorbital region. Therefore, more derived Testudines than these must have been included in the scoring of this character. When questioned about the taxa used to score the Testudines in the analysis, Olivier Rieppel (pers. comm. 2003) said that he used more than just *Proganochelys*, because of the difficulties of interpreting the material of that taxon. He said that in addition to *Proganochelys* he used all of the taxa present in Gaffney's (1979b) review of North American Jurassic turtles. Although this explains the reasons for the many polymorphic scorings within the Testudines, Rieppel's explanation does not resolve the inconsistencies discussed above, because none of the Jurassic forms described by Gaffney (1979b) possesses a lacrimal.

There are many similar examples in the scorings of Rieppel and Reisz (1999) and Lee (2001) that suggest the scorings for many taxa are based on different sets of species for different characters. The taxa that suffer the most are the Testudines, Squamata, Rhynchocephalia, Prolacertiformes, Archosauriformes and Rhynchosauria. For this reason, in the new matrix exemplars of each of these groups were scored (see Table 7.2). Although the exemplar method of coding inclusive taxa is not always accurate (Bininda-Emonds *et al.*, 1998), it has the advantage over the groundplan method in that it is more easily

repeatable and less subjective. Also, if the exemplars chosen are basal member of the groups involved, the chances of attaining an accurate result are increased, and, as stated by Bininda-Emonds *et al.* (1998) such taxa are very similar to the groundplan of the group they represent. It is also preferable to the method used by Rieppel and Reisz (1999) and Lee (2001), because it reduces the number of polymorphic taxa, and hence increases the proportion of informative characters.

Taxon Replaced	Exemplar(s) (References used for coding shown in parentheses)
Testudines	*Proganochelys quenstedti* (Gaffney, 1990; Kordikova, 2002) *Kayentachelys aprix* (Gaffney *et al*., 1987)
Rhynchosephalia	*Clevosaurus hudsoni* (Fraser, 1988)
Squamata	*Huehuecuetzpalli mixtecus* (Reynoso, 1998)
Prolacertiformes	*Prolacerta broomi* (Gow, 1975; Modesto and Sues, 2004) *Pamelaria dolichotrachela* (Sen, 2003)
Rhynchosauria	*Mesosuchus browni* (Dilkes, 1998) *Rhynchosaurus articeps* (Benton, 1990)
Archosauriformes	*Proterosuchus vanhoepeni* (Broom, 1903a; Cruickshank, 1972) *Euparkeria capensis* (Ewer, 1965; Gower and Weber, 1998; Senter, 2003) *Stagonolepis robertsoni* (Walker, 1961; Gower and Walker, 2002)
Eosauropterygia	*Simosaurus galliardoti* (Rieppel, 1994)

Table 7.2. Exemplars used as a replacement for some of the more inclusive taxa in the analyses of Rieppel and Reisz (1999) and Lee (2001).

The exemplars used in the new matrix (Table 7.2) were chosen because they are all early members of their respective groups, and are relatively complete, which allows scoring of most of the characters. This reworking of the matrix had the added benefit of allowing turtles to plot within the Archosauriformes, as has been suggested by a number of molecular analyses (e.g. Hedges and Poling, 1999; Cao *et al*., 2000). For this reason an aetosaur (*Stagonolepis robertsoni*), was included in the new analysis, because this group has been suggested as a possible turtle ancestor (Hedges and Poling, 1999). Similarly, the exemplar pareiasaurian, procolophonoid and placodont taxa included in the analysis of Rieppel and Reisz (1999) and combined by Lee (2001) were restored. Lee's (2001) 'outgroup' (which was based on the Diadectomorpha) was also replaced with Rieppel and Reisz's (1999) Seymouriidae and Diadectomorpha.

7.4 Analyses

7.4.1 Coding Methods and Character Strength

Mann-Whitney tests were used to see if any of the types of character coding discussed in Chapter 2 were removed earlier than expected by chance alone in the boildowns of the data of Rieppel and Reisz (1999) or Lee (2001) carried out in Chapter 6. The characters were ranked based on their average removal rank in the 'Number', CCSR and LQP boildowns of the matrices of Rieppel and Reisz (1999) and Lee (2001). Mann-Whitney tests were then carried out on each type of coding present in the data.

The results of the Mann-Whitney tests are shown in table 7.3.

SHAPE	RATIO	EXTENT	POSITIONAL	INAPPLICABLE DATA	UNIFYING	LOSS
0.0771	0.3159	**0.0004**	0.4485	0.0724	0.5839	0.4966

Table 7.3. Results of Mann-Whitney tests to see if characters of each coding type were removed earlier than expected by chance alone during the boildowns in Chapter 6. Significant results are highlighted in bold.

7.4.2 Phylogenetic Analysis of the New Matrix

The analysis of the rescored matrix (Appendix 3) was carried out using the same methods described in Chapter 6, and yielded two most parsimonious trees (TL = 656, CI = 0.309, RI = 0.690), the strict consensus of which is shown in figure 7.2. Average bootstrap and decay values for the tree are 60.1 and 4.4 respectively, with many relationships within the Diapsida showing low support. The topology is generally consistent with the results of the analyses of Rieppel and Reisz (1999) and Lee (2001). However, the new tree does differ from those of previous analyses in the position of the Testudines, which interestingly plot inside the Diapsida, nested within the Archosauriformes as the sister-group to the aetosaur, *Stagonolepis*. As with the analysis of Rieppel and Reisz (1999), Templeton tests show this diapsid placement is not a significantly better fit to the data than the most parsimonious tree with the two turtle taxa within the Anapsida (Templeton p-values range from 0.5383 to 0.5501).

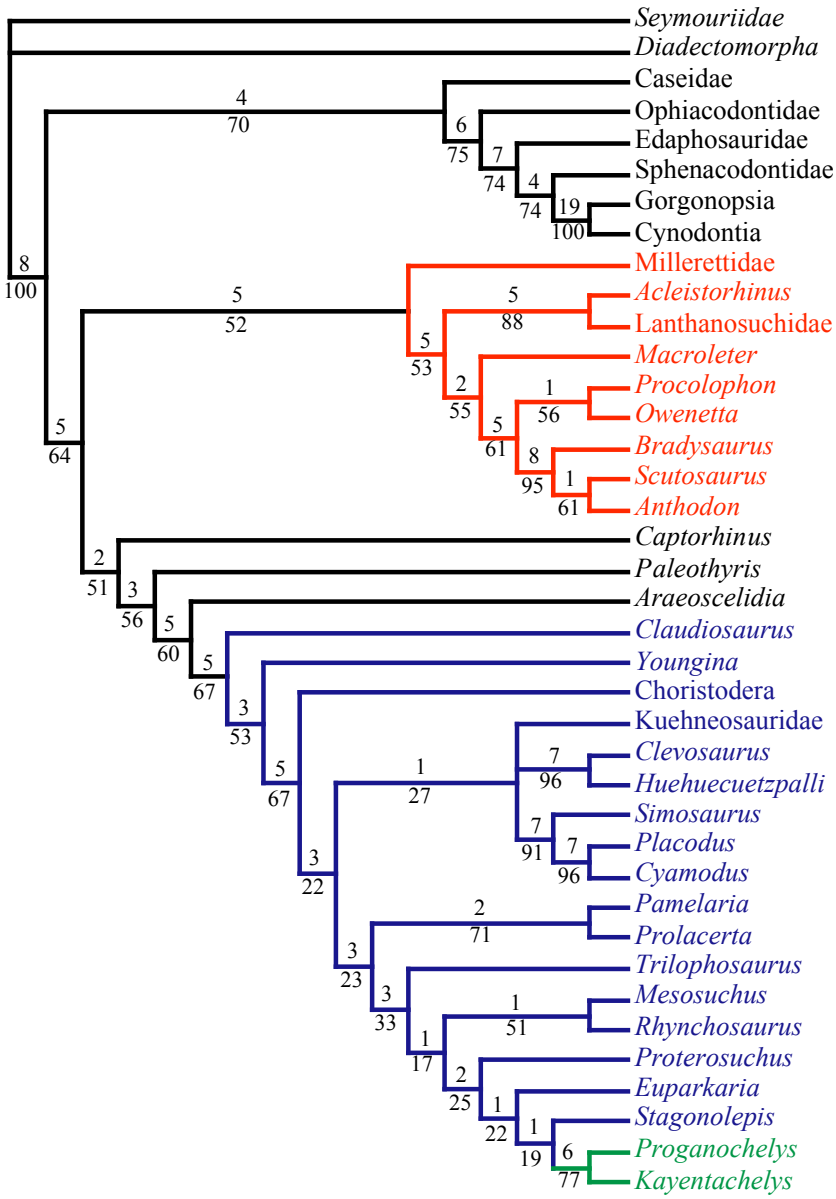


Figure 7.2. Strict consensus of two MPTs yielded by analysis of the rescored matrix which includes a number of exemplar taxa in replacement of highly inclusive taxa in the analyses of Rieppel and Reisz (1999) and Lee (2001). Numbers above nodes represent decay values, and those below nodes represent bootstrap proportions. The Anapsida are highlighted in red, the Diapsida in blue and turtles in green.

Results of a leaf stability analysis of the bootstrap trees (Table 7.4) show that the Gorgonopsia and Cynodontia are the most unstable taxa, closely followed by the two turtle taxa, *Proganochelys* and *Kayentachelys*. However, the average stability, which is a meaure of overall phylogenetic stability, is relatively high.

Leaf	Maximum	Difference	Entropy
Seymouriidae	1.000	1.000	1.000
Diadectomorpha	0.9997	0.9995	0.9978
Ophiacodontidae	0.9603	0.9253	0.8838
Edaphosauridae	0.9603	0.9253	0.8838
Sphenacodontidae	0.9603	0.9253	0.8840
Caseidae	0.9551	0.9134	0.8723
Youngina	0.9397	0.8884	0.8193
Claudiosaurus	0.9393	0.8875	0.8177
Araeoscelidia	0.9299	0.8662	0.8018
Protorothyrididae	0.9270	0.8620	0.7928
Nyctoleterida	0.9197	0.8534	0.7704
Procolophon	0.9195	0.8529	0.7719
Owenetta	0.9195	0.8530	0.7718
Bradysaurus	0.9146	0.8445	0.7600
Scutosaurus	0.9146	0.8445	0.7600
Anthodon	0.9146	0.8445	0.7600
Average	0.9116	0.8435	0.7699
Euparkeria	0.9114	0.8463	0.7754
Proterosuchus	0.9110	0.8455	0.7744
Stagonolepis	0.9109	0.8455	0.7743
Millerettidae	0.9103	0.8419	0.7398
Rhynchosaurus	0.9103	0.8442	0.7730
Mesosuchus	0.9096	0.8431	0.7719
Simosaurus	0.9069	0.8415	0.7692
Placodus	0.9064	0.8407	0.7705
Cyamodus	0.9063	0.8406	0.7705
Trilophosaurus	0.9020	0.8309	0.7568
Acleistorhinus	0.8990	0.8186	0.7145
Huehuecuetzpalli	0.8980	0.8253	0.7541
Clevosaurus	0.8978	0.8249	0.7540
Pamelaria	0.8976	0.8234	0.7548
Prolacerta	0.8973	0.8229	0.7547
Lanthanosuchidae	0.8968	0.8151	0.7079
Kuehneosauridae	0.8942	0.8212	0.7446
Captorhinidae	0.8846	0.8091	0.6794
Choristodera	0.8833	0.8001	0.7394
Kayentachelys	**0.8535**	**0.7396**	**0.6473**
Proganochelys	**0.8339**	**0.6984**	**0.6164**
Gorgonopsia	0.8289	0.6968	0.5683
Cynodontia	0.8289	0.6968	0.5683

Table 7.4. Maximum, difference and entropy leaf stability measures for the bootstrap trees in order of decreasing stability. The two turtle taxa are highlighted in bold.

Sequential sister-group removal was used to assess the strength of the nesting of the turtles in the Archosauriformes. Analysis of the data with *Stagonolepis* removed led to seven MPTs (TL = 649, CI = 0.313, RI = 0.684), with turtles plotting in a polytomy with the rhynchosaurs, Archosauriformes and *Trilophosaurus*. Removal of the five taxa that make up these groups led to the production a single MPT (TL = 567, CI = 0.356, RI = 0.682), which placed the turtles as the sister-group of the Sauropterygians, within the Diapsida. Removal of the three sauropterygian taxa led to the production of two MPTs, (TL = 503, CI = 0.402, RI = 0.693), in which the turtles are the sister-group of the Lepidosauria. When the two lepidosaurs are removed, a single tree is produced (TL = 455, CI = 0.431, RI = 0.703), in which the turtles plot within the Anapsida, as the sister-group of the pareiasaurs.

7.4.3 Taxon Jacknife

The matrix contains 8890 pairwise incompatibilities. The results of first- and second-order taxon jacknifes are shown in table 7.5. As with the matrices of Rieppel and Reisz (1999) and Lee (2001), and the consensus matrix, the taxa that are the cause of most incompatibility in the new matrix are the cynodontians, turtles and gorgonopsians.

Of the two turtles, *Proganochelys* is by far the cause of most incompatibilities. However, this is probably because *Kayentachelys* is scored as unknown for a large number of characters (66.3%). As with the analysis of Rieppel and Reisz (1999), the three pareiasaur taxa appear relatively compatible, but as discussed in Chapter 6, this is likely to be due to their similarity with one another.

Taxon Excluded	Uninf.	% Missing	1st Order	Exp	P	2nd Order
Cynodontia	9	10.5	613	341	0.999	946
Proganochelys	**8**	**13.3**	**600**	**301**	**0.999**	**924**
Clevosaurus	8	11.6	268	307	0.341	606
Gorgonopsia	8	9.4	257	326	0.216	609
Simosaurus	7	27.1	223	186	0.762	566
Placodus	8	16.6	213	253	0.285	557
Procolophon	8	8.3	172	339	0.007	515
Trilophosaurus	7	22.1	153	237	0.079	494
Proterosuchus	7	9.9	145	298	0.013	489
Choristodera	8	18.2	128	257	0.024	470
Huehuecuetzpalli	7	29.3	116	195	0.076	461
Bradysaurus	7	7.2	111	335	0.001	457
Claudiosaurus	7	14.9	111	270	0.006	456
Scutosaurus	9	6.6	104	322	0.001	450
Kuehneosauridae	7	33.1	99	153	0.156	445
Younginiformes	7	12.2	98	322	0.001	443
Captorhinidae	7	8.3	92	338	0.001	435
Lanthanosuchidae	8	43.6	90	111	0.312	433
Araeoscelidia	7	8.3	90	347	0.001	435
Rhynchosaurus	8	14.9	87	271	0.004	432
Edaphosauridae	7	15.5	82	292	0.001	428
Cyamodus	7	27.1	81	193	0.007	431
Sphenacodontidae	7	6.1	71	366	0.001	419
Nyctoleteridae	7	18.8	66	239	0.001	412
Stagonolepis	7	11	64	301	0.001	410
Anthodon	7	8.8	57	311	0.001	405
Millerettidae	7	17.1	49	269	0.001	395
Owenetta	8	28.2	49	168	0.003	396
Mesosuchus	8	15.5	42	250	0.001	388
Caseidae	7	5	39	379	0.001	386
Acleistorhinus	7	49.2	37	90	0.032	383
Seymouriidae	8	6.6	35	352	0.001	387
Ophiacodontidae	7	7.7	30	340	0.001	378
Kayentachelys	**7**	**66.3**	**28**	**30**	**0.513**	**374**
Pamelaria	7	11.6	19	291	0.001	366
Euparkeria	7	12.7	17	280	0.001	366
Diadectomorpha	8	9.9	16	331	0.001	368
Prolacerta	7	8.3	16	323	0.001	364
Protorothyrididae	7	10.5	3	320	0.001	350

Table 7.5. Results of first- and second-order taxon jacknife analyses of the new matrix in order of decreasing first-order jacknife incompatibilities. Uninf. = number of uninformative characters in matrix when the taxon is removed, % missing = percentage of taxon character scorings that are missing, caused = change in the number of incompatibilities in the data when the taxon is removed, expected = average number of incompatibilities caused by randomised versions of the taxon, p = proportion of randomised versions of the taxon that cause the same or fewer incompatibilities than the actual taxon, 2nd order = average number of incompatibilities caused by all pairs of taxa including the taxon in question. Turtles are highlighted in bold.

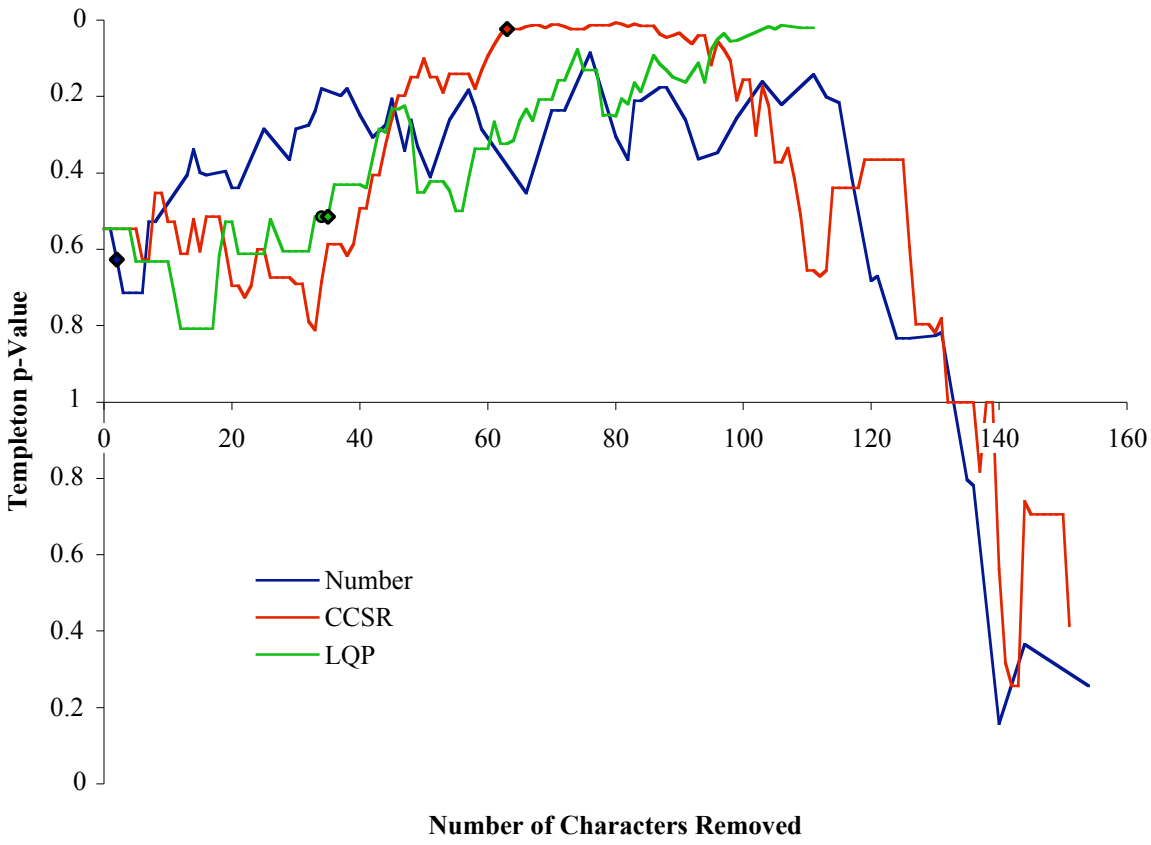


Figure 7.3. Line chart showing the Templeton test p-value when comparing the fit of the new matrix to trees with backbone constraints forcing turtles within the Anapsida and Diapsida at each stage of number, CCSR and LQP boildowns. The x-axis represents the number of characters removed by the boildown. The y-axis is the Templeton test p-value. Values above the x-axis indicate that the diapsid hypothesis is more parsimonious than the anapsid hypothesis, whereas those values below the x-axis indicate that the anapsid hypothesis is more parsimonious. Coloured diamonds indicate the point at which the boildown bootstrap identified maximum bootstrap values for each boildown. The green circle is the point at which characters removed in the LQP boildown have an LQP value lower than 0.05, and are therefore significantly more compatible than random characters.

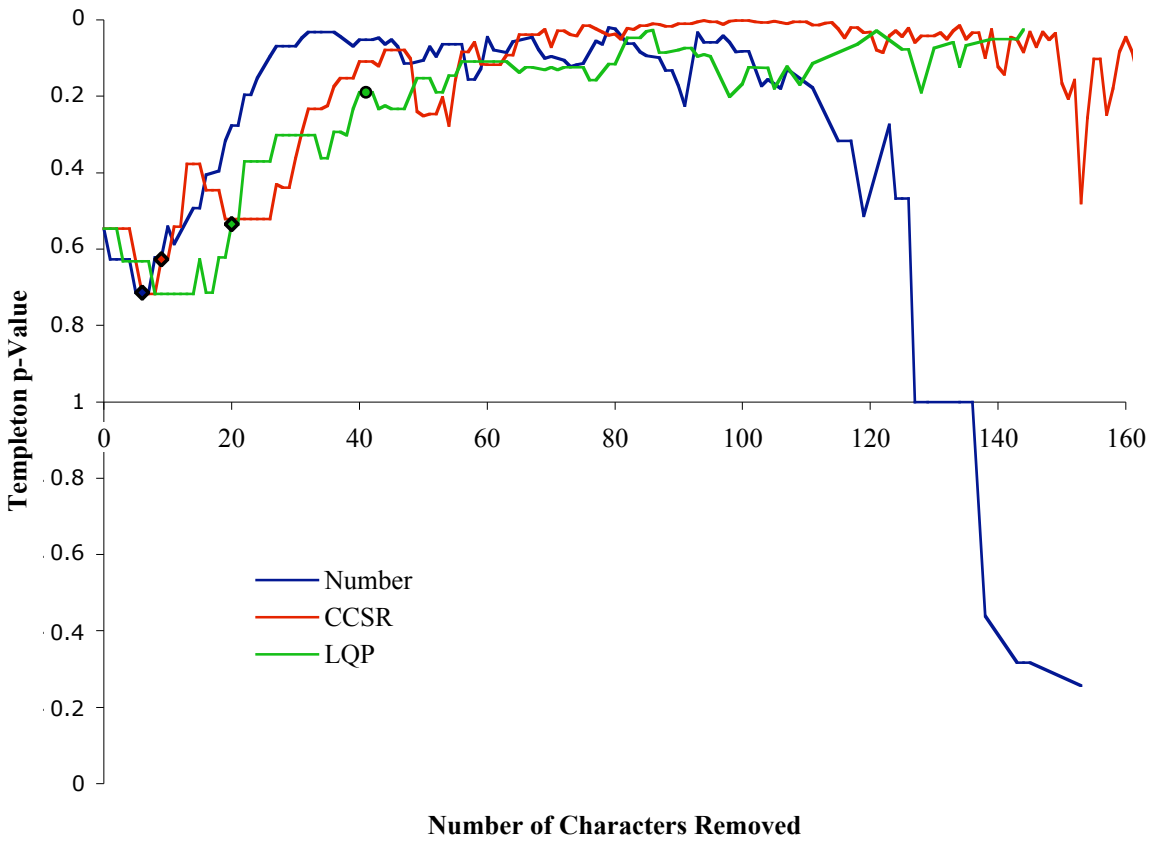


Figure 7.4. Line chart showing the Templeton test p-value when comparing the fit of the new matrix to trees with backbone constraints forcing turtles within the Anapsida and Diapsida at each stage of fuzzy number, CCSR and LQP boildowns. The x-axis represents the number of characters removed by the boildown. The y-axis is the Templeton test p-value. Values above the x-axis indicate that the diapsid hypothesis is more parsimonious than the anapsid hypothesis, whereas those values below the x-axis indicate that the anapsid hypothesis is more parsimonious. Coloured diamonds indicate the point at which the boildown bootstrap identified maximum bootstrap values for each boildown. The green circle is the point at which characters removed in the LQP boildown have an LQP value lower than 0.05, and are therefore significantly more compatible than random characters.

7.4.4 Boildown Analysis

Results of boildown analyses on the new matrix are shown in figure 7.3. Templeton tests at each stage of the boildown show an initial trend towards significant support for a diapsid placement for turtles. Significance is reached in the CCSR boildown when 62 characters are removed, and in the LQP analysis when 97 characters are removed. In the CCSR and 'Number' boildowns there is a late swing back towards support for the anapsid hypothesis.

Results of a fuzzy boildown on the new matrix are shown in figure 7.4. Again there is an initial trend towards significant support for the diapsid hypothesis. Significance is reached when 31 characters are removed in the 'Number' boildown, when 65 characters are removed in the CCSR boildown, and when 82 characters are removed in the LQP boildown. There is a late swing back towards the anapsid hypothesis in the 'Number' analysis.

7.5 Discussion

Re-examination of the characters used in the analyses of Rieppel and Reisz (1999) and Lee (2001) identified many character construction and scoring problems. Given the seniority of the authors involved, their reputations in the field of vertebrate palaeontology, and the fact that their matrices have become the basis of the debate over the affinities of turtles, the abundance of these problems was extremely surprising and shocking. The problems were in three main areas, clarity in character descriptions and scorings, the scoring of supraspecific taxa, and poor treatment of linkage between characters. The first of these, the lack of clarity in character descriptions was a major problem when trying to replicate the results attained by the previous authors. It should not be necessary to contact authors in order to understand the meanings of their characters or to attain corrections that make the published results and matrices agree. As discussed above, the scoring of supraspecific taxa in their analyses was confusing, and in the new matrix a number of these taxa were rescored using the exemplar method. The method used by Rieppel and Reisz (1999) and Lee (2001) for coding linked characters is potentially more problematic. Many characters in the data were identified as being logically and/or biologically linked, either completely or incompletely (see Chapter 3). An extreme example that illustrates the type of poor codings employed can be seen in characters 30 and 53. Both of these characters have a state for supratemporal absent. Obviously such a coding upweights the relationship

between taxa that lack supratemporals. Furthermore, these characters also highlight the problems in character scoring, because the scoring of taxa for the two identical states was not the same. The solutions to all of the identified linkages are explained in the character descriptions above. Another problematic issue that was identified was that of the coding of inapplicable data. The history of this often encountered and hotly debated area of character coding is discussed in depth in Chapter 2. In the characters used by Rieppel and Reisz (1999) and Lee (2001) there were many multistate inapplicable characters. Although some workers argue that multistate coding is the best available treatment for inapplicable data, here it is considered that the potential of such a coding to waste useful information about relationships makes it an undesirable method. Such characters were therefore recoded using the 'missing method' of coding inapplicable characters, as discussed in Chapter 2.

7.5.1 Weak Character Coding Types

Mann-Whitney tests showed that only characters coded using the extent method were removed significantly earlier in the boildown process than would be expected by chance alone ($p=0.0004$). However, both shape and positional codings had p-values very close to significance at the 5% level (0.0771 and 0.0724 respectively). These results suggest that characters coded based on the shape and position of elements are more incompatible than those coded using more conventional coding methods, such as presence or absence. Such a result is not altogether surprising. The shape of a bone or process and their relative positions seem, on the basis of biological knowledge, to be more likely to change rapidly during the evolution of a group, leading to saturation of character states and/or reversals and convergences, whereas the gain or loss of an entire bony element seems less likely to occur, and therefore may be subject to less homoplasy.

7.5.2 The Phylogeny

The analysis presented here is the first quantitative osteological analysis to suggest that the origins of turtles are located within the Archosauriformes. However, as previously mentioned, this placement is not novel. Almost all recent molecular analyses of mitochondrial data have placed turtles as the sister-group of archosaurs (e.g. Zardoya and Meyer, 1998; Kumazawa and Nishida, 1999; Zardoya and Meyer, 2001; Rest *et al*., 2003), while analyses of nuclear genes have placed them within this group as the sister-group of crocodiles (e.g. Hedges and Poling, 1999; Cao *et al*., 2000). Such a result led Hedges and Poling (1999) to speculate that the Triassic aetosaurs, which are herbivorous, heavily

armoured relatives of the crocodiles, may be closely related to Testudines. Unfortunately, until now, none of the larger morphological analyses aimed at finding the origin of turtles has allowed a placement within the crown-group Archosauria. All recent analyses have included a single aggregate 'Archosauriformes' taxon, which precludes the possibility of turtles plotting within the group, as the molecular evidence might be suggesting. In the analysis presented here, three archosauriform taxa were scored for the existing characters. Because the original data contained only a single archosauriform, there were no characters within the data that defined this group. However, even without the addition of many possible characters to define the Archosauriformes and their internal resolution, the interactions of the existing characters managed to resolve the group into the expected position (e.g. see Rieppel and Reisz, 1999; Lee, 2001) and with the expected internal resolution (e.g. see Gower and Wilkinson, 1996; Dilkes, 1998), although the lack of additional characters means that the group has very few synapomorphies (see Fig. 7.5). For the same reasons, the positioning of turtles within the Archosauriformes is unexpected. With no additional characters included there are few synapomorphies linking turtles with archosaurs or aetosaurs. Characters highlighted by Hedges and Poling (1999), such as similar cervical spines possessed by *Proganochelys* and some aetosaurs were not included. The obvious link between these two groups in the existing data is the fact that both possess a large carapace covering their entire back. However, these are not the only two taxa possessing this trait, and if this were the only evidence supporting such a placement, it would be expected that once the aetosaur was removed during sequential sister-group removal, the turtles would revert to one of their more familiar positions. This was not the case. When the aetosaur, *Stagonolepis*, was removed, turtles still plotted within the archosauromorphs, although the resolution of this group became poor.

Despite the interesting result thrown up by this analysis, it cannot be claimed that it satisfactorily resolves the problem of the phylogenetic affinities of the Testudines. Templeton tests again show that the diapsid placement found by this analysis is not a significantly better fit to the data than the most parsimonious tree in which turtles plot within the Anapsida. Similarly, tracing the characters onto the two alternative trees (Fig. 7.5) shows that virtually all characters show at least some homoplasy, and therefore almost no clades on the phylogeny are supported by true, unreversed and non-convergent synapomorphies. It is likely that this high level of homoplasy within the data is a result of the vast scale of the analysis, encompassing the whole of the amniotes and spanning huge periods of time, which increases the chance of homoplasy occurring. However, as with all

of the matrices analysed in Chapter 6, compatibility methods suggest that the data supporting the diapsid placement are less homoplastic than those supporting the anapsid placement. This is shown in the results of the boildown analyses, which all show an initial trend towards greater support for the diapsid hypothesis. As with the data analysed in Chapter 6, the late trend towards the anapsid hypothesis is likely to be due to the dearth of characters remaining in the data (see Chapter 6 for more detail).

At the very least, the agreement of this analysis, however weak, with that of molecular data deserves further attention, and shows that exclusion of taxa may prohibit the attainment of more parsimonious phylogenies and lead to support for erroneous relationships.

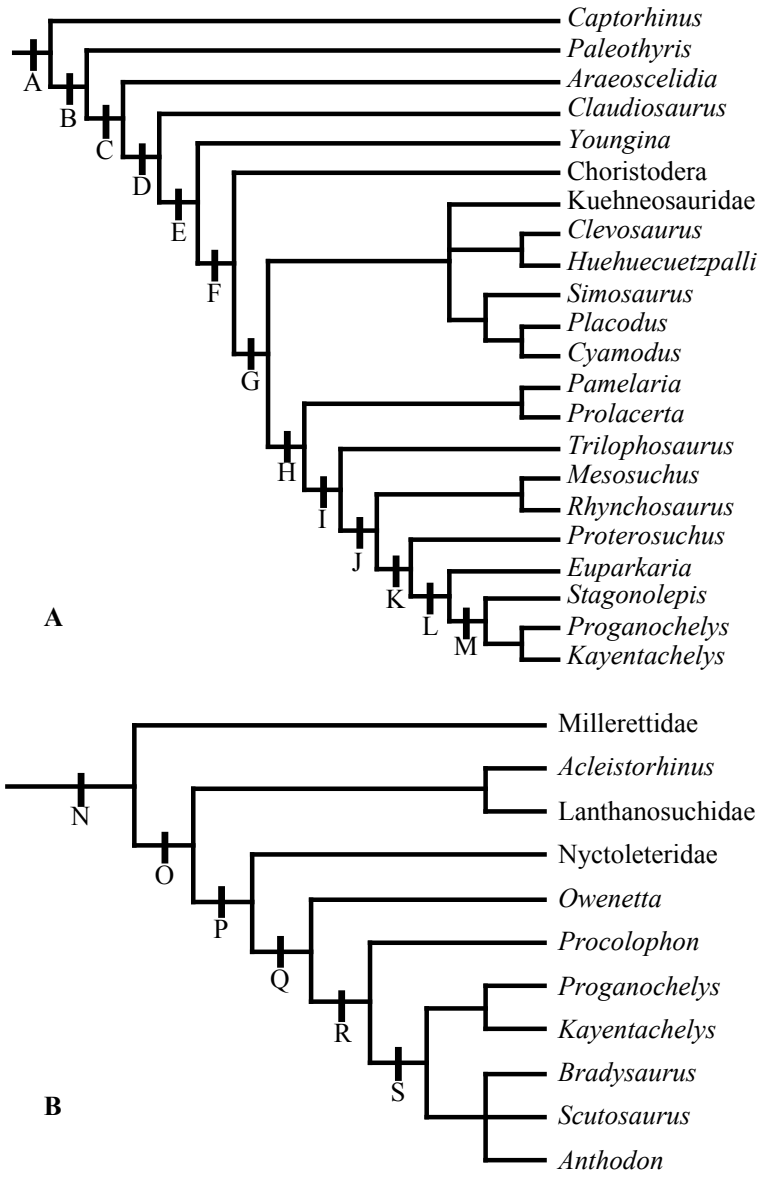


Figure 7.5. The most parsimonious diapsid (A) and anapsid (B) placements for turtles in the new matrix, showing synapomorphies for all clades including the Testudines.

A) 36) Reversed in Rhynchocephalia. Convergent in *Owenetta*.

41) Reversed in Testudines.

53b) Convergent in some synapsids.

B) 62) Reversed in *Cyamodus*. Convergent in some synapsids and *Acleistorhinus*.

103) Convergent in some synapsids and at node R.

162) Reversed in Placodontia and *Proganochelys*, unknown in *Kayentachelys*. Convergent in Cynodontia, Nycteroleteridae and *Owenetta*.

163) Reversed in Placodontia and *Proganochelys*, unknown in *Kayentachelys*. Convergent in Cynodontia and Millerettidae.

C) 19) State – 1. Reversed to state 0 in Kuehneosauridae, Rhynchocephalia and *Mesosuchus*. Derived to state 2 in Placodontia, Testudines, *Rhynchosaurus* and Choristodera. Convergent in Lanthanosuchidae and some synapsids.

50) Reversed in Testudines.

51a) Reversed in *Trilophosaurus* and Testudines. Convergent in Synapsida, *Acleistorhinus* and Lanthanosuchidae.

88b) Reversed in Placodontia, *Stagonolepis* and *Rhynchosaurus*. Convergent in Seymouriidae and at node O, but reversed in Pareiasauridae.

116) Reversed in Placodontia. Convergent in Gorgonopsia and Cynodontia.

160) Convergent at node R.

D) 17b) State – 1. Derived to state 2 in Squamata and *Rhynchosaurus*. Convergent in Sphenacodontidae, Gorgonopsia, Cynodontia and at node O excluding Pareiasauridae.

32) Reversed in Placodontia, Lepidosauria, Testudines and *Rhynchosaurus*. Convergent in Synapsida and Lanthanosuchidae.

55a) Polymorphic in *Youngina*. Convergent in *Captorhinus* and at node P.

79) Reversed in Placodontia, Choristodera, Rhynchosauria and Archosauriformes. Convergent in Edaphosauridae and at node S.

95) Convergent in Edaphosauridae, Lanthanosuchidae and at node R.

102)

111) Convergent in Gorgonopsia, Cynodontia and Pareiasauridae.

113) Convergent in Cynodontia, *Procolophon*, *Scutosaurus* and *Anthodon*.

114) Reversed in Sauropterygia and Testudines. Convergent in Cynodontia.

115) State – 1. Derived to state 2 in Lepidosauria. Convergent in Gorgonopsia, Cynodontia and at node O.

119)

120) Defines ingroup, but reversed at node D. Convergent reversal in Millerettidae.

125) Reversed in Kuehneosauridae. Convergent in Millerettidae and *Anthodon*.

130) Reversed in Lepidosauria and *Trilophosaurus*. Convergent in Cynodontia and at node Q.

132) Reversed at node I and in *Prolacerta*. Convergent in Cynodontia, *Stagonolepis* and *Euparkaria*.

137) Reversed in *Proterosuchus*, *Mesosuchus* and *Proganochelys*, unknown in *Kayentachelys*. Convergent in Gorgonopsia, Cynodontia and at node R.

139) Reversed in Eosauropterygia. Convergent in Cynodontia.

142) Reversed in Eosauropterygia and *Proganochelys*, unknown in *Kayentachelys*. Convergent in Gorgonopsia and Cynodontia.

144) Convergent in Gorgonopsia, Cynodontia and at node P.

E) 38b) Convergent at node P except Pareiasauridae.

39)

45) Equivocal in *Owenetta*, *Claudiosaurus* and *Kayentachelys*. Reversed in *Proganochelys*. Convergent in *Procolophon*.

46) Equivocal in *Claudiosaurus* and Nycteroleteridae. Convergent in Cynodontia and at node Q.

92) Convergent in some synapsids, *Captorhinus* and at node N.

110) Reversed in Eosauropterygia. Convergent in Gorgonopsia, Cynodontia, and at node R.

145) Reversed in Eosauropterygia. Convergent in *Araeoscelis*.

154) Reversed in Sauropterygia and *Rhynchosaurus*.

F) 1) Reversed in Kuehneosauridae, Squamata and Testudines.

18) Reversed in Rhynchosauria. Convergent in Cynodontia.

42b) State – 2. Derived to state 1 at node H. Convergent in some synapsids, *Claudiosaurus*, *Euparkaria* and *Mesosuchus*.

52a) Convergent in *Procolophon*.

67) Equivocal from node B to node F. Convergent in Diadectomorpha and some synapsids.

69) State 0 is derived. Reversed to state 1 in Prolacertiformes and *Mesosuchus*. Convergent in Pareiasauridae.

70) Convergent in some synapsids.

75) Reversed in Prolacertiformes. Convergent in Gorgonopsia, Cynodontia and at node Q.

99) Reversed in *Placodus* and Rhynchocephalia. Convergent in some synapsids.

105) Reversed in Lepidosauria. Convergent in *Claudiosaurus*.

128) Reversed in Lepidosauria. Convergent in *Owenetta*.

129) Reversed in Rhynchocephalia and at node I.

131) Reversed in *Trilophosaurus*. Convergent in Gorgonopsia and Cynodontia.

153) Convergent in Gorgonopsia, Cynodontia and at node N.

155) Reversed in Sauropterygia and *Rhynchosaurus*. Convergent in Cynodontia.

G) 27) Reversed in Placodontia, Rhynchosauria and at node M.

35) State – 1. Reversed to state 0 in Rhynchocephalia and *Mesosuchus*. Derived to state 2 in Kuehneosauridae. Convergent in Procolophonoidea.

63) Equivocal from node B to G. Convergent in Pareiasauridae and some synapsids.

65) State – 1. Derived to state 2 at node K. Convergent at node Q, but derived to state 2 at node S.

80) Equivocal from node E to G. Convergent in Pareiasauridae.

140) Reversed at node K. Convergent in Gorgonopsia, Cynodontia, *Proganochelys* and at node S.

H) 6) Reversed in *Placodus*, Rhynchocephalia and at node M. Convergent in Cynodontia and at node O excluding Testudines.

11) Convergent in some synapsids and at node O.

37) Character defines ingroup, reversed at node H, then reversed back at node L.

42b) State – 1. Reversed to state 2 in *Euparkaria* and *Mesosuchus*. Convergent in Procolophonoidea.

60) State – 1. Reversed to state 0 in *Kayentachelys* and *Euparkaria*. Convergent in Choristodera, *Cyamodus* and Rhynchocephalia.

136)

149) State – 2. Reversed to state 1 in *Proganochelys*, unknown in *Kayentachelys*. Convergent in Gorgonopsia and Cynodontia.

I) 5a) Convergent in Kuehneosauridae, Sauropterygia and *Bradysaurus*.

61) Convergent in Gorgonopsia, Cynodontia and Pareiasauridae.

66) Convergent in some synapsids, Rhynchocephalia, Choristodera and at node R.

78a) Reversed in *Proterosuchus*. Convergent in Gorgonopsia, Cynodontia, Sauropterygia, Lepidosauria and *Pamelaria*.

85) Convergent in Placodontia and at node Q. Equivocal from node O to Q.

151) Reversed in *Proganochelys*, unknown in *Kayentachelys*. Convergent in Gorgonopsia, Cynodontia and Choristodera.

157) Reversed in *Euparkaria* and *Stagonolepis*. Convergent in *Araeoscelis* and *Youngina*.

J) 33) Reversed at node M. Convergent in Rhynchocephalia.

71) State –1. Equivocal from node H to J.

82) Reversed to state 0 at node M. Convergent in Choristodera and Sauropterygia.

123) Convergent in Cynodontia, *Anthodon* and Sauropterygia.

147) Convergent in Gorgonopsia, Cynodontia and Lepidosauria.

164) Convergent in *Pamelaria* and at node R.

K) 10) Reversed in Testudines. Convergent in Prolacertiformes, Choristodera, and some synapsids. State – 2. Convergent in Pareiasauridae and some synapsids.

76) Convergent in some synapsids, *Paleothyris*, Choristodera, Eosauropterygia, Lepidosauria, Lanthanosuchidae and at node Q.

108) Convergent in Kuehneosauridae and Placodontia.

127) State – 1. Derived to state 2 in *Proganochelys*, unknown in *Kayentachelys*. Convergent in *Procolophon* and *Claudiosaurus*.

158) State – 2. Convergent in Lepidosauria and at node R.

165a) Convergent in Pareiasauridae and Placodontia.

L) 37) Character defines ingroup, reversed at node H, then reversed back at node L.

49a) Convergent in Squamata, *Trilophosaurus*, *Rhynchosaurus*, *Pamelaria* and Choristodera.

64) Convergent in Gorgonopsia, Cynodontia, *Claudiosaurus*, *Placodus*, Prolacertiformes and Pareiasauridae.

117) State – 1. Derived to state 2 in *Proganochelys*, unknown in *Kayentachelys*. Convergent in Synapsida, *Trilophosaurus* and at node P.

159) Convergent in Cynodontia and Pareiasauridae.

M) 83) State – 1. Uncertain between states 1 and 2 in *Stagonolepis*. Convergent in outgroups, *Youngina*, Prolacertiformes, *Cyamodus*, Squamata, *Trilophosaurus*, *Rhynchosaurus* and at node O.

165b) Convergent in Placodontia and *Anthodon*.

167) Convergent in *Scutosaurus* and *Anthodon*.

N) 13) Reversed in Testudines.

20) Equivocal in Testudines. Convergent in *Placodus*, but equivocal in all other taxa above node F.

82) State – 2. Reversed to state 0 in *Proganochelys*. Polymorphic in Lanthanosuchidae. Convergent in Cynodontia and *Trilophosaurus*.

92) Reversed in *Acleistorhinus*.Convergent in some synapsids, *Captorhinus* and at node E.

134) Reversed in *Proganochelys*, unknown in *Kayentachelys*. Convergent in *Youngina* and Eosauropterygia.

153) Convergent in Gorgonopsia, Cynodontia and at node F.

O) 6) Reversed in Testudines. Convergent in Cynodontia and at node H.

11) Convergent in some synapsids and at node H.

16) Reversed in Procolophonoidea and Testudines.

17b) Reversed in Pareiasauridae. Convergent in Sphenacodontidae, Gorgonopsia, Cynodontia and at node D.

46) Equivocal in *Claudiosaurus* and Nycteroleteridae. Convergent in Cynodontia and at node E.

49b) State – 2. Equivocal in many diapsids. Polymorphic in Lanthanosuchidae. Convergent in some synapsids, Kuehneosauridae, Placodontia, Lepidosauria and *Mesosuchus*.

83) State – 1. Convergent in outgroups, *Youngina*, Prolacertiformes, *Cyamodus*, Squamata, *Trilophosaurus*, *Rhynchosaurus* and at node M.

88b) Reversed in Pareiasauridae. Convergent in Seymouriidae and at node C, but reversed in Placodontia, *Stagonolepis* and *Rhynchosaurus*.

106) State – 2.

115) State – 1. Convergent in Gorgonopsia, Cynodontia and at node D.

P) 38b) Reversed in Pareiasauridae. Convergent at node E.

55a) Convergent in *Captorhinus* and at node D.

86) Equivocal from node N to P. Reversed in *Proganochelys*, unknown in *Kayentachelys*. Convergent in Cynodontia, *Placodus*, Rhynchocephalia and Rhynchosauria.

109) Equivocal from node O to P. Reversed in Testudines. Convergent in some synapsids and Sauropterygia.

117) State – 1. Derived to state 2 in *Anthodon* and *Proganochelys*, unknown in *Kayentachelys*. Convergent in Synapsida, *Trilophosaurus* and at node L.

144) Equivocal from node N to P. Convergent in Gorgonopsia, Cynodontia and at node D.

Q) 14) Reversed in Testudines. Convergent in many diapsid taxa.

56) Convergent in Rhynchocephalia, *Trilophosaurus* and Rhynchosauria.

65) State – 1. Derived to state 2 at node S. Convergent at node G, but derived to state 2 at node K.

75) Convergent in Gorgonopsia, Cynodontia and at node F.

76) Convergent in some synapsids, *Paleothyris*, Choristodera, Eosauropterygia Lepidosauria, Lanthanosuchidae and at node K.

77) Reversed in *Kayentachelys*. Convergent in *Araeoscelis*, *Claudiosaurus*, *Trilophosaurus* and *Stagonolepis*.

85) Equivocal from node O to Q. Convergent in Placodontia and at node I.

87) Convergent in Eosauropterygia, and *Cyamodus*.

130) Convergent in Cynodontia and at node D.

R) 66) Convergent in some synapsids, Rhynchocephalia, Choristodera and at node I.

95) Convergent in Edaphosauridae, Lanthanosuchidae and at node D.

103) Convergent in some synapsids and at node B.

110) Equivocal from node N to R. Convergent in Gorgonopsia, Cynodontia, and at node E.

113) Equivocal from node Q to R. Reversed in *Bradysaurus*. Convergent in Cynodontia and at node D.

126a) Convergent in Cynodontia, *Captorhinus*, *Claudiosaurus*, Eosauropterygia, *Trilophosaurus* and *Mesosuchus*.

137) Equivocal from node Q to R. Reversed in *Proganochelys*, unknown in *Kayentachelys*. Convergent in Gorgonopsia, Cynodontia and at node D.

138) Convergent in Gorgonopsia, Cynodontia and Nycteroleteridae.

149) State – 1. Equivocal from node Q to R. Convergent in Lepidosauria.

150) State – 1. Convergent in Cynodontia and Lepidosauria.

158) State – 2. Equivocal from node Q to R. Convergent in Lepidosauria and at node K.

160) Equivocal from node Q to R. Convergent at node C.

164) Equivocal from node Q to R. Convergent in *Pamelaria* and at node J.

S) 8) Convergent in Placodontia, Gorgonopsia and Cynodontia.

22) Reversed in *Kayentachelys*. Convergent in the outgroups, Cynodontia, *Placodus* and *Trilophosaurus*.

26) Convergent in Lanthanosuchidae and *Placodus*.

61) Convergent in Gorgonopsia, Cynodontia and at node I.

63) Equivocal from node B to G. Convergent in some synapsids and at node G.

64) Convergent in Gorgonopsia, Cynodontia, *Claudiosaurus*, *Placodus*, Prolacertiformes and at node L.

65) State – 2. Convergent in some synapsids and at node K.

69) State 0 is derived. Convergent at node F.

79) Convergent in Edaphosauridae and at node D, but reversed in Placodontia, Choristodera, Rhynchosauria and Archosauriformes.

80) Convergent at node G.

97)

98) Reversed in *Bradysaurus*.

111) Convergent in Gorgonopsia, Cynodontia and at node D.

112)

127) State – 2. Convergent in Gorgonopsia, Cynodontia and Lepidosauria.

140) Convergent in Gorgonopsia, Cynodontia and at node G.

159) Convergent in Cynodontia and at node L.

161) State – 2. Convergent in Gorgonopsia and Cynodontia.

165a) Equivocal from node R to S. Convergent in Placodontia and at node K.

7.6 Conclusions: Anapsid or Diapsid?

The cautious conclusion of the study of the origins of turtles presented in this thesis would be that without new specimens bridging the morphological gap between turtles and their true relatives we cannot provide compelling support for any phylogenetic position for the group. We cannot even confidently assign Testudines to the Anapsida or Diapsida. However, being bolder, it can be suggested that the agreement between theoretical methods of signal enhancement and morphological methods of reassessing hypotheses of

homology, character constructions and scorings, and taxon selection should be, in light of the problematic nature of the debate, considered persuasive. Therefore, until further evidence is presented, it is concluded here that the best explanation of the current data is that turtles originated from somewhere within the Diapsida, especially when results of molecular analyses are also taken into account. Once the diapsid origin is accepted, pinpointing the position within this group has also proved problematic. The analysis presented here supports the result of many molecular analyses, placing turtles within the Archosauria, as close relatives of the Aetosauria. This was an unexpected result, and is likely to be considered controversial by many morphologists. However, the debate over turtle origins has never been anything other than controversial, and the potential archosaurian affinities of turtles should at the least be investigated further in future.

7.7 Future Work

In order to improve our knowledge of, and hopefully to ultimately solve, the problem of the affinities of turtles, it is recommended here that future work should be focussed on a number of areas:

- The character construction and scoring issues raised here should be reconsidered and improved. Further checks should also be carried out on the current characters, because, given the abundance of scoring errors and character construction problems found in this study, some mistakes are likely to have been missed. Poor characters should be removed, and further characters identified. One limitation of this study was that the scorings of taxa were made from the literature rather than specimens. This means that not all scorings could be checked, and in the newly coded taxa, some scorings could not be made. A future analysis would be improved by checking all scorings on the actual specimens.
- The re-examination of the character constructions used in the current analyses showed that a high proportion involved splitting continuous data into discrete states. This is a well-known problem in character construction. In the current analyses many such characters were constructed using relative lengths or extents of structures. This data could be further explored using morphometric techniques (e.g. see Wiens, 2000).
- The taxon set upon which future studies are focussed should be different to those of the past. It is generally accepted today that turtles are either derived from the Anapsida or Diapsida. Therefore, including synapsids in the analysis is of little use, and leads to

extra difficulty in character construction and scoring, because of the increased diversity they bring to the analysis, which increases homology assessment problems. It was shown that some of the most unstable taxa in the current analyses were members of this group, whose inclusion essentially added nothing to our knowledge of amniote phylogeny. Removal of the synapsids would mean that some characters in the current matrix could be removed, and might make finding new characters easier. Furthermore, removing this set of taxa would allow more taxa to be added in the Anapsida and Diapsida. Most importantly, more archosauriform exemplars should be added, especially in the crown-group archosaurs where this analysis and many molecular analyses place turtles. Although the analysis presented here supports aetosaurs as the turtle sister-group, no other crown-group archosaurs are included, and the addition of birds and crocodiles, for example, is necessary to understand the implications of the results found here. Along with the addition of new taxa, new characters would have to be sought to provide resolution in the tree.

- The debate could also be explored further by carrying out more restrictive analyses, including turtles and just the Diapsida or Anapsida. This would allow a different set of characters to be identified, and may reduce the levels of homoplasy in the data caused by attempting to homologise across such a diverse taxon set as found in the current analyses. Reducing this homoplasy may allow better resolution within the groups and provide greater support for a stable position of the turtles.

Chapter 8: Summary and Conclusions

In light of the vast and increasing number of phylogenetic analyses that are carried out every year, it seems surprising that little attention has been paid to discussion of the processes of character construction and scoring that, along with weighting, completely determine the results of any analysis of morphology. The aims of this thesis, identified in Chapter 1.4, were mainly concerned with attempting to address some of the issues that frequently occur during this stage of phylogenetic analyses.

Most of the literature concerning character construction revolves around the philosophy and definitions of some of the most important concepts in the field: homology and the character concept. These two issues are discussed in Chapter 1.3, and I concluded that a concise and precise definition of either word is not essential. More important are the practical aspects of morphological character construction and scoring that have received relatively little discussion.

In Chapter 2, a survey of the types of characters used in a sample of 100 analyses revealed a great deal of variation in how various methods of character construction are applied in the zoological literature. It was also noted that very few authors discussed their reasons for their character construction choices. These results were consistent with a similar survey of botanical analyses (Hawkins, 2000). It was also shown that there were significant differences in the types of characters used by workers in the fields of palaeontology and neontology, but even more obvious were the dramatic differences between individual studies by different authors. As long as such inconsistencies are prevalent in character construction, repetition of studies will be all but impossible, and comparing results of cladistic analyses by different authors will be difficult. Also in Chapter 2, the strengths and weaknesses of a number of alternative character construction methods were highlighted, and some suggestions as to the best method in certain circumstances were made. For instance, the much-debated inapplicable data problem was discussed in some detail. Although it was concluded that there is no perfect treatment currently available for inapplicable data, I recommended the use of the “missing” method, which, although in certain circumstances suffers from its own problems, maximises the potential information in the data that can be utilised for phylogeny reconstruction.

Chapter 3 confronted some further issues of character construction, particularly the treatment of characters involving structures that are repeated within a single organism (intraorganismal homologues). The chapter was written in conjunction with Dr Mark

Wilkinson and Dr David Gower of the Natural History Museum, London as a paper that was published in Systematic Biology (Harris *et al*., 2003a). The chapter started with a review of the three currently published phylogenies of the Aetosauria, a group of Triassic archosaurs. These three published analyses had very different results, so the characters they used were looked at in detail. It was noted that a number of mistakes had been made in the character construction and scoring stages of each of the analyses. The most common problem, and that which had most effect on the outcome of the analyses, was regarding the treatment of the intraorganismally homologous dermal osteoderms that form a carapace covering the back of aetosaurs. The authors of the published studies tended to split this carapace into regions for character construction and scoring based purely on standard vertebral regions (cervical, dorsal and caudal). In many cases there was no difference in the scoring of the osteoderms across these different regions, meaning that regional subdivision simply upweighted potential homologues in cladistic analysis. It was concluded in Chapter 3 that the best treatment for features of intraorganismal homologues was to combine them into single characters in the absence of evidence for independent evolutionary transformations, such as where the distributions of the states of those characters among taxa differed in different regions of the body. This conclusion is consistent with Hennig's auxiliary principle, which is an important theoretical and operational basis of homology assessment and phylogenetic inference using parsimony. A new matrix that combined all independent characters from the three published studies and resolved all of the character construction and scoring problems identified in the review was produced and analysed. This is the most in depth phylogenetic analysis of the Aetosauria to date, and shows that little can be concluded about the internal resolution of the group from the currently available data. The authors of one of the original aetosaur studies published a reply to the published version of Chapter 3 (Heckert and Lucas, 2003) complaining that it is more fossils that are needed to resolve the problems in aetosaur phylogeny, and not more analyses. In our own reply (Harris *et al*., 2003b) we agreed that more fossils are needed, but countered that without proper analysis of their morphology the phylogeny produced is still likely to be misleading.

Chapters 4 to 6 looked at theoretical methods available for assessing the 'strength' of signal in phylogenetic characters, especially in debates between two or more alternative phylogenetic hypotheses for a group have reached stalemate. Generally, parsimony analysis and related tree-based methods are the only techniques used to explore the signals present in morphological data, and this narrow view may mean that interesting patterns are

being overlooked. Chapter 4 described a number of little-used compatibility-based analysis techniques that can be applied to the data matrix prior to tree production in an attempt to identify the characters that fit least well with the rest of the data. Two new methods, fuzzy compatibility, and the boildown bootstrap, were also introduced in this Chapter. These methods are an attempt to improve upon the previously available techniques. Fuzzy compatibility provides a measure of how incompatible two incompatible characters are, giving a more accurate picture of the levels of conflict in the data. The boildown bootstrap is a method for trying to identify the point in a boildown analysis (see Chapter 4) at which signal is being removed from the data rather than random noise. Chapter 5 introduces my new *Boildown* computer program that is designed to carry out the methods described in Chapter 4, including the two new analytical techniques. Although previous programs were available for carrying out some of these methods (e.g. Wilkinson, 2001a), this new program is more user-friendly, with a menu-based Mac OS X interface, uses new algorithms to improve upon previous programs, and includes the new methods described in Chapter 4. A CD containing the *Boildown* program can be found in the back cover of this thesis, and will also be made freely available online.

Chapter 6 is a case study that illustrates how the techniques described in Chapters 4 and 5 can be used. It is based on the age-old debate over the affinities of turtles. Compatibility methods were applied to the two most recent analyses of turtle affinities that support the two main alternative hypotheses, which position turtles within the Anapsida and Diapsida. In virtually all tests carried out, the results suggested that characters supporting the historically popular anapsid hypothesis were the cause of significantly more conflict in the data than characters supporting the diapsid hypothesis. This result was unexpected, both because of the great deal of support for the anapsid hypothesis from most workers in the field, and because only the hypothesis supporting an anapsid turtle positioning has been supported by statistical significance in previous studies. The compatibility tests used identified a number of taxa including turtles, pareiasaurs, sauropterygians, gorgonopsids and cynodontids as particularly problematic. Interestingly, most of the subsignals identified within the data stemmed from the instability of these five taxa, which often changed position dramatically, usually to form clades with one another in various combinations. This, it was concluded, suggests that the two contradictory results produced by parsimony analyses of similar data may be the result of long-branch attraction, a well-known problem in analyses of molecular data, but rarely identified with morphological data.

This case study, although not necessarily solving the problem of turtle affinities difinitively, does highlight the utility of the compatibility methods described in Chapter 4 for data exploration beyond parsimony analysis. Compatibility methods have generally been forgotten, because they appear to be less suitable for phylogeny reconstruction than parsimony, and it is not suggested here that compatibility should be used as an alternative to parsimony. However, it is known that parsimony has its own flaws and biases, and is itself simply a model of evolution that may not be accurate. For that reason, it is recommended here that phylogenetic analyses should not be composed simply of a parsimony analysis. Using alternative approaches, such as compatibility, likelihood and Bayesian methods alongside parsimony can highlight when data are strong and/or weak. If a strong signal is present in a dataset, all methods would be expected to identify that signal and produce results that are in agreement. Where the results using the different methods disagree, there should be less confidence placed in those results. However, compatibility has even more potential for identifying noise and subsignals in a dataset, and for this reason should be afforded more attention in future work.

Chapter 7 looked at the same turtle debate in a more empirical way, by reconsidering character constructions and scorings used in recent analyses. A number of problematic character construction issues were raised, especially regarding linkage between characters and the methods used for scoring supraspecific taxa. Scorings that were considered poor or incorrect were also corrected, and resolutions to the scorings over which previous authors have disagreed were made, where possible. A revised matrix containing a number of exemplar taxa in place of the groundplan scorings of supraspecific taxa in the original matrices and all corrections to character constructions and scorings was produced. Analysis of this matrix agreed with the analytical results in Chapter 6, in that it supported a diapsid placement for turtles, although support for this resolution over the best anapsid placement was not statistically significant.

Although the results from Chapters 6 and 7 would not necessarily be considered convincing alone, the fact that analytical and empirical techniques give highly correlated results (supporting a diapsid affinity of turtles) does add weight to those findings. It appears that the current morphological data cannot satisfactorily resolve the debate, and because of the diversity of extinct data between the major lineages of living reptiles, molecular evidence is also of limited use. It is possible that new data could be collected from the currently available taxa that could lead to a acceptable solution, but it would be useful, and may be essential to find new fossils intermediate between turtles and their

closest relatives. However, it seems that the best explanation of the current data is that it supports the diapsid hypothesis of turtle origins.

Conclusions

- Currently, cladistic analyses are suffering from a lack of consistency at the initial character construction stage that is the most important part of the analysis. This area requires a great deal of future work.
- The lack of consistency stems from a lack of discussion of character construction issues in the literature, with far more effort being afforded to attempts to concisely define concepts such as homology and the cladistic character than to providing practical guidelines as to the best way to construct characters in certain circumstances.
- Along with the inconsistency in character construction, there is also a lack of discussion by authors of cladistic analyses as to why they chose to construct and score their characters in a certain way.
- One of the most common types of character coding in the literature is that dealing with the issue of inapplicable data caused by the hierarchical nature of characters. Although at present there is no perfect way to code such characters, the best appears to be the 'missing' method, because it preserves the maximum information in the data and is only problematic in extreme circumstances.
- Another problematic character construction issue arises when organisms possess intraorganismal homologues. However, a simple rule can be applied to these features: attributes of intraorganismally homologous structures should be treated as single characters unless the distribution of the attributes of the structure among taxa differs in different regions (or between different individual structures) of the organisms.
- Parsimony should not be considered the be-all and end-all of phylogenetic analysis. Many other methods are available that allow further investigation of the data.
- Compatibility provides a useful tool for data exploration that can provide information that cannot be identified using parsimony alone. It is especially useful for identifying the characters in the data that cause conflict and for recognising competing subsignals in datasets. However, the available methods need further testing and application to many more case studies before their attributes can be fully understood.
- The treatment of supraspecific taxa needs to be carefully considered when creating a data matrix. If a groundplan method is used, the make-up of that groundplan must be

explicitly explained to allow future workers to understand the results of the analysis and check and reanalyse the data. Including a number of exemplars is preferable to creating a groundplan if, as in the case of the turtle analysis, generating the groundplan taxa leads to the inclusion of many polymorphic scorings in the dataset, which reduces the amount of signal in the data. This can lead to problems finding the correct phylogeny, or separating that tree from other trees, which can in turn cause the formation of an unresolved consensus tree.

- The debate over the affinities of turtles appears to be in a state of deadlock that cannot be satisfactorily solved with the currently available data. However, The best explanation of the current data is that turtles are a group of diapsid reptiles in which the temporal foramina have fused closed, probably in conjunction with the development of the shell, to provide protection from predators.
- The affinities of turtles may be solved in future by reassessing the morphology of known taxa, including additional anapsid and diapsid taxa and excluding synapsid taxa, and identifying new characters. However, the solution may require identification of new fossils that bridge the large morphological gap between turtles and their closest relatives.

References

ALROY, J. 1994. Four permutation tests for the presence of phylogenetic structure. *Systematic Biology* **43**:430-437.

ARCHIE, J. W. 1985. Methods for coding variable morphological features for numerical taxonomic analysis. *Systematic Zoology* **34**:326-345.

ARCHIE, J. W. 1989. A randomization test for phylogenetic information in systematic data. *Systematic Zoology* **38**:219-252.

BAUM, B. R. 1992. Combining trees as a way of combining data sets for phylogenetic inference, and the desirability of combining gene trees. *Taxon* **41**:3-10.

BAUR, G. 1887. Ueber den Ursprung der Extremitäten der Ichthyopterygia. *Jahresberichte und Mitteilungen des Oberrheinischen Geologischen Vereines* **20**:17-20.

BAUR, G. 1889. On the morphology of the vertebrate skull. *Journal of Morphology* **3**:467-474.

BENTON, M. J. 1990. The species of *Rhynchosaurus*, a rhynchosaur (Reptilia, Diapsida) from the Middle Triassic of England. *Philosophical Transactions of the Royal Society of London* **328**:213-306.

BININDA-EMONDS, O. R. P., H. N. BRYANT, and A. P. RUSSELL. 1998. Supraspecific taxa as terminals in cladistic analysis: implicit assumptions of monophyly and a comparison of methods. *Biological Journal of the Linnean Society* **64**:101-133.

BISHOP, M. J., and A. E. FRIDAY. 1988. Estimating the interrelationships of tetrapod groups on the basis of molecular sequence data. Pages 33-58 *in* The Phylogeny and Classification of the Tetrapods (M. J. Benton, ed.) Clarendon Press, Oxford.

BOCK, H.-H. 1963. Evolution and phylogeny in morphologically uniform groups. *American Naturalist* **97**:265-285.

BOCK, W. J. 1973. Philosophical foundations of classical evolutionary classification. *Systematic Zoology* **22**:375-392.

BOONSTRA, L. D. 1929a. Pareiasaurian studies. Part III. On the pareiasaurian manus. *Annals of the South African Museum* **28**:97-111.

BOONSTRA, L. D. 1929b. Pareiasaurian studies. Part IV. On the pareiasaurian pes. *Annals of the South African Museum* **28**:113-122.

BOONSTRA, L. D. 1932a. Pareiasaurian studies. Part VII. On the hind limb of two little-known pareiasaurian genera: *Anthodon* and *Pareiasaurs*. *Annals of the South African Museum* **28**:429-435.

BOONSTRA, L. D. 1932b. Pareiasaurian studies. Part VIII. The osteology and myology of the locomotor apparatus. B. Fore limb. *Annals of the South African Museum* **28**:437-503.

BOONSTRA, L. D. 1934a. Pareiasaurian studies. Part IX. The cranial osteology. *Annals of the South African Museum* **31**:1-38.

BOONSTRA, L. D. 1934b. Pareiasaurian studies. Part X. The dermal armour. *Annals of the South African Museum* **31**:39-48.

BOONSTRA, L. D. 1934c. Pareiasaurian studies. Part XI. The vertebral column and ribs. *Annals of the South African Museum* **31**:49-66.

BOYDEN, A. 1973. Perspectives in zoology. Pergamon Press, Oxford.

BREMER, K. 1988. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. *Evolution* **42**:795-803.

BRIGANDT, I. 2002. Homology and the origin of correspondence. *Biology and Philosophy* **17**:389-407.

BROOM, R. 1903a. On a new species of reptile (*Proterosuchus fergusi*) from the Karoo Beds of Terkastad, South Africa. *Annals of the South African Museum* **4**:159-163.

BROOM, R. 1903b. On an almost perfect skeleton of *Pareiasaurus serridens* Owen. *Annals of the South African Museum* **4**:123-138.

BROOM, R. 1912. On a new species of *Propappus, and on the pose of the pareiasaurian limbs*. *Annals of the South African Museum* **7**:323-330.

BROOM, R. 1913. On the manus and pes of *Pareiasaurus*. *Annals of the South African Museum* **7**:353-357.

BROOM, R. 1922. On the temporal arches of the Reptilia. *Proceedings of the Zoological Society of London* **1922**:17-27.

BROOM, R. 1924. On the classification of the reptiles. *Bulletin of the American Museum of Natural History* **51**:39-65.

CAMIN, J. H., and R. R. SOKAL. 1965. A method for deducing branching sequences in phylogeny. *Evolution* **19**:311-326.

CAO, Y., M. D. SORENSON, Y. KUMAZAWA, D. P. MINDELL, and M. HASEGAWA. 2000. Phylogenetic position of turtles among amniotes: Evidence from mitochondrial and nuclear genes. *Gene* **259**:139-148.

CARROLL, R. L. 1969a. A middle Pennsylvanian captorhinomorph, and the interrelationships of primitive reptiles. *Journal of Paleontology* **43**:151-170.

CARROLL, R. L. 1969b. Origin of reptiles. Pages 1-44 *in* Biology of the Reptilia (C. Gans, A. d. A. Bellairs, and T. S. Parsons, eds.). Academic Press, London and New York.

CARROLL, R. L. 1975. Permo-Triassic "lizards" from the Karroo. *Palaeontologia Africana* **18**:71-87.

CARROLL, R. L. 1988. Vertebrate paleontology and evolution. W. H. Freeman and Co.

CARROLL, R. L., and P. M. GALTON. 1977. 'Modern' lizards from the Upper Triassic of China. *Nature* **266**:252-255.

CARROLL, R. L., and W. LINDSAY. 1985. Cranial anatomy of the primitive reptile *Procolophon*. *Canadian Journal of Earth Sciences* **22**:1571-1587.

CHATTERJEE, S. 1974. A Rhynchosaur from the Upper Triassic Maleri Formation of India. *Philosophical Transactions of the Royal Society of London* **267**:209-261.

CHEN, D., L. DIAO, O. EULENSTEIN, D. FERNÁNDEZ-BACA, and M. J. SANDERSON. 2003. Flipping: a supertree construction method. Pages 135-160 *in* Bioconsensus, volume 61 of DIMACS: Series in Discrete Mathematics and Theoretical Computer Sciences (M. Janowitz, F.-J. Lapointe, F. R. McMorris, B. Mirkin, and F. S. Roberts, eds.). American Mathematical Society, Providence.

CLARK, J., and R. L. CARROLL. 1973. Romeriid reptiles from the Lower Permian. *Bulletin of the Museum of Comparative Zoology* **144**:353-407.

COLBERT, E. H. 1966. A gliding reptile from the Triassic of New Jersey. *American Museum Novitates* **2246**:1-23.

COLBERT, E. H. 1970. The Triassic gliding reptile *Icarosaurus*. *Bulletin of the American Museum of Natural History* **143**:85-142.

COLBERT, E. H., and M. MORALES. 1991. Evolution of the vertebrates. John Wiley and Sons Inc., New York.

COLLESS, D. H. 1985. On "character" and related terms. *Systematic Zoology* **34**:229.

COPE, E. D. 1892. On the homologies of the posterior cranial arches in the reptilia. *Transactions of the American Philosophical Society* **17**:11-26.

COX, C. B. 1969. The problematic Permian reptile *Eunotosaurus*. *Bulletin of the British Museum (Natural History) Geology* **18**:166-196.

CRUICKSHANK, A. R. I. 1972. The proterosuchian thecodonts. Pages 89-119 *in* Studies in vertebrate evolution (K. A. Joysey, and W. P. Kemp, eds.). Oliver and Boyd, Edinburgh.

DARWIN, C. 1859. The origin of species. 1968 Edition, Penguin Books Ltd., Harmondsworth.

DE LA FUENTE, M. S., G. W. ROUGIER, and A. B. ARCUCCI. 1996. New specimens of *Palaeochersis talampayensis*: additions to its anatomy and relationships. *Journal of Vertebrate Paleontology* **16**:31A.

DE PINNA, C. C. 1991. Concepts and tests of homology in the cladistic paradigm. *Cladistics* **7**:367-394.

DEBRAGA, M. 2003. The postcranial skeleton, phylogenetic position, and probable lifestyle of the Early Triassic reptile *Procolophon trigonoceps*. *Canadian Journal of Earth Sciences* **40**:527-556.

DEBRAGA, M., and R. R. REISZ. 1995. A new diapsid reptile from the uppermost Carboniferous of Kansas. *Palaeontology* **38**:199-212.

DEBRAGA, M., and R. R. REISZ. 1996. The Early Permian reptile *Acleistorhinus pteroticus* and its phylogenetic position. *Journal of Vertebrate Paleontology* **16**:384-395.

DEBRAGA, M., and O. RIEPPEL. 1997. Reptile phylogeny and the interrelationships of turtles. *Zoological Journal of the Linnean Society* **120**:281-354.

DILKES, D. W. 1995. The rhynchosaur *Howesia* from the Lower Triassic of South Africa. *Palaeontology* **38**:665-685.

DILKES, D. W. 1998. The early Triassic rhynchosaur *Mesosuchus brownii* and interrelationships of basal archosauromorph reptiles. *Philosophical Transactions of the Royal Society of London* **353**:501-541.

DODICK, J. T., and S. P. MODESTO. 1995. The cranial anatomy of the captorhinid reptile *Labidosaurikos meachami* from the Lower Permian of Oklahoma. *Palaeontology* **38**:687-711.

DONOGHUE, M. J., R. G. OLMSTEAD, J. F. SMITH, and J. D. PALMER. 1992. Phylogenetic relationships of Dipsacales based on rbcL sequences. *Annals of the Missouri Botanical Gardens* **79**:333-345.

DUNLOP, J. A., and S. J. BRADDY. 2001. Scorpions and their sister-group relationships. Pages 1-24 *in* Scorpions 2001. In Memoriam Gary A. Polis (V. Fet, and P. A. Selden, eds.). British Arachnological Society, Burnham Beeches.

EERNISSE, D. J., J. S. ALBERT, and F. E. ANDERSON. 1992. Annelida and Arthropoda are not sister taxa: A phylogenetic analysis of spiralian metazoan morphology. *Systematic Biology* **41**:305-330.

ELDREDGE, N. 1979. Cladism and common sense. Pages 165-198 *in* Phylogenetic analysis and paleontology (J. Cracraft, and N. Eldredge, eds.). Columbia University Press, New York.

ELDREDGE, N., and J. CRACRAFT. 1980. Phylogenetic patterns and the evolutionary process. Columbia University Press, New York.

EMERSON, S. B., and P. HASTINGS. 1998. Morphological correlations in evolution: Consequences for phylogenetic analysis. *Quarterly Review of Biology* **73**:141-162.

ESTABROOK, G. F. 1983. The causes of character incompatibility. Pages 279-295 *in* Numerical taxonomy (J. Felsenstein, ed.) Springer-Verlag, Berlin.

ESTABROOK, G. F. 1984. Phylogenetic trees and character state trees. Pages 135-151 *in* Cladistics: perspectives on the reconstruction of evolutionary history (T. Duncan, and T. F. Stuessy, eds.). Columbia University Press, New York.

ESTABROOK, G. F., C. S. JOHNSON JR., and F. R. MCMORRIS. 1975. An idealized concept of the true cladistic character. *Mathematical Biosciences* **23**:263-272.

ESTABROOK, G. F., C. S. JOHNSON JR., and F. R. MCMORRIS. 1976. A mathematical foundation for the analysis of cladistic character compatibility. *Mathematical Biosciences* **29**:181-187.

ESTABROOK, G. F., and L. R. LANDRUM. 1975. A simple test for the possible simultaneous evolutionary divergence of two amino acid positions. *Taxon* **24**:609-613.

ESTABROOK, G. F., and F. R. MCMORRIS. 1977. When are two qualitative taxonomic characters compatible? *Journal of Mathematical Biology* **4**:195-200.

ESTABROOK, G. F., J. G. STRAUCH JR., and K. L. FIALA. 1977. An application of compatibility analysis to the Blackiths' data on orthopteroid insects. *Systematic Zoology* **26**:269-276.

EVANS, S. E., G. V. R. PRASAD, and B. K. MANHAS. 2001. Rhynchocephalians (Diapsida: Lepidosauria) from the Jurassic Kota Formation of India. *Zoological Journal of the Linnean Society* **133**:309-334.

EWER, R. F. 1965. The anatomy of the thecodont reptile *Euparkeria capensis* Broom. *Philosophical Transactions of the Royal Society of London* **248**:379-435.

FAITH, D. P., and P. S. CRANSTON. 1991. Could a cladogram this short have arisen by chance alone?: On permutation tests for cladistic structure. *Cladistics* **7**:1-28.

FARRIS, J. S. 1969. A successive approximations approach to character weighting. *Systematic Zoology* **19**:83-92.

FARRIS, J. S. 1973a. On comparing the shapes of taxonomic trees. *Systematic Zoology* **22**:50-54.

FARRIS, J. S. 1973b. A probability model for inferring evolutionary trees. *Systematic Zoology* **22**:250-256.

FARRIS, J. S. 1988. Hennig86. Port Jefferson, New York.

FARRIS, J. S., V. A. ALBERT, M. KÄLLERAJÖ, D. LIPSCOMB, and A. G. KLUGE. 1996. Parsimony jacknifing outperforms neighbour-joining. *Cladistics* **12**:99-124.

FARRIS, J. S., A. G. KLUGE, and M. J. ECKHARDT. 1970. A numerical approach to phylogenetic systematics. *Systematic Zoology* **19**:172-189.

FELSENSTEIN, J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. *Systematic Zoology* **27**:401-410.

FELSENSTEIN, J. 1982. Numerical methods for inferring evolutionary trees. *Quarterly Review of Biology* **57**:379-404.

FELSENSTEIN, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **39**:783-791.

FELSENSTEIN, J. 1995. PHYLIP: phylogenetic inference package, version 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle.

FIALA, K. L. 1984. CLINCH, version 6.2. State University of New York.

FITCH, W. M. 1975. Toward finding the tree of maximum parsimony. Pages 189-230 *in* Proceedings of the 8th international Freeman Conference in numerical taxonomy (G. F. Estabrook, ed.) W. H. Freeman, San Francisco.

FITCH, W. M. 1977. On the problem of discovering the most parsimonious tree. *American Naturalist* **111**:223-257.

FITCH, W. M. 2000. Homology. A personal view on some of the problems. *Trends in Genetics* **16**:227-231.

FRASER, N. 1988. The osteology and relationships of *Clevosaurus* (Reptilia: Sphenodontida). *Philosophical Transactions of the Royal Society of London* **321**:125-178.

GAFFNEY, E. S. 1975a. *Solnhofia parsoni*, a new cryptodiran turtle from the Late Jurassic of Europe. *American Museum Novitates* **2576**:1-25.

GAFFNEY, E. S. 1975b. A taxonomic revision of the Jurassic turtles *Portlandemys* and *Plesiochelys*. *American Museum Novitates* **2574**:1-19.

GAFFNEY, E. S. 1979a. Comparative cranial morphology of recent and fossil turtles. *Bulletin of the American Museum of Natural History* **164**:65-376.

GAFFNEY, E. S. 1979b. The Jurassic turtles of North America. *Bulletin of the American Museum of Natural History* **162**:91-136.

GAFFNEY, E. S. 1990. The comparative osteology of the Triassic turtle *Proganochelys*. *Bulletin of the American Museum of Natural History* **194**:1-263.

GAFFNEY, E. S., J. H. HUTCHISON, F. A. JENKINS JNR., and L. J. MEEKER. 1987. Modern turtle origins: the oldest known cryptodire. *Science* **237**:289-291.

GAFFNEY, E. S., and J. W. KITCHING. 1995. The morphology and relationships of *Australochelys*, an Early Jurassic turtle from South Africa. *American Museum Novitates* **3130**:1-29.

GAFFNEY, E. S., and P. A. MEYLAN. 1988. A phylogeny of turtles. Pages 157-219 *in* The Phylogeny and Classification of the Tetrapods (M. J. Benton, ed.) Clarendon Press, Oxford.

GARDINER, B. G. 1982. Tetrapod classification. *Zoological Journal of the Linnean Society* **74**:207-232.

GARLAND JR., T., P. H. HARVEY, and A. R. IVES. 1992. Procedures for the analysis of comparative data using phylogenetically independent characters. *Systematic Biology* **41**:18-32.

GAULD, I., and G. UNDERWOOD. 1986. Some applications of the LeQuesne compatibility test. *Biological Journal of the Linnean Society* **29**:191-222.

GAUTHIER, A. G. KLUGE, and T. ROWE. 1988a. Amniote phylogeny and the importance of fossils. *Cladistics* **4**:105-209.

GAUTHIER, J. A., A. G. KLUGE, and T. ROWE. 1988b. The early evolution of the amniota. Pages 103-155 *in* The Phylogeny and Classification of the Tetrapods (M. J. Benton, ed.) Clarendon Press, Oxford.

GEOFFROY SAINT-HILAIRE, É. 1818. Philosophie anatomique. Méquignon-Marvis, Paris.

GHISELIN, M. T. 1969. The distinction between similarity and homology. *Systematic Zoology* **18**:148-149.

GHISELIN, M. T. 1976. The nomenclature of correspondence: A new look at "homology" and "analogy". Pages 129-142 *in* Evolution, brain and behaviour: Persistent problems. (R. B. Masterson, W. Hodos, and H. Jerison, eds.). Lawrence Erlbaum, Hillsdale, New Jersey.

GOLOBOFF, P. A. 1998. NONA, version 2.0. INSUE Fundación e Instituto Miguel Lillo, Miguel Lillo 205, 4000 S. M. de Tucumon, Argentina.

GOLOBOFF, P. A., and D. POL. 2002. Semi-strict supertrees. *Cladistics* **18**:514-525.

GOODRICH, E. S. 1916. On the classification of the Reptilia. *Proceeding of the Royal Society of London B* **89**:261-276.

GOW, C. E. 1972. The osteology and relationships of the Millerettidae (Reptilia: Cotylosauria). *Journal of Zoology, London* **167**:219-264.

GOW, C. E. 1975. The morphology and relationships of *Youngina capensis* Broom and *Prolacerta broomi* Parrington. *Palaeontologia Africana* **18**:89-131.

GOWER, D. J., and A. D. WALKER. 2002. New data on the braincase of the aetosaurian archosaur *Stagonolepis robertsoni* Agassiz. *Zoological Journal of the Linnean Society* **136**:7-23.

GOWER, D. J., and E. WEBER. 1998. The braincase of *Euparkeria*, and the evolutionary relationships of birds and crocodilians. *Biological Reviews* **73**:367-411.

GOWER, D. J., and M. WILKINSON. 1996. Is there any consensus on basal archosaur phylogeny? *Proceedings of the Royal Society of London B: Biological Sciences* **263**:1399-1406.

GRANT, T., and A. G. KLUGE. 2004. Transformation series as an ideographic character concept. *Cladistics* **20**:23-31.

GREGORY, W. K. 1946. Pareiasaurs versus placodonts as near ancestors to the turtles. *Bulletin of the American Museum of Natural History* **86**:278-326.

GUISE, A., D. PEACOCK, and T. GLEAVES. 1982. A method for identification of parallelism in discrete character sets. *Zoological Journal of the Linnean Society* **74**:293-303.

GÜNTHER, A. 1867. Contribution to the anatomy of *Hatteria* (Rhynchocephalus, Owen). *Philosophical Transactions of the Royal Society of London* **157**:595-629.

GUYER, C., and J. B. SLOWINSKI. 1991. Comparison of observed phylogenetic topologies with null expectations among three monophyletic lineages. *Evolution* **45**:340-350.

GUYER, C., and J. B. SLOWINSKI. 1993. Adaptive radiation and the topology of large phylogenies. *Evolution* **47**:253-263.

HALL, B. K. 2003. Descent with modification: the unity underlying homology and homoplasy as seen through an analysis of development and evolution. *Biological Reviews* **78**:409-433.

HARCOURT-BROWN, K. G. 2002. Phylogenetic tree shape with special reference to the phylogeny of the Cretaceous globotruncanid foraminifera. Ph.D. Thesis, University of Bristol.

HARCOURT-BROWN, K. G., P. N. PEARSON, and M. WILKINSON. 2001. The imbalance of paleontological trees. *Paleobiology* **27**:188-204.

HARRIS, S. R., D. J. GOWER, and M. WILKINSON. 2003a. Intraorganismal homology, character construction, and the phylogeny of aetosaurian archosaurs (Reptilia, Diapsida). *Systematic Biology* **52**:239-252.

HARRIS, S. R., D. J. GOWER, and M. WILKINSON. 2003b. Phylogenetic methods and aetosaur interrelationships: a rejoinder. *Systematic Biology* **52**:851-852.

HAUGHTON, S. H., and L. D. BOONSTRA. 1930a. Pareiasaurian studies. Part V. On the pareiasaurian mandible. *Annals of the South African Museum* **28**:261-289.

HAUGHTON, S. H., and L. D. BOONSTRA. 1930b. Pareiasaurian studies. Part VI. The osteology and myology of the locomotor apparatus. A. Hind limb. *Annals of the South African Museum* **28**:297-367.

HAUSER, D. L., and W. PRESCH. 1989. The effect of ordered characters on phylogeny reconstruction. *Cladistics* **7**:243-265.

HAWKINS, J. A. 2000. A survey of primary homology assessment: different botanists perceive and define characters in different ways. Pages 22-53 *in* Homology and Systematics. (R. Scotland, and T. R. Pennington, eds.). Taylor and Francis.

HAWKINS, J. A., C. E. HUGHES, and R. W. SCOTLAND. 1997. Primary homology assessment, characters and character states. *Cladistics* **13**:275-283.

HEATON, M. J., and R. R. REISZ. 1980. A skeletal reconstruction of the Early Permian captorhinid reptile *Eocaptorhinus laticeps* (Williston). *Journal of Paleontology* **54**:136-143.

HECHT, M. K., and J. L. EDWARDS. 1977. The methodology of phylogenetic inference above the species level. Pages 3-51 *in* Major patterns in vertebrate evolution. (M. K. Hecht, P. C. Goody, and B. M. Hecht, eds.). Plenum Press, New York.

HECKERT, A. B., A. P. HUNT, and S. G. LUCAS. 1996. Redescription of *Redondasuchus reseri*, a Late Triassic aetosaur (Reptilia: Archosauria) from New Mexico (USA), and the biochronology and phylogeny of aetosaurs. *Geobios* **29**:619-632.

HECKERT, A. B., and S. G. LUCAS. 1999. A new aetosaur (Reptilia: Archosauria) from the Upper Triassic of Texas and the phylogeny of aetosaurs. *Journal of Vertebrate Paleontology* **19**:50-68.

HECKERT, A. B., and S. G. LUCAS. 2000. Taphonomy, phylogeny, biostratigraphy, biochronology, paleobiogeograhy, and evolution of the Late Triassic Aetosauria (Archosauria: Crurotarsi). *Zentralblatt für Geologie und Paläontologie* **11-12**:1539-1587.

HECKERT, A. B., and S. G. LUCAS. 2003. Clarifying aetosaur phylogeny requires more fossils, not more trees-reply to intraorganismal homology, character construction, and the phylogeny of aetosaurian archosaurs (Reptilia, Diapsida). *Systematic Biology* **52**:253-255.

HEDGES, S. B., and L. L. POLING. 1999. A molecular phylogeny of reptiles. *Science* **283**:998-1001.

HENDY, M. D., and D. PENNY. 1989. A framework for the quantitative study of evoutionary trees. *Systematic Zoology* **38**:297-309.

HENNIG, W. 1966. Phylogenetic systematics. University of Illinois Press, Urbana.

HILLIS, D. M., J. P. HUELSENBECK, and D. L. SWOFFORD. 1994. Hobgoblin of phylogenetics? *Nature* **369**:363-364.

HUELSENBECK, J. P. 1995. The performance of phylogenetic methods in simulation. *Systematic Biology* **44**:17-48.

HUELSENBECK, J. P., and D. M. HILLIS. 1993. Success of phylogenetic methods in the four-taxon case. *Systematic Biology* **42**:247-264.

HUELSENBECK, J. P., and M. KIRKPATRICK. 1996. Do phylogenetic methods produce trees with biased shapes? *Evolution* **50**:1418-1424.

HUELSENBECK, J. P., and R. NIELSEN. 1999. Effect of nonindependent substitution on phylogenetic accuracy. *Systematic Biology* **48**:317-328.

HUNT, A. P., and S. G. LUCAS. 1991. A new aetosaur from the Redonda Formation (Late Triassic: Middle Norian) of east-central New Mexico, USA. *Neues Jahrbuch für Geologie und Paläontologie, Monatshefte* **1991**:728-736.

JAEKEL, O. 1902. Über *Placochelys* n.g. und ihre Bedeutung für die Stammesgeschichte der Schildkröten. *Neues Jahrbuch für Mineralogie, Geologie und Paläontologie* **1**:127-144.

JALIL, N.-E. 1997. A new prolacertiform diapsid from the Triassic of North Africa and the interrelationships of the Prolacertiformes. *Journal of Vertebrate Paleontology* **17**:506-525.

JOYCE, W. G. 2000. The first complete skeleton of *Solnhofia parsonsi* (Cryptodira, Eurysternidae) from the Upper Jurassic of Germany and its taxonomic implications. *Journal of Paleontology* **74**:684-700.

KÄLLERAJÖ, M., V. A. ALBERT, and J. S. FARRIS. 1999. Homoplasy *increases* phylogenetic structure. *Cladistics* **15**:91-93.

KEARNEY, M. 2002. Fragmentary taxa, missing data, and ambiguity: Mistaken assumptions and conclusions. *Systematic Biology* **51**:369-381.

KEMP, T. S. 1978. The primitive cynodont *Procynosuchus*: functional anatomy of the skull and relationships. *Philosophical Transactions of the Royal Society of London* **285**:73-122.

KEMP, T. S. 1979. The primitive cynodont *Procynosuchus*: structure, function and evolution of the postcranial skeleton. *Philosophical Transactions of the Royal Society of London* **288**:217-258.

KOLLAR, E. J., and C. FISHER. 1980. Tooth induction in chick epithelium: expression of quiescent genes for enamel synthesis. *Science* **207**:993-995.

KORDIKOVA, E. G. 2002. Comparative morphology of the palate dentition in *Progeanochelys questedti* Baur 1887 from the Upper Triassic of Germany and chelonian ancestry. *Neues Jahrbuch für Geologie und Paläontologie, Abhandlungen* **225**:195-249.

KORNET, D. J., and H. TURNER. 1999. Coding polymorphism in phylogeny reconstruction. *Systematic Biology* **48**:365-379.

KUMAZAWA, Y., and M. NISHIDA. 1999. Complete mitochondrial DNA sequences of the Green Turtle and Blue-Tailed Skink: Statistical evidence for archosaurian affinity of turtles. *Mol Biol Evol* **16**:784-792.

LAKJER, T. 1926. Studien über die Trigeminus-versorgte Kaumuskulatur der Sauropsiden. C. A. Reitzel, Copenhagen.

LANKESTER, E. R. 1870. On the use of the term homology in modern zoology. *Annals and Magazine of Natural History* **6**:34-43.

LAURIN, M., and R. R. REISZ. 1993. The origin of turtles. *Journal of Vertebrate Paleontology* **13**:46A.

LAURIN, M., and R. R. REISZ. 1995. A reevaluation of early amniote phylogeny. *Zoological Journal of the Linnean Society* **113**:165-223.

LE QUESNE, W. J. 1969. A method of selection of characters in numerical taxonomy. *Systematic Zoology* **18**:201-205.

LE QUESNE, W. J. 1972. Further studies based on the uniquely derived character concept. *Systematic Zoology* **21**:281-288.

LE QUESNE, W. J. 1982. Compatibility analysis and its applications. *Zoological Journal of the Linnean Society* **74**:267-275.

LE QUESNE, W. J. 1989. The normal deviate test of phylogenetic value of a data matrix. *Systematic Zoology* **38**:51-54.

LEE, D.-C., and H. N. BRYANT. 1999. A reconsideration of the coding of inapplicable characters: Assumptions and problems. *Cladistics* **15**:373-378.

LEE, M. S. Y. 1993a. The origin of the turtle body plan: Bridging a famous morphological gap. *Science* **261**:1716-1720.

LEE, M. S. Y. 1993b. The origin of turtles reconsidered (again). *Abstract of the Second World Congress of Herpetology*:150.

LEE, M. S. Y. 1994. The turtle's long-lost relatives. *Natural History* **6**:63-65.

LEE, M. S. Y. 1995. Historical burden in systematics and the interrelationships of 'parareptiles'. *Biological Reviews* **70**:459-547.

LEE, M. S. Y. 1996a. Correlated progression and the origin of turtles. *Nature* **379**:812-815.

LEE, M. S. Y. 1996b. The homologies and early evolution of the shoulder girdle in turtles. *Proceedings of the Royal Society of London B* **263**:111-117.

LEE, M. S. Y. 1997a. Pareiasaur phylogeny and the origin of turtles. *Zoological Journal of the Linnean Society* **120**.

LEE, M. S. Y. 1997b. Reptile relationships turn turtle... *Nature* **389**:245-246.

LEE, M. S. Y. 2001. Molecules, morphology, and the monophyly of diapsid reptiles. *Contributions to Zoology* **70**:1-22.

LEWIS, P. O. 2001. A likelihood approach to estimating phylogeny from discrete morphological characters. *Systematic Biology* **50**:913-925.

LONG, R. A., and K. L. BALLEW.1985. Aetosaur dermal armor from the late Triassic of southwestern North America, with special reference to material from the Chinle Formation of Petrified Forest National Park *in* Mus N Ariz Bull. 47:45-68.

LONG, R. A., and P. A. MURRAY. 1995. Late Triassic (Carnian and Norian) Tetrapods from the Southwestern United States. *New Mexicon Museum of Natural History Bulletin* **4**:1-254.

LORENZEN, S. 1993. The role of parsimony, outgroup analysis, and theory of evolution in phylogenetic systematics. *Zeitschrift für Zoologische Systematik und Evolutionsforschung* **31**:1-20.

LØVTRUP, S. 1977. The Phylogeny of Vertebrata. Wiley, London.

LUCAS, S. G., A. B. HECKERT, and A. P. HUNT. 2000. Probable turtle from the Upper Triassic of East-Central New Mexico. *Neues Jahrbuch für Geologie und Paläontologie, Monatshefte* **2000**:287-300.

LYONS-WEILER, J., and G. A. HOELZER. 1997. Escaping from the Felsenstein Zone by detecting long branches in phylogenetic data. *Molecular Phylogenetics and Evolution* **8**:375-384.

MADDISON, W. P. 1990. A method for testing the correlated evolution of two binary characters: Are gains or losses concentrated on certain branches of a phylogenetic tree? *Evolution* **44**:539-557.

MADDISON, W. P. 1993. Missing data versus missing characters in phylogenetic analysis. *Systematic Biology* **42**:576-581.

MADDISON, W. P. 2000. Testing character correlations using pairwise comparisons on a phylogeny. *Journal of Theoretical Biology* **202**:195-204.

MANNEN, H., and S. S.-L. LI. 1999. Molecular evidence for a clade of turtles. *Molecular Phylogenetics and Evolution* **13**:144-148.

MAYR, E. 1982. The growth of biological thought. The Belknap Press of Harvard University Press, Cambridge, Massachusettes.

MCMORRIS, F. R. 1975. Compatibility criteria for cladistic and qualitative taxonomic characters. Pages 399-415 *in* Proceedings of the 8th international Freeman Conference in numerical taxonomy (G. F. Estabrook, ed.) W. H. Freeman, San Francisco.

MEACHAM, C. A. 1984. Evaluating characters by character compatibility analysis. Pages 152-165 *in* Cladistics: perspectives on the reconstruction of evolutionary history (T. Duncan, and T. F. Stuessy, eds.). Columbia University Press, New York.

MEACHAM, C. A. 1994. Phylogenetic relationships at the basal radiation of Angiosperms, further study by probability of character compatibility. *Systematic Botany* **19**:506-522.

MEYER, A. 1995. Molecular evidence on the origin of tetrapods and the relationships of the coelacanth. *Trends in Ecology and Evolution* **10**:111-116.

MEYER, A., and R. ZARDOYA. 2003. Recent advances in the (molecular) phylogeny of vertebrates. *Annual Review of Ecology, Evolution and Systematics* **34**:311-338.

MIKKELSEN, P. M. 1996. Phylogenetics of the Cephalaspidea. *Malacologia* **37**:375-442.

MODESTO, S. P., and R. M. H. SMITH. 2001. A new Late Permian captorhinid reptile: a first record from the South African Karoo. *Journal of Vertebrate Paleontology* **21**:405-409.

MODESTO, S. P., and H.-D. SUES. 2004. The skull of the Early Triassic archosauromorph reptile *Prolacerta broomi* and its phylogenetic significance. *Zoological Journal of the Linnean Society* **140**:335-351.

MOOERS, A. Ø., R. D. M. PAGE, A. PURVIS, and P. H. HARVEY. 1995. Phylogenetic noise leads to unbalanced cladistic tree reconstructions. *Systematic Biology* **44**:332-342.

MOTANI, R., N. MINOURA, and T. ANDO. 1998. Ichthyosaurian relationships illuminated by new primitive skeletons from Japan. *Nature* **393**:255-257.

NAYLOR, G. J. P., and D. C. ADAMS. 2001. Are fossil data really at odds with the molecular data? Morphological evidence for Cetartiodactyla phylogeny reexamined. *Systematic Biology* **50**:444-453.

NELSON, G., and N. I. PLATNICK. 1991. Three-taxon statements: a more precise use of parsimony? *Cladistics* **7**:351-366.

O'KEEFE, F. R., and P. J. WAGNER. 2001. Inferring and testing hypotheses of cladistic character dependence by using character compatibility. *Systematic Biology* **50**:657-675.

O'LEARY, M. A., and J. H. GEISLER. 1999. The position of Cetacea within Mammalia: Phylogenetic analysis of morphological data from extinct and extant taxa. *Systematic Biology* **48**:455-490.

OLSON, E. C. 1947. The family Diadectidae and its bearing on the classification of reptiles. *Fieldiana: Geology* **11**:1-53.

OLSON, E. C. 1965. Relationships of *Seymouria*, *Diadectes* and Chelonia. *American Zoologist* **5**:295-307.

OSBORN, H. F. 1903. The reptilian subclasses Diapsida and Synapsida and the early history of the Diaptosauria. *Memoirs of the American Museum of Natural History* **1**:451-507.

OWEN, R. 1843. Lectures on the comparative anatomy and physiology of invertebrate animals. Longman, Brown, Green, and Longmans, London.

OWEN, R. 1847. Report on the archetype and homologies of the vertebrate skeleton. *Report of the British Association for the Advancement of Science* **16**:169-340.

PAGEL, M. 1994. Detecting correlated evolution on phylogenies: A general method for the comparative analysis of discrete characters. *Proceedings of the Royal Society of London B:* **255**:37-45.

PARRISH, J. M. 1994. Cranial osteology of *Longosuchus meadei* and the phylogeny and distribution of the Aetosauria. *Journal of Vertebrate Paleontology* **14**:196-209.

PATTERSON, C. 1982. Morphological characters and homology. Pages 21-74 *in* Problems of Phylogenetic Reconstruction (K. A. Joysey, and A. E. Friday, eds.). Academic Press, London and New York.

PATTERSON, C. 1988. Homology in classical and molecular biology. *Molecular Biology and Evolution* **5**:603-625.

PENNY, D. 1982. Graph theory, evolutionary trees and classification. *Zoological Journal of the Linnean Society* **74**:305-328.

PENNY, D., and M. D. HENDY. 1985. Testing methods of evolutionary tree construction. *Cladistics* **1**:266-272.

PENNY, D., and M. D. HENDY. 1986. Estimating the reliability of evolutionary trees. *Molecular Biology and Evolution* **3**:403-417.

PIMENTEL, R. A., and R. RIGGINS. 1987. The nature of cladistic data. *Cladistics* **3**:201-209.

PLATNICK, N. I. 1978. Classifications, historical narratives, and hypotheses. *Systematic Zoology* **27**:365-369.

PLATNICK, N. I. 1979. Philosophy and the transformation of cladistics. *Systematic Zoology* **28**:537-546.

PLATNICK, N. I., C. E. GRISWOLD, and J. A. CODDINGTON. 1991. On missing entries in cladistic-analysis. *Cladistics* **7**:337-343.

PLEIJEL, F. 1995. On character coding for phylogeny reconstruction. *Cladistics* **11**:309-315.

POE, S., and J. J. WIENS. 2000. Character selection and the methodology of morphological phylogenetics. Pages 20-36 *in* Phylogenetic analysis of morphological data (J. J. Wiens, ed.) Smithsonian Institution Press, Washington DC.

POPPER, K. R. 1962. The open society and Its enemies. Vol. 2. Hegel and Marx, 4th edition. Routledge & Keegan Paul, London.

PURVIS, A. 1995. A modification to Baum and Regan's method for combining phylogenetic trees. *Systematic Biology* **44**:251-255.

RAE, T. C. 1998. The logical basis for the use of continuous characters in phylogenetic systematics. *Cladistics* **14**:221-228.

RAGAN, M. A. 1992. Phylogenetic inference based on matrix representation of trees. *Molecular Phylogenetics and Evolution* **1**:53-58.

REISZ, R. R., D. S. BERMAN, and D. SCOTT. 1984. The anatomy and relationships of the Lower Permian reptile *Araeoscelis*. *Journal of Vertebrate Paleontology* **4**:57-67.

REISZ, R. R., D. W. DILKES, and D. S. BERMAN. 1998. Anatomy and relationships of *Elliotsmitha longiceps*, a small synapsid. *Journal of Vertebrate Paleontology* **18**:602-611.

REISZ, R. R., and M. LAURIN. 1991. *Owenetta* and the origin of turtles. *Nature* **349**:324-326.

REISZ, R. R., and M. LAURIN. 2001. The reptile *Macroleter*: first vertebrate evidence for correlation of Upper Permian continental strata of North America and Russia. *Geological Society of America Bulletin* **113**:1229-1233.

REISZ, R. R., and D. SCOTT. 2002. *Owenetta kitchingorum*, sp. nov., a small parareptile (Procolophonia: Owenettidae) from the Lower Triassic of South Africa. *Journal of Vertebrate Paleontology* **22**:244-256.

REST, J. S., J. C. AST, C. C. AUSTIN, P. J. WADDEL, E. A. TIBBETTS, J. M. HAY, and D. P. MINDELL. 2003. Molecular systematics of primary reptilian lineages and the tuatara mitochondrial genome. *Molecular Phylogenetics and Evolution* **29**:289-297.

REYNOSO, V.-H. 1998. *Huehuecuetzpalli mixtecus* gen. et sp. nov: a basal squamate (Reptilia) form the Early Cretaceous of Tepexi de Rodríguez, Central México. *Philosophical Transactions of the Royal Society of London* **353**:477-500.

RICHARDS, R. 2003. Character individuation in phylogenetic inference. *Philosophy of Science* **70**:264-279.

RIEPPEL, O. 1980. Homology, a deductive concept? *Zeitschrift für Zoologische Systematik und Evolutionsforschung* **18**:315-319.

RIEPPEL, O. 1990. The structure and development of the jaw adductor musculature in the turtle *Chelydra serpentina*. *Zoological Journal of the Linnean Society* **98**:27-62.

RIEPPEL, O. 1993. Studies on skeleton formation in reptiles: Patterns of ossification in the skeleton of *Chelydra serpentina* (Reptilia, Testudines). *Journal of Zoology, London* **231**:487-509.

RIEPPEL, O. 1994. Osteology of *Simosaurus gaillardoti* and the relationships of stem-group Sauropterygia. *Fieldiana* **28**:1-85.

RIEPPEL, O. 1995. Studies on skeleton formation in reptiles: Implications for turtle relationships. *Zoology* **98**:298-308.

RIEPPEL, O. 1996. Testing homology by congruence: The pectoral girdle of turtles. *Proceedings of the Royal Society of London B* **263**:1395-1398.

RIEPPEL, O. 1998. *Corosaurus alcovensis* Case and the phylogenetic interrelationships of Triassic stem-group Sauropterygia. *Zoological Journal of the Linnean Society* **124**:1-41.

RIEPPEL, O. 2000a. *Paraplacodus* and the phylogeny of the Placodontia (Reptilia: Sauropterygia). *Zoological Journal of the Linnean Society* **130**:635-659.

RIEPPEL, O. 2000b. Turtles as diapsid reptiles. *Zoologica Scripta* **29**:199-212.

RIEPPEL, O., and M. DEBRAGA. 1996. Turtles as diapsid reptiles. *Nature* **384**:453-455.

RIEPPEL, O., and M. KEARNEY. 2001. The origin of snakes: limits of a scientific debate. *Biologist* **48**:110-114.

RIEPPEL, O., and M. KEARNEY. 2002. Similarity. *Biological Journal of the Linnean Society* **75**:59-82.

RIEPPEL, O., and R. R. REISZ. 1999. The origin and early evolution of turtles. *Annual Review of Ecology and Systematics* **30**:1-22.

ROBINSON, P. L. 1962. Gliding lizards from the Upper Keuper beds of Great Britain. *Proceedings of the Geological Society of London* **1601**:137-146.

ROMER, A. S. 1956. Osteology of the Reptiles. University of Chicago Press, Chicago.

ROMER, A. S. 1964. *Diadectes*, an amphibian. *Copeia* **4**:718-719.

ROTH, V. L. 1984. On homology. *Biological Journal of the Linnean Society* **22**:13-29.

ROTH, V. L. 1988. The biological basis of homology. Pages 1-26 *in* Ontogeny and Systematics (C. J. Humphries, ed.) Columbia University Press, New York.

ROUGIER, G. W., M. S. DE LA FUENTE, and A. B. ARCUCCI. 1995. Late Triassic turtles from South America. *Science* **268**:855-858.

ROUGIER, G. W., J. R. WIBLE, and J. A. HOPSON. 1996. Basicranial anatomy of *Priacodon friutaensis* from the Late Jurassic of Colorado, and a reappraisal of mammaliaform interrelationships. *American Museum Novitates* **3183**.

RUBIDGE, B. S., and C. A. SIDOR. 2001. Evolutionary patterns among Permo-Triassic theraspids. *Annual Review of Ecology and Systematics* **32**:449-480.

SEELEY, H. G. 1892. On a new reptile from Welte Vreden (Beaufort West), *Eunotosaurus africanus* (Seeley). *Quarterly Journal of the Geological Society of London* **48**:583-585.

SEN, K. 2003. *Pamelaria dolichotrachela*, a new prolacertid reptile from the Middle Triassic of India. *Journal of Asian Earth Sciences* **21**:663-681.

SENTER, P. 2003. New information on cranial and dental features of the Triassic archosauriform reptile *Euparkeria capensis*. *Palaeontology* **46**:613-621.

SHARKEY, M. J. 1989. A hypothesis independent method of character weighting for cladistic analysis. *Cladistics* **5**:63-86.

SHARKEY, M. J. 1994. Discriminate compatibility measures and the reduction routine. *Systematic Biology* **43**:526-542.

SIDOR, C. A., and R. M. H. SMITH. 2004. A new galesaurid (Therapsida: Cynodontia) from the Lower Triassic of South America. *Palaeontology* **47**:535-556.

SMALL, B. J. 2002. Cranial anatomy of *Desmatosuchus haplocerus* (Reptilia: Archosauria: Stagonolepididae). *Zoological Journal of the Linnean Society* **136**:97-111.

SNEATH, P. H. A., and R. R. SOKAL. 1973. Numerical Taxonomy. W H Freeman, San Francisco.

SOKAL, R. R., and F. J. ROHLF. 1981. Taxonomic congruence in the Leptopodomorpha reexamined. *Systematic Zoology* **30**:309-325.

STEVENS, P. F. 1991. Character states, morphological variation, and phylogenetic analysis: a review. *Systematic Botany* **16**:553-583.

STORRS, G. W. 1991. Anatomy and relationships of Corosaurus alcovensis (Diapsida: Sauropterygia) and the Triassic Alcova Limestone of Wyoming. *Peabody Museum Bulletin* **44**:1-151.

STRONG, E. E., and D. LIPSCOMB. 1999. Character coding and inapplicable data. *Cladistics* **15**:363-371.

SUES, H. D., N. H. SHUBIN, and P. E. OLSEN. 1994. A new sphenodontian from the McCoy Brook Formation (Lower Jurassic) of Nova Scotia, Canada. *Journal of Vertebrate Paleontology* **14**:327-340.

SWOFFORD, D. L. 2003. PAUP*. Phylogenetic analysis using parsimony (*and other methods), version 4.10b. Sinauer Associates, Sunderland, Massachusetts.

SWOFFORD, D. L., and S. POE. 1999. Taxon sampling revisited. *Nature* **398**:299-300.

TEMPLETON, A. R. 1983. Phylogenetic inference from restriction endonuclease clevage site maps with particular reference to the evolution of humans and apes. *Evolution* **37**:221-244.

THIELE, K. 1993. The holy grail of the perfect character: the cladistic treatment of morphometric data. *Cladistics* **9**:275-304.

THORLEY, J. L., and R. D. M. PAGE. 2000. RadCon: Phylogenetic tree comparison and consensus. *Bioinformatics* **16**:486-487.

THORLEY, J. L., and M. WILKINSON. 1999. Testing the phylogenetic stability of early tetrapods. *Journal of Theoretical Biology* **200**:343-344.

THORLEY, J. L., and M. WILKINSON. 2003. A view of supertree methods. Pages 185-193 *in* Bioconsensus (M. Janowitz, F.-J. Lapointe, F. R. McMorris, B. Mirkin, and F. S. Roberts, eds.). DIMACS series in discrete mathematics and theoretical computer science, American Mathematical Society, Providence.

THORLEY, J. L., M. WILKINSON, and M. A. CHARLESTON. 1998. The information content of consensus trees. Pages 91-98 *in* Advances in data science and classification (A. Rizi, Vichi M, and H.-H. Bock, eds.). Springer, Berlin.

TONG, H., E. BUFFETAUT, and V. SUTEETHORN. 2002. Middle Jurassic turtles from southern Thailand. *Geological Magazine* **139**:687-697.

TRUEMAN, J. W. H. 1998. Reverse successive weighting. *Systematic Biology* **47**:733-737.

VAN VALEN, L. M. 1982. Homology and causes. *Journal of Morphology* **173**:305-312.

VANE-WRIGHT, R. I., S. SCHULTZ, and M. BOPPRE. 1992. The cladistics of Amauris butterflies: congruence, consensus and total evidence. *Cladistics* **8**:125-138.

WAGNER, P. J. 1998. A likelihood approach for estimating phylogenetic relationships among fossil taxa. *Paleobiology* **24**:430-449.

WAGNER, P. J. 2001. The character concept in evolutionary biology. Academic Press, San Diego, California.

WAGNER, P. J., and B. Y. MISOF. 1993. How can a character be developmentally constrained despite variation in developmental pathways? *Journal of Evolutionary Biology* **6**:449-455.

WALKER, A. D. 1961. Triassic reptiles from the Elgin area: *Stagonolepis*, *Dasygnathus* and their allies. *Philosophical Transactions of the Royal Society of London B* **244**:103-204.

WATSON, D. M. S. 1914. *Eunotosaurus africanus* Seeley, and the ancestry of the Chelonia. *Proceedings of the Zoological Society of London* **1**:1011-1020.

WATSON, D. M. S. 1941. On the skull of a pariasaurian reptile, and on the relationships of that type. *Proceedings of the Zoological Society of London* **1**:155-180.

WHEELER, W. C., P. CARTWRIGHT, and C. Y. HAYASHI. 1993. Arthropod phylogeny: a combined approach. *Cladistics* **9**:1-39.

WHITING, M. F., S. BRADLER, and T. MAXWELL. 2003. Loss and recovery of wings in stick insects. *Nature* **421**:264-267.

WIENS, J. J. 1995. Polymorphic characters in phylogenetic systematics. *Systematic Biology* **44**:482-500.

WIENS, J. J. 2000. Phylogenetic analysis of morphological data. Smithsonian Institution Press, Washington DC.

WIENS, J. J. 2001. Character analysis in morphological phylogenetics: Problems and solutions. *Systematic Biology* **50**:689-699.

WIENS, J. J., and D. M. HILLIS. 1996. Accuracy of parsimony analysis using morphological data: a repraisal. *Systematic Botany* **21**:237-243.

WILEY, E. O. 1975. Karl R. Popper, systematics, and classification: a reply to Walter Bock and other evolutionary taxonomists. *Systematic Zoology* **24**:233-243.

WILEY, E. O. 1981. Phylogenetics. John Wiley and Sons, New York.

WILKINSON, M. 1992a. Consensus, compatibility and missing data in phylogenetic inference. Ph.D. Thesis, University of Bristol.

WILKINSON, M. 1992b. Ordered versus unordered characters. *Cladistics* **8**:375-385.

WILKINSON, M. 1994a. Common cladistic information and its consensus representation: Reduced Adams and reduced cladistic consensus trees and profiles. *Systematic Biology* **43**:343-368.

WILKINSON, M. 1994b. The permutation method and character compatibility. *Systematic Biology* **43**:274-277.

WILKINSON, M. 1994c. Three-taxon statements: When is parsimony analysis also a clique analysis? *Cladistics* **10**:221-223.

WILKINSON, M. 1995a. A comparison of two methods of character construction. *Cladistics* **11**:297-308.

WILKINSON, M. 1995b. Coping with abundant missing entries in phylogenetic inference using parsimony. *Systematic Biology* **44**:501-514.

WILKINSON, M. 1996. Majority-rule reduced consensus methods and their use in bootstrapping. *Molecular Biology and Evolution* **13**:437-444.

WILKINSON, M. 1997a. Characters, congruence and quality: A study of neuroanatomical and traditional data in caecilian phylogeny. *Biological Reviews of the Cambridge Philosophical Society* **72**:423-470.

WILKINSON, M. 1997b. Limitations on the use of compatibility methods for polarizing ordered characters. *Systematic Biology* **46**:230-231.

WILKINSON, M. 2001a. PICA software and documentation. School of Biological Sciences, University of Bristol. Bristol

WILKINSON, M. 2001b. TAXEQ3: software and documentation. Department of Zoology, The Natural History Museum. London

WILKINSON, M., and J. L. THORLEY. 2001a. The efficiency of strict consensus trees. *Systematic Biology*.

WILKINSON, M., and J. L. THORLEY. 2001b. No compromise on consensus. *Taxon* **50**:181-184.

WILKINSON, M., and J. L. THORLEY. 2003. Reduced consensus. Pages 195-204 *in* Bioconsensus (M. Janowitz, F.-J. Lapointe, F. R. McMorris, B. Mirkin, and F. Roberts, eds.). American Mathematical Society, Providence.

WILKINSON, M., J. L. THORLEY, and M. J. BENTON. 1997. Uncertain turtle relationships. *Nature* **387**:466.

WILKINSON, M., J. L. THORLEY, D. T. J. LITTLEWOOD, and R. A. BRAY. 2001. Towards a phylogenetic supertree for the Platyhelminthes? Pages 292-301 *in* Interrelationships of the Platyhelminthes (D. T. J. Littlewood, and R. A. Bray, eds.). Chapman-Hall, London.

WILKINSON, M., J. L. THORLEY, and P. UPCHURCH. 2000. A chain is no stronger than its weakest link: Double decay analysis for phylogenetic hypotheses. *Systematic Biology* **49**:754-776.

WILLISTON, S. W. 1917. The phylogeny and classfication of reptiles. *Journal of Geology* **25**:411-421.

WILSON, E. O. 1965. A consistency test for phylogenies based on contemporaneous species. *Systematic Zoology* **14**:214-220.

WU, X. C., H. D. SUES, and Z. M. DONG. 1997. *Sichuanosuchus shuhanensis*, a new ?Early Cretaceous protosuchian from Sichuan (China) and the monophyly of the protosuchia. *Journal of Vertebrate Paleontology* **17**:89-103.

YEATES, D. K. 1995. Groundplans and exemplars: Paths to the tree of life. *Cladistics* **11**:343-357.

ZARDOYA, R., and A. MEYER. 1998. Complete mitochondrial genome sequence suggests diapsid affinities of turtles. *Proceedings of the National Academy of the United States of America* **95**:14226-14231.

ZARDOYA, R., and A. MEYER. 2001. The evolutionary position of turtles revised. *Naturwissenschaften* **88**:193-200.

Appendix 1

References Used in the Survey of Character Construction Methods

References are listed in the order they appear in the results, under their code number (see Appendix 2). E = Extant taxa i.e. analysis of neontological data; F = Fossil taxa i.e. analysis of palaeontological data.

E1) GAUTHIER, J., A. G. KLUGE, and T. ROWE. 1988. Amniote phylogeny and the importance of fossils. *Cladistics* **4**:105-209.

E2) BELLWOOD, O. 1998. The phylogeny of the box-crab genera with notes on their fossil record, biogeography and depth distribution. *Journal of the Zoological Society of London* **244**:459-471.

E3) CARLSON, S. J. 1995. Phylogenetic relationships of amongst extant brachiopods. *Cladistics* **11**:13-197.

E4) WHEELER, W. C., P. CARTWRIGHT, and C. Y. HAYASHI. 1993. Arthropod phylogeny: a combined approach. *Cladistics* **9**:1-39.

E5) MILLER, J. S. 1996. Phylogeny of the neotropical moth tribe Josiini: a hidden case of Mullerian mimicry. *Zoological Journal of the Linnean Society* **118**:1-45.

E6) SIVALDADOTTIR, E., A. S. Y. MACKIE, and F. PLEIJEL. 1997. Generic interrelations within the Sponidae. *Zoological Journal of the Linnean Society* **473-500**.

E7) OOSTERBROEK, P., and G. CORTNEY. 1995. Phylogeny of the nematoceras families of Diptera. *Zoological Journal of the Linnean Society* **115**:267-311.

E8) HALLERMANN, J. 1998. The ethmoidal region of *Dibamus taylori* with a phylogenetic hypothesis on dibamid relationships within the Squamata. *Zoological Journal of the Linnean Society* **122**:385-426.

E9) SCHULTZ, J. W. 1990. Evolutionary morphology and phylogeny of the Arachnida. *Cladistics* **6**:1-38.

E10) SUNDBERG, P. 1989. Phylogeny and classification of the paramonostiliferous family Plectonemertidae. *Cladistics* **5**:87-100.

E11) DOMÍNGUEZ, E. 1999. Systematics, cladistics and biogeography of the American genus Farrodes (Ephemeroptera : Leptophlebiidae : Atalophlebiinae). *Zoological Journal of the Linnean Society* **126**:155-189.

E12) VANE-WRIGHT, R. I., S. SCHULTZ, and M. BOPPRE. 1992. The cladistics of Amauris butterflies: congruence, consensus and total evidence. *Cladistics* **8**:125-138.

E13) LADIGE, P. Y., S. M. PROBER, and G. NELSON. 1992. Cladistics and biogeographic analysis of the 'Blue Ash' eucalypts. *Cladistics* **8**:103-124.

E14) SMITH, A. B., G. L. J. PATTERSON, and B. LAFAY. 1995. Ophiuroid phylogeny and higher taxonomy: morphological, molecular and palaeontological perspectives. *Zoological Journal of the Linnean Society* **114**:213-243.

E15) MCHUGH, D. 1995. Phylogenetic analysis of the Amphinitritinae. *Zoological Journal of the Linnean Society* **114**:405-429.

E16) KLUGE, A. G. 1993. *Calabria* and the phylogeny of erycine snakes. *Zoological Journal of the Linnean Society* **107**:293-351.

E17) PLEIJEL, F. 1993. Phylogeny of *Phyllodoce* (Polychaeta, Phyllodocidae). *Zoological Journal of the Linnean Society* **108**:287-299.

E18) BRYANT, H. N., A. P. RUSSELL, and W. D. FITCH. 1993. Phylogenetic relationships within the extant Mustelidae: appraisal of the cladistic status of the Simpsonian subfamilies. *Zoological Journal of the Linnean Society* **108**:301-334.

E19) WHITING, M. F., J. C. CARPENTER, Q. D. WHEELER, and W. C. WHEELER. 1997. The Strepsiptera: phylogeny of the Holometabolous insect orders inferred from 18S and 28S ribosomal DNA sequences and morphology. *Systematic Biology* **46**:1-68.

E20) SHAFFER, H. B., P. A. MEYLAN, and M. L. MCKNIGHT. 1997. Tests of turtle phylogeny: molecular, morphological and paleontological approaches. *Systematic Biology* **46**:235-268.

E21) BROCHU, C. A. 1997. Morphology, fossils, divergence timing, and the phylogenetic relationships of *Gavialis*. *Systematic Biology* **46**:479-522.

E22) LIVEZEY, B. C. 1996. A phylogenetic analysis of geese and swans including selected fossil species. *Systematic Biology* **45**:415-450.

E23) LAFAY, B., A. B. SMITH, and R. CHRISTEN. 1995. A combined morphological and molecular approach to the phylogeny of asteroids. *Systematic Biology* **44**:190-208.

E24) LYDEARD, C., M. C. WOOTEN, and A. MEYER. 1995. Molecules, morphology and area cladograms: a cladistic biogeographic analysis of *Gambusia*. *Systematic Biology* **44**:221-236.

E25) TITUS, T. A., and A. LARSON. 1995. A molecular phylogenetic perspective on the evolutionary radiation of the salamander family Salamandridae. *Systematic Biology* **44**:125-151.

E26) MIKKELSEN, P. M. 1996. Phylogenetics of the Cephalaspidea. *Malacologia* **37**:375-442.

E27) BURKE, R. L., T. E. LEUTERITZ, and A. J. WOLF. 1996. Phylogenetic relationships of emydine turtles. *Herpetologica* **52**:572-584.

E28) GREER, A. E. 1997. *Leptoseps*: A new genus of Scincid lizards from southeast Asia. *Journal of Herpetology* **31**:393-398.

E29) MORRONE, J. J., and R. S. ANDERSON. 1995. The *Falklandius* generic group : cladistic analysis with description of new taxa (Coleoptera: Curculionidae: Rhytirrhinini). *American Museum Novitates* **3121**.

E30) MORRONE, J. J. 1994. Systematics, cladistics and biogeography of the Andean weevil genera *Macrostyphlus*, *Adioristidius*, *Puranius*, and *Amathynetoides*, new genus (Coleoptera: Curculionidae). *American Museum Novitates* **3104**.

E31) ROIG-JUNENT, S. 1995. Cladistic analysis of Barypus Dejean 1828 (Coleoptera: Carabidae: Broscini). *American Museum Novitates* **3117**.

E32) ROZEN, J. G. 1996. Phylogenetic analysis of the cleptoparasitic bees belonging to the nomadinae based on mature larvae (Apoidea: Apidae). *American Museum Novitates* **3180**.

E33) EERNISSE, D. J., J. S. ALBERT, and F. E. ANDERSON. 1992. Annelida and Arthropoda are not sister taxa - a phylogenetic analysis of spiralian metazoan morphology. *Systematic Biology* **41**:305-330.

E34) GOULD, G. C. 1995. Hedgehog phylogeny - the reciprocal illumination of the quick and the dead. *American Museum Novitates* **3131**.

E35) HINX, P. 1996. Phylogenetic relationships amongst box turtles, *Terrapene*. *Herpetologica* **52**:584-597.

E36) WYSS, A. R., and J. F. FLYNN. 1993. A phylogenetic analysis and definition of the Carnivora. Pages 32-52 *in* Mammal Phylogeny: Mesozoic Differentiation, Multituberculates, Monotremes, Early Therians and Marsupials (F. S. Szalay, M. J. Novacek, and M. C. McKenna, eds.). Springer-Verlag, New York.

E37) WIBLE, J. R., and J. A. HOPSON. 1993. Basicranial evidence for early mammal phylogeny. Pages 45-62 *in* Mammal Phylogeny: Mesozoic Differentiation,

Multituberculates, Monotremes, Early Therians and Marsupials (F. S. Szalay, M. J. Novacek, and M. C. McKenna, eds.). Springer-Verlag, New York.

E38) PRESCH, W. 1988. Phylogenetic relationships of the Scincomorpha. Pages 471-492 *in* Phylogenetic Relationships of the Lizard Families (R. Estes, and G. Pregill, eds.). Stanford University Press, Stanford, California.

E39) GRISMER, L. L. 1988. Phylogeny, taxonomy, classification, and biogeography of Eublepharid Geckos. Pages 369-469 *in* Phylogenetic Relationships of the Lizard Families (R. Estes, and G. Pregill, eds.). Stanford University Press, Stanford, California.

E40) SIMMONS, N. B., and J. H. GEISLER. 1998. Phylogenetic relationships of *Icaronycteris*, *Archeonycteris*, *Hassianycteris*, and *Palaeochiropteryx* to extant bat lineages, with comments on the evolution of echolocation and foraging strategies in microchiroptera. *Bulletin of the American Museum of Natural History* **235**:1-182.

E41) YEATES, D. K. 1994. The cladistics and classification of the Bombyliidae (Diptera: Asiloidea). *Bulletin of the American Museum of Natural History* **219**:1-191.

E42) ALBERT, J. S., and W. L. FINK. 1996. *Sternopygus xingu*, a new species of electric fish from Brazil (Teleostei: Gymnotoidei), with comments on the phylogenetic position of *Sternopygus*. *Copeia* **1**:85-102.

E43) MAYDEN, R. L., and B. R. KUHAJDA. 1996. Systematics, taxonomy and conservation status of the endangered Alabama sturgeon *Scaphirhynchus suttkusi* Williams and Clemmer (Actinopterygii, Acipenseridae). *Copeia* **2**:241-273.

E44) WILKINSON, M., and R. A. NUSSBAUM. 1996. On the phylogenetic position of the Uraeotyphlidae (Amphibia: Gymnophiona). *Copeia* **3**:550-562.

E45) GOSLINER, T. M., and R. F. JOHNSON. 1999. Phylogeny of *Hypselodoris* (Nudibranchia: Chromodorididae) with a review of the monophyletic clade of Indo-Pacific species, including descriptions of twelve new species. *Zoological Journal of the Linnean Society* **125**:1-114.

E46) PLEIJEL, F., and T. DAHLGREN. 1998. Position and delineation of Chrysopetalidae and Hesionidae (Annelida, Polychaeta). *Cladistics* **14**:129-150.

E47) CIGLIANO, M. M., R. A. RONDEROS, and W. P. KEMP. 1996. Phylogenetic relationships of *Scotussa* and *Leiotettix* Orthoptera: Acrididae). *Cladistics* **12**:125-138.

E48) GUISU, R. C., and R. WINTERBOTTOM. 1993. Osteological evidence for the phylogeny of the recent genera of sturgeonfishes. *Copeia* **2**:300-312.

E49) WIENS, J. J. 1993. Phylogenetic relationships of Phrynosomatid lizards and monophyly of the *Sceloporus* group. *Copeia* **2**:287-299.

E50) HAROLD, A. S. 1993. Phylogenetic relationships of the Sternoptychid *Argyropelecus* (Teleostei: Stomiiformes). *Copeia* **1**:123-133.

F1) WU, X. C., and S. CHATTERJEE. 1993. *Dibothrosuchus elaphros*, a crocodylomorph from the Lower Jurassic of China, and the phylogeny of the Sphenosuchia. *Journal of Vertebrate Paleontology* **13**:58-89.

F2) NOVAS, F. E. 1993. New information on the systematics and postcranial skeleton of *Herrerasaurus ischigualensis* from the Ischigualasto Formation of Argetina. *Journal of Vertebrate Paleontology* **13**:400-423.

F3) DEBRAGA, M., and R. R. REISZ. 1995. A new diapsid reptile from the uppermost Carboniferous of Kansas. *Palaeontology* **38**:199-212.

F4) SUES, H. D., N. H. SHUBIN, and P. E. OLSEN. 1994. A new sphenodontian from the McCoy Brook Formation (Lower Jurassic) of Nova Scotia, Canada. *Journal of Vertebrate Paleontology* **14**:327-340.

F5) CHIAPPE, L. M., and J. O. CALVO. 1994. *Neuquenornis volans*, a new Late Cretaceous bird from Patagonia, Argentina. *Journal of Vertebrate Paleontology* **14**.

F6) DILKES, D. W. 1995. The rhynchosaur *Howesia* from the Lower Triassic of South Africa. *Palaeontology* **38**:665-685.

F7) DODICK, J. T., and S. P. MODESTO. 1995. The cranial anatomy of the captorhinid reptile *Labidosaurikos meachami* from the Lower Permian of Oklahoma. *Palaeontology* **38**:687-711.

F8) LI, G. Q., and M. V. H. WILSON. 1996. The discovery of Heterodontinae from the Paleocene Paskapo Formation of Alberta, Canada. *Journal of Vertebrate Paleontology* **16**:198-209.

F9) SERENO, P. C. 1991. Basal archosaurs: Phylogenetic relationships and functional implications. *Journal of Vertebrate Paleontology* **11**:Supp.

F10) LEE, M. S. Y. 1997. Pareiasaur phylogeny and the origin of turtles. *Zoological Journal of the Linnean Society* **120**.

F11) DILKES, D. W. 1998. The early Triassic rhynchosaur *Mesosuchus brownii* and interrelationships of basal archosauromorph reptiles. *Philosophical Transactions of the Royal Society of London* **353**:501-541.

F12) ERICSON, P. G. P. 1997. Systematic relationships of the palaeogene family Presbyornithidae. *Zoological Journal of the Linnean Society* **121**:429-483.

F13) YATES, A. M. 1999. The Lappilopsidae: a new family of small temnospondyls from the Early Triassic of Australia. *Journal of Vertebrate Paleontology* **19**:302-320.

F14) BAEZ, A. M., and L. A. PUGENER. 1998. a new Palaeogene Pipid frog from north-west Patagonia. *Journal of Vertebrate Paleontology* **18**:511-524.

F15) STEPPAN, S. J., and U. F. J. PARDINAS. 1998. Two new fossil muroids from the early Pleistocene of Argentina: phylogeny and palaeoecology. *Journal of Vertebrate Paleontology* **18**:640-649.

F16) REISZ, R. R., D. W. DILKES, and D. S. BERMAN. 1998. Anatomy and relationships of *Elliotsmitha longiceps*, a small synapsid. *Journal of Vertebrate Paleontology* **18**:602-611.

F17) GAFFNEY, E. S., P. A. MEYLAN, and A. R. WYSS. 1991. A computer assisted analysis of the relationships of the higher categories of turtles. *Cladistics* **7**:313-335.

F18) LAURIN, M., and R. R. REISZ. 1995. A reevaluation of early amniote phylogeny. *Zoological Journal of the Linnean Society* **113**:165-223.

F19) ZHANG, J. Y. 1998. Morphology and phylogenetic relationships of *Kuntulunia.* *Journal of Vertebrate Paleontology* **18**:280-300.

F20) CORIA, R. A. 1996. A basal iguanodontian from the Late Cretaceous of South America. *Journal of Vertebrate Paleontology* **16**:445-457.

F21) STEPPAN, S. J. 1996. A new species of *Holochilus* from the Middle Pleistocene of Bolivia, and its phylogenetic significance. *Journal of Vertebrate Paleontology* **16**:522-530.

F22) HOOKS, G. E. 1998. Systematic revision of the Protostgeidae, with a redescription of *Calcarichelys gemma* Zangerl, 1953. *Journal of Vertebrate Paleontology* **18**:85-98.

F23) WILSON, J. A., and P. C. SERENO. 1998. Early evolution and higher level phylogeny of Sauropod dinosaurs. *Journal of Vertebrate Paleontology* **18**:1-67.

F24) SUES, H. D. 1997. On *Chirostenotes*, a Late Cretaceous oviraptor from western North America. *Journal of Vertebrate Paleontology* **17**:698-716.

F25) JALIL, N. E. 1997. A new prolacertiforme diapsid from the Triassic of North Africa and the interrelationships of the Prolacertiformes. *Journal of Vertebrate Paleontology* **17**:506-525.

F26) GERAADS, D., and E. GULEC. 1997. Relationships of *Barbourofelis piveteaui*, a Late Miocene nimravid from central Turkey. *Journal of Vertebrate Paleontology* **17**:370-375.

F27) GOWER, D. J., and A. G. SENNIKOV. 1997. *Sarmatosuchus* and the early history of the Archosauria. *Journal of Vertebrate Paleontology* **17**:60-73.

F28) WU, X. C., H. D. SUES, and Z. M. DONG. 1997. *Sichuanosuchus shuhanensis*, a new ?Early Cretaceous protosuchian from Sichuan (China) and the monophyly of the protosuchia. *Journal of Vertebrate Paleontology* **17**:89-103.

F29) NOVAS, F. E. 1997. Anatomy of *Patagonykus puertai* from the Late Cretaceous of Patagonia. *Journal of Vertebrate Paleontology* **17**:137-166.

F30) MEYLAN, P. A. 1996. Skeletal morphology and the relationships of the Early Cretaceous side necked turtle *Araripemys barretoi* from the Santana formations of Brazil. *Journal of Vertebrate Paleontology* **16**:20-33.

F31) SCHULTZE, H. 1994. Comparison of hypotheses on the relationships of sarcopterygians. *Systematic Biology* **43**:155-173.

F32) ALROY, J. 1995. Continuous track analysis: a new phylogenetic and biogeographic method. *Systematic Biology* **44**:152-178.

F33) PRADO, J., and M. T. ALBERDI. 1996. A cladistic analysis of the hors tribe Equini. *Palaeontology* **39**:663-680.

F34) WU, X. C., D. B. BRINKMAN, and A. P. RUSSELL. 1996. A new alligator from the Upper Cretaceous of Canada and the relationships of early eusuchians. *Palaeontology* **39**:351-376.

F35) JANVIER, P. 1996. The dawn of vertebrates: characters versus common ascent in the rise of current vertebrate phylogenies. *Palaeontology* **39**:259-288.

F36) RUTA, M., and J. R. THERON. 1997. Two Devonian mitrates from South Africa. *Palaeontology* **40**:201-244.

F37) ROUGIER, G. W., J. R. WIBLE, and J. A. HOPSON. 1996. Basicranial anatomy of *Priacodon friutaensis* from the Late Jurassic of Colorado, and a reappraisal of mammaliaform interrelationships. *American Museum Novitates* **3183**.

F38) MACPHEE, R. D. E., I. HOROVITZ, O. ARREDONDO, and J. V. OSVALDO. 1995. A new genus for the extinct Hispaniolan monkey Saimiri berensis with notes on its systematic position. *American Museum Novitates* **1314**.

F39) WANG, Y., W. A. CLEMENS, Y. HU, and C. LI. 1998. A probable pseudo-tribosphenic upper molar from the Late Jurassic of China, and the early radiation of the holotheria. *Journal of Vertebrate Paleontology* **18**:777-787.

F40) HEAD, J. J. 1998. A new species of basal hadrosaurid from the Cenomanian of Texas. *Journal of Vertebrate Paleontology* **18**:718-738.

F41) MAKOVICKY, P. J., and H. D. SUES. 1998. Anatomy and phylogenetic relationships of the theropod dinosaur Microvenator celer from the Lower Cretaceous of Montana. *American Museum Novitates* **3240**.

F42) FISCHER, M. S., and P. TASSAY. 1993. The interrelationships between Proboscidea, Sirenia, Hyracoidea and Mesaxonia: the morphological evidence. Pages 217-234 *in* Mammal Phylogeny: Placentals (F. S. Szalay, M. J. Novacek, and M. C. McKenna, eds.). springer-Verlag, New York.

F43) BEARD, K. C. 1993. Phylogenetic systematics of the Primatomorpha, with special reference to Dermoptera. Pages 129-150 *in* Mammal Phylogeny: Placentals (F. S. Szalay, M. J. Novacek, and M. C. McKenna, eds.). Springer-Verlag, New York.

F44) CIFELLI, R. L. 1993. The phylogeny of native South American ungulates. Pages 195-216 *in* Mammal Phylogeny: Placentals (F. S. Szalay, M. J. Novacek, and M. C. McKenna, eds.). Springer-Verlag, New York.

F45) GAFFNEY, E. S. 1996. The postcranial morphology of *Meiolanaia platyceps* and a review of the Meiolaniidae. *Bulletin of the American Museum of Natural History* **229**:1-165.

F46) MACPHEE, R. D. E. 1994. Morphology, adaptations and relationships of Plesioryctus, and a diagnosis of a new order of Eutherian mammals. *Bulletin of the American Museum of Natural History* **220**:1-214.

F47) WILLIAMS, A., L. E. POPOV, L. E. HOLMER, and M. CUSACK. 1998. The diversity and phylogeny of the paterinate brachiopods. *Palaeontology* **41**:221-263.

F48) FROELICH, D. J. 1999. Phylogenetic systematics of basal perissodactyls. *Journal of Vertebrate Paleontology* **19**:140-159.

F49) HECKERT, A. B., and S. G. LUCAS. 1999. A new aetosaur from the upper Triassic of Texas and the phylogeny of aetosaurs. *Journal of Vertebrate Paleontology* **19**:50-68.

F50) WILSON, M. V. H., and G. Q. LI. 1999. Osteology and systematic position of the Eocene salmonid *Eosalmo driftwoodensis* from western North America. *Zoological Journal of the Linnean Society* **125**:279-311.

Appendix 2

Results of the survey of character coding methods. Appendix 2a shows the number of each type of unconventional character present in each study. Appendix 2b shows the proportion of characters in each study that are of each unconventional type

Appendix 2a

	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10
Total Number of Characters	**207**	**55**	**112**	**-**	**86**	**25**	**98**	**144**	**64**	**38**
Behavioural*	0	0	0	-	0	0	0	0	0	0
Composite	5	0	3	-	9	1	0	3	0	0
Conjunction	0	0	0	-	0	0	0	0	0	0
Developmental*	0	0	0	-	0	0	0	0	0	0
Extent	4	0	1	-	0	0	0	7	0	0
Inapplicable Data (missing)	10	1	7	-	1	2	2	9	0	2
(result)	16	1	10	-	1	6	7	12	0	2
Inapplicable Data (multistate)	5	3	19	-	8	3	3	6	8	1
(cryptic)	0	0	2	-	6	0	0	0	1	1
Landmark*	0	0	0	-	0	0	0	0	0	0
Logically Related	0	0	0	-	0	0	1	0	1	0
Nominal Variable	0	0	0	-	0	0	0	0	0	0
Positional	1	0	2	-	0	0	6	6	2	0
Ratio	9	5	2	-	6	0	0	2	0	1
Repetition	2	0	0	-	0	0	0	1	0	0
Single Structure Ratio	0	2	2	-	0	0	0	0	0	0
Unifying	0	1	4	-	3	0	4	7	3	0
Unspecified Homologue	0	1	2	-	1	0	0	0	0	1
Total Unconventional	**34**	**11**	**34**	**-**	**23**	**5**	**14**	**38**	**9**	**5**

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
Total Number of Characters	**65**	**51**	**50**	**34**	**73**	**28**	**39**	**28**	**36**	**128**
Behavioural*	0	0	0	0	0	0	0	0	0	0
Composite	0	4	1	0	0	0	3	1	0	5
Conjunction	0	0	0	0	0	0	0	0	0	0
Developmental*	0	0	0	0	0	0	0	0	0	0
Extent	2	3	3	1	1	0	1	0	0	3
Inapplicable Data (missing)	2	0	2	0	0	0	0	0	0	7
(result)	2	0	2	0	0	0	0	0	0	7
Inapplicable Data (multistate)	6	0	1	0	1	0	2	1	1	15
(cryptic)	2	0	0	0	0	0	0	0	1	4
Landmark*	0	0	0	0	0	0	0	0	0	0
Logically Related	0	0	0	0	1	0	0	0	0	1
Nominal Variable	0	0	0	0	0	0	0	0	0	0
Positional	0	1	0	0	3	1	4	0	0	5
Ratio	2	10	9	3	10	5	8	2	10	2
Repetition	0	0	0	0	0	0	0	0	0	0
Single Structure Ratio	2	1	3	0	2	2	0	1	1	3
Unifying	0	2	0	2	0	0	0	0	3	1
Unspecified Homologue	0	0	0	0	2	0	0	0	0	0
Total Unconventional	**13**	**20**	**18**	**6**	**12**	**8**	**15**	**5**	**15**	**35**

	E11	E12	E13	E14	E15	E16	E17	E18	E19	E20
Total Number of Characters	**27**	**-**	**36**	**43**	**22**	**75**	**23**	**46**	**176**	**115**
Behavioural*	0	-	0	0	0	1	0	0	3	0
Composite	3	-	0	4	1	5	0	1	3	0
Conjunction	0	-	0	0	0	0	1	0	0	0
Developmental*	0	-	0	0	0	0	0	0	0	0
Extent	0	-	0	2	2	0	0	1	0	4
Inapplicable Data (missing)	0	-	1	4	1	2	0	0	0	1
(result)	0	-	1	6	3	2	0	0	0	1
Inapplicable Data (multistate)	10	-	2	8	5	11	3	4	23	2
(cryptic)	6	-	0	5	1	3	3	0	4	0
Landmark*	0	-	0	0	0	0	0	0	0	0
Logically Related	0	-	0	0	0	0	0	1	0	0
Nominal Variable	0	-	0	0	0	0	0	0	0	0
Positional	3	-	0	1	1	2	0	4	1	7
Ratio	1	-	0	0	0	1	0	2	0	4
Repetition	0	-	0	0	0	0	0	0	0	0
Single Structure Ratio	2	-	1	0	0	0	0	1	0	0
Unifying	1	-	1	2	3	5	0	2	3	8
Unspecified Homologue	0	-	0	0	0	0	0	0	0	0
Total Unconventional	**18**	**-**	**4**	**17**	**9**	**25**	**4**	**15**	**30**	**26**

	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20
Total Number of Characters	**144**	**50**	**68**	**47**	**45**	**31**	**39**	**124**	**67**	**40**
Behavioural*	0	0	0	0	0	0	0	0	0	0
Composite	1	2	2	1	2	3	0	10	1	0
Conjunction	0	0	0	0	0	0	0	2	0	0
Developmental*	0	0	0	0	0	0	0	0	0	0
Extent	0	6	6	2	3	0	1	6	0	0
Inapplicable Data (missing)	0	0	0	0	0	0	1	3	1	0
(result)	0	0	0	0	0	0	1	3	4	0
Inapplicable Data (multistate)	14	3	2	6	18	3	0	22	9	0
(cryptic)	2	0	1	1	0	2	0	4	1	0
Landmark*	0	0	0	0	0	0	0	0	0	0
Logically Related	0	0	0	0	0	0	2	0	0	0
Nominal Variable	0	0	0	0	0	0	1	0	0	0
Positional	0	2	2	2	0	2	2	4	4	2
Ratio	10	1	5	4	5	4	0	5	1	3
Repetition	0	0	0	0	0	0	0	0	0	0
Single Structure Ratio	6	0	2	1	0	1	0	0	0	0
Unifying	0	2	2	2	6	2	4	3	0	1
Unspecified Homologue	0	0	0	0	0	0	0	0	1	0
Total Unconventional	**33**	**15**	**19**	**16**	**23**	**12**	**10**	**47**	**16**	**6**

	E21	E22	E23	E24	E25	E26	E27	E28	E29	E30
Total Number of Characters	**164**	**175**	**43**	**55**	**48**	**-**	**38**	**41**	**20**	**37**
Behavioural*	0	9	0	0	1	-	5	0	0	0
Composite	6	0	5	0	1	-	0	0	0	0
Conjunction	0	0	0	0	0	-	0	0	0	0
Developmental*	0	0	0	0	0	-	0	0	0	0
Extent	12	8	0	0	3	-	1	0	0	1
Inapplicable Data (missing)	1	1	2	0	0	-	0	0	0	0
(result)	1	1	3	0	0	-	0	0	0	0
Inapplicable Data (multistate)	15	19	2	3	10	-	2	4	4	9
(cryptic)	9	2	0	0	1	-	0	1	2	2
Landmark*	0	0	0	0	0	-	9	0	0	0
Logically Related	1	2	0	1	0	-	0	0	0	0
Nominal Variable	0	0	0	0	0	-	0	0	0	0
Positional	18	3	2	1	1	-	1	1	0	0
Ratio	5	0	0	0	1	-	1	0	1	2
Repetition	0	0	0	0	0	-	0	0	0	0
Single Structure Ratio	1	0	0	0	0	-	0	0	0	0
Unifying	8	5	3	0	0	-	3	0	0	0
Unspecified Homologue	0	0	0	0	0	-	0	0	0	0
Total Unconventional	**63**	**45**	**13**	**5**	**17**	**-**	**22**	**5**	**5**	**12**

	F21	F22	F23	F24	F25	F26	F27	F28	F29	F30
Total Number of Characters	**27**	**27**	**109**	**47**	**71**	**22**	**35**	**128**	**75**	**35**
Behavioural*	0	0	0	0	0	0	0	0	0	0
Composite	3	1	4	0	2	1	0	4	3	0
Conjunction	0	0	0	0	0	0	0	0	0	0
Developmental*	0	0	0	0	0	0	0	0	0	0
Extent	4	1	2	0	1	0	0	2	0	0
Inapplicable Data (missing)	0	0	2	0	2	0	0	1	1	0
(result)	0	0	2	0	3	0	0	1	1	0
Inapplicable Data (multistate)	4	0	1	0	0	1	1	12	0	6
(cryptic)	0	0	1	0	0	0	0	9	0	2
Landmark*	0	0	0	0	0	0	0	0	0	0
Logically Related	0	2	1	0	0	0	0	1	1	0
Nominal Variable	0	0	0	0	0	0	0	0	0	0
Positional	0	2	0	1	0	1	0	1	1	0
Ratio	1	1	21	3	1	3	1	14	8	0
Repetition	0	0	0	0	0	0	0	0	0	0
Single Structure Ratio	0	0	8	0	0	0	1	4	1	0
Unifying	1	0	2	4	4	1	0	3	3	2
Unspecified Homologue	0	27	0	0	0	0	0	0	1	0
Total Unconventional	**9**	**27**	**36**	**8**	**9**	**7**	**3**	**35**	**18**	**7**

	E31	E32	E33	E34	E35	E36	E37	E38	E39	E40
Total Number of Characters	**41**	**29**	**-**	**103**	**32**	**64**	**38**	**91**	**96**	**195**
Behavioural*	0	0	-	0	0	0	0	0	0	0
Composite	3	0	-	3	0	0	0	0	7	0
Conjunction	0	0	-	0	0	0	0	2	0	1
Developmental*	0	0	-	0	0	0	0	0	0	0
Extent	0	3	-	1	1	3	1	4	2	4
Inapplicable Data (missing)	0	1	-	1	0	0	1	3	2	16
(result)	0	1	-	4	0	0	3	5	3	24
Inapplicable Data (multistate)	11	1	-	15	5	7	3	4	0	18
(cryptic)	4	0	-	1	1	0	0	0	0	0
Landmark*	0	0	-	0	0	0	0	0	0	0
Logically Related	0	0	-	0	0	0	0	0	1	0
Nominal Variable	0	0	-	0	0	0	0	0	0	0
Positional	1	6	-	12	1	5	9	2	1	21
Ratio	0	2	-	7	2	3	1	0	1	17
Repetition	0	0	-	0	0	0	0	0	0	0
Single Structure Ratio	1	1	-	0	2	1	1	0	1	4
Unifying	0	1	-	14	2	6	2	0	1	16
Unspecified Homologue	0	0	-	0	0	1	0	0	0	0
Total Unconventional	**12**	**12**	**-**	**48**	**12**	**23**	**16**	**13**	**16**	**82**

	F31	F32	F33	F34	F35	F36	F37	F38	F39	F40
Total Number of Characters	**213**	**56**	**20**	**46**	**88**	**38**	**51**	**32**	**20**	**20**
Behavioural*	0	0	0	0	0	0	0	0	0	0
Composite	3	6	0	1	0	0	1	1	0	1
Conjunction	0	0	0	0	0	0	0	0	0	0
Developmental*	0	0	0	0	0	0	0	0	0	0
Extent	1	1	0	3	0	0	2	0	0	0
Inapplicable Data (missing)	15	1	0	0	10	0	0	0	0	0
(result)	21	6	0	0	20	0	0	0	0	0
Inapplicable Data (multistate)	1	21	0	0	1	4	17	5	1	0
(cryptic)	1	1	0	0	0	0	6	0	0	0
Landmark*	0	0	0	0	0	0	0	0	0	0
Logically Related	1	0	1	0	0	0	0	0	0	0
Nominal Variable	0	0	0	0	5	0	0	0	0	0
Positional	27	7	2	3	1	0	7	4	1	0
Ratio	2	3	3	8	0	2	1	5	0	0
Repetition	0	0	0	0	0	0	0	0	0	0
Single Structure Ratio	0	2	0	1	0	1	1	0	0	0
Unifying	6	13	2	5	0	3	1	2	0	1
Unspecified Homologue	0	0	0	0	0	3	0	0	0	0
Total Unconventional	**55**	**39**	**7**	**20**	**16**	**10**	**25**	**15**	**2**	**2**

	E41	E42	E43	E44	E45	E46	E47	E48	E49	E50
Total Number of Characters	**154**	**51**	**36**	**40**	**59**	**88**	**29**	**44**	**45**	**37**
Behavioural*	2	0	0	0	1	0	0	0	3	0
Composite	4	2	1	3	3	0	0	0	1	1
Conjunction	0	0	0	0	1	0	1	0	1	0
Developmental*	0	0	0	0	0	0	0	0	0	1
Extent	2	2	0	1	0	0	1	0	2	1
Inapplicable Data (missing)	4	0	1	1	1	10	0	1	1	0
(result)	7	0	1	1	1	21	0	2	2	0
Inapplicable Data (multistate)	7	0	5	8	11	0	0	0	1	3
(cryptic)	4	0	3	2	0	0	0	0	0	0
Landmark*	0	0	0	0	0	0	0	0	0	0
Logically Related	10	2	0	0	0	0	0	0	0	0
Nominal Variable	0	0	0	0	0	16	0	0	0	0
Positional	7	2	0	2	3	0	0	0	2	2
Ratio	5	11	1	0	6	0	2	3	6	1
Repetition	0	0	0	0	0	0	0	0	0	0
Single Structure Ratio	2	0	0	0	0	0	0	1	0	0
Unifying	7	0	0	6	0	0	0	0	9	3
Unspecified Homologue	1	0	0	0	0	0	0	44	0	0
Total Unconventional	**43**	**17**	**7**	**18**	**23**	**25**	**4**	**44**	**22**	**11**

	F41	F42	F43	F44	F45	F46	F47	F48	F49	F50
Total Number of Characters	**95**	**77**	**29**	**173**	**40**	**32**	**39**	**125**	**60**	**54**
Behavioural*	0	0	0	0	0	0	0	0	0	0
Composite	7	0	0	1	0	2	1	4	1	1
Conjunction	0	0	0	0	0	0	0	0	0	0
Developmental*	0	0	0	0	0	0	0	0	0	0
Extent	5	0	1	5	1	1	0	3	0	1
Inapplicable Data (missing)	1	2	1	0	0	0	1	11	2	3
(result)	1	2	1	0	0	0	2	13	3	3
Inapplicable Data (multistate)	0	7	4	6	0	2	14	14	0	0
(cryptic)	0	0	1	0	0	0	1	2	0	0
Landmark*	0	0	0	0	0	0	0	0	0	0
Logically Related	0	0	3	2	0	0	0	0	1	0
Nominal Variable	0	0	0	0	0	0	0	0	0	0
Positional	6	11	3	9	1	1	2	5	0	5
Ratio	7	2	1	6	1	1	0	12	4	1
Repetition	0	0	0	0	0	0	0	0	0	0
Single Structure Ratio	3	0	0	1	0	2	0	1	2	3
Unifying	11	1	1	6	4	5	1	0	2	5
Unspecified Homologue	0	1	0	0	0	0	2	0	0	0
Total Unconventional	**37**	**21**	**13**	**33**	**7**	**12**	**17**	**40**	**10**	**19**

Total (E)	%age (E)
3285	
25	2.62
78	8.16
7	0.73
1	0.10
74	7.74
90	9.41
148	15.48
291	30.44
64	6.69
9	0.94
20	2.09
16	1.67
139	14.54
111	11.61
3	0.31
23	2.41
133	13.91
51	5.33
808	**29.10**

Total (F)	%age (F)	Total (All)	%age (All)
3071		**6356**	
-	-	25	1.25
84	9.30	162	8.75
2	0.22	9	0.47
-	-	1	0.05
72	7.97	146	7.86
69	7.64	159	8.49
98	10.85	246	13.07
226	25.03	517	27.62
42	4.65	106	5.63
0	0.00	9	0.45
17	1.88	37	1.98
6	0.66	22	1.15
135	14.95	274	14.75
211	23.37	322	17.73
0	0.00	3	0.15
56	6.20	79	4.38
119	13.18	252	13.53
37	4.10	88	4.69
903	**29.40**	**2105**	**29.26**

Appendix 2b

	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10
Total Number of Characters	**207**	**55**	**112**	**-**	**86**	**25**	**98**	**144**	**64**	**38**
Behavioural*	0.00	0.00	0.00	-	0.00	0.00	0.00	0.00	0.00	0.00
Composite	2.42	0.00	2.68	-	10.47	4.00	0.00	2.08	0.00	0.00
Conjunction	0.00	0.00	0.00	-	0.00	0.00	0.00	0.00	0.00	0.00
Developmental*	0.00	0.00	0.00	-	0.00	0.00	0.00	0.00	0.00	0.00
Extent	1.93	0.00	0.89	-	0.00	0.00	0.00	4.86	0.00	0.00
Inapplicable Data (missing)	4.83	1.82	6.25	-	1.16	8.00	2.04	6.25	0.00	5.26
Inapplicable Data (multistate)	2.42	5.45	16.96	-	9.30	12.00	3.06	4.17	12.50	2.63
Landmark	0.00	0.00	0.00	-	0.00	0.00	0.00	0.00	0.00	0.00
Logically Related	0.00	0.00	0.00	-	0.00	0.00	1.02	0.00	1.56	0.00
Nominal Variable	0.00	0.00	0.00	-	0.00	0.00	0.00	0.00	0.00	0.00
Positional	0.48	0.00	1.79	-	0.00	0.00	6.12	4.17	3.13	0.00
Ratio	4.35	9.09	1.79	-	6.98	0.00	0.00	1.39	0.00	2.63
Repetition	0.97	0.00	0.00	-	0.00	0.00	0.00	0.69	0.00	0.00
Single Structure Ratio	0.00	3.64	1.79	-	0.00	0.00	0.00	0.00	0.00	0.00
Unifying	0.00	1.82	3.57	-	3.49	0.00	4.08	4.86	4.69	0.00
Unspecified Homologue	0.00	1.82	1.79	-	1.16	0.00	0.00	0.00	0.00	2.63

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
Total Number of Characters	**65**	**51**	**50**	**34**	**73**	**28**	**39**	**28**	**36**	**128**
Behavioural*	-	-	-	-	-	-	-	-	-	-
Composite	0.00	7.84	2.00	0.00	0.00	0.00	7.69	3.57	0.00	3.91
Conjunction	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Developmental*	-	-	-	-	-	-	-	-	-	-
Extent	3.08	5.88	6.00	2.94	1.37	0.00	2.56	0.00	0.00	2.34
Inapplicable Data (missing)	3.08	0.00	4.00	0.00	0.00	0.00	0.00	0.00	0.00	5.47
Inapplicable Data (multistate)	9.23	0.00	2.00	0.00	1.37	0.00	5.13	3.57	2.78	11.72
Landmark	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Logically Related	0.00	0.00	0.00	0.00	1.37	0.00	0.00	0.00	0.00	0.78
Nominal Variable	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Positional	0.00	1.96	0.00	0.00	4.11	3.57	10.26	0.00	0.00	3.91
Ratio	3.08	19.61	18.00	8.82	13.70	17.86	20.51	7.14	27.78	1.56
Repetition	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Single Structure Ratio	3.08	1.96	6.00	0.00	2.74	7.14	0.00	3.57	2.78	2.34
Unifying	0.00	3.92	0.00	5.88	0.00	0.00	0.00	0.00	8.33	0.78
Unspecified Homologue	0.00	0.00	0.00	0.00	2.74	0.00	0.00	0.00	0.00	0.00

	E11	E12	E13	E14	E15	E16	E17	E18	E19	E20
Total Number of Characters	**27**	**-**	**36**	**43**	**22**	**75**	**23**	**46**	**176**	**115**
Behavioural*	0.00	-	0.00	0.00	0.00	1.33	0.00	0.00	1.70	0.00
Composite	11.11	-	0.00	9.30	4.55	6.67	0.00	2.17	1.70	0.00
Conjunction	0.00	-	0.00	0.00	0.00	0.00	4.35	0.00	0.00	0.00
Developmental*	0.00	-	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Extent	0.00	-	0.00	4.65	9.09	0.00	0.00	2.17	0.00	3.48
Inapplicable Data (missing)	0.00	-	2.78	9.30	4.55	2.67	0.00	0.00	0.00	0.87
Inapplicable Data (multistate)	37.04	-	5.56	18.60	22.73	14.67	13.04	8.70	13.07	1.74
Landmark	0.00	-	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Logically Related	0.00	-	0.00	0.00	0.00	0.00	0.00	2.17	0.00	0.00
Nominal Variable	0.00	-	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Positional	11.11	-	0.00	2.33	4.55	2.67	0.00	8.70	0.57	6.09
Ratio	3.70	-	0.00	0.00	0.00	1.33	0.00	4.35	0.00	3.48
Repetition	0.00	-	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Single Structure Ratio	7.41	-	2.78	0.00	0.00	0.00	0.00	2.17	0.00	0.00
Unifying	3.70	-	2.78	4.65	13.64	6.67	0.00	4.35	1.70	6.96
Unspecified Homologue	0.00	-	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20
Total Number of Characters	**144**	**50**	**68**	**47**	**45**	**31**	**39**	**124**	**67**	**40**
Behavioural*	-	-	-	-	-	-	-	-	-	-
Composite	0.69	4.00	2.94	2.13	4.44	9.68	0.00	8.06	1.49	0.00
Conjunction	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.61	0.00	0.00
Developmental*	-	-	-	-	-	-	-	-	-	-
Extent	0.00	12.00	8.82	4.26	6.67	0.00	2.56	4.84	0.00	0.00
Inapplicable Data (missing)	0.00	0.00	0.00	0.00	0.00	0.00	2.56	2.42	1.49	0.00
Inapplicable Data (multistate)	9.72	6.00	2.94	12.77	40.00	9.68	0.00	17.74	13.43	0.00
Landmark	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Logically Related	0.00	0.00	0.00	0.00	0.00	0.00	5.13	0.00	0.00	0.00
Nominal Variable	0.00	0.00	0.00	0.00	0.00	0.00	2.56	0.00	0.00	0.00
Positional	0.00	4.00	2.94	4.26	0.00	6.45	5.13	3.23	5.97	5.00
Ratio	6.94	2.00	7.35	8.51	11.11	12.90	0.00	4.03	1.49	7.50
Repetition	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Single Structure Ratio	4.17	0.00	2.94	2.13	0.00	3.23	0.00	0.00	0.00	0.00
Unifying	0.00	4.00	2.94	4.26	13.33	6.45	10.26	2.42	0.00	2.50
Unspecified Homologue	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.49	0.00

	E21	E22	E23	E24	E25	E26	E27	E28	E29	E30
Total Number of Characters	**164**	**175**	**43**	**55**	**48**	**-**	**38**	**41**	**20**	**37**
Behavioural*	0.00	5.14	0.00	0.00	2.08	-	13.16	0.00	0.00	0.00
Composite	3.66	0.00	11.63	0.00	2.08	-	0.00	0.00	0.00	0.00
Conjunction	0.00	0.00	0.00	0.00	0.00	-	0.00	0.00	0.00	0.00
Developmental*	0.00	0.00	0.00	0.00	0.00	-	0.00	0.00	0.00	0.00
Extent	7.32	4.57	0.00	0.00	6.25	-	2.63	0.00	0.00	2.70
Inapplicable Data (missing)	0.61	0.57	4.65	0.00	0.00	-	0.00	0.00	0.00	0.00
Inapplicable Data (multistate)	9.15	10.86	4.65	5.45	20.83	-	5.26	9.76	20.00	24.32
Landmark	0.00	0.00	0.00	0.00	0.00	-	23.68	0.00	0.00	0.00
Logically Related	0.61	1.14	0.00	1.82	0.00	-	0.00	0.00	0.00	0.00
Nominal Variable	0.00	0.00	0.00	0.00	0.00	-	0.00	0.00	0.00	0.00
Positional	10.98	1.71	4.65	1.82	2.08	-	2.63	2.44	0.00	0.00
Ratio	3.05	0.00	0.00	0.00	2.08	-	2.63	0.00	5.00	5.41
Repetition	0.00	0.00	0.00	0.00	0.00	-	0.00	0.00	0.00	0.00
Single Structure Ratio	0.61	0.00	0.00	0.00	0.00	-	0.00	0.00	0.00	0.00
Unifying	4.88	2.86	6.98	0.00	0.00	-	7.89	0.00	0.00	0.00
Unspecified Homologue	0.00	0.00	0.00	0.00	0.00	-	0.00	0.00	0.00	0.00

	F21	F22	F23	F24	F25	F26	F27	F28	F29	F30
Total Number of Characters	**27**	**27**	**109**	**47**	**71**	**22**	**35**	**128**	**75**	**35**
Behavioural*	-	-	-	-	-	-	-	-	-	-
Composite	11.11	3.70	3.67	0.00	2.82	4.55	0.00	3.13	4.00	0.00
Conjunction	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Developmental*	-	-	-	-	-	-	-	-	-	-
Extent	14.81	3.70	1.83	0.00	1.41	0.00	0.00	1.56	0.00	0.00
Inapplicable Data (missing)	0.00	0.00	1.83	0.00	2.82	0.00	0.00	0.78	1.33	0.00
Inapplicable Data (multistate)	14.81	0.00	0.92	0.00	0.00	4.55	2.86	9.38	0.00	17.14
Landmark	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Logically Related	0.00	7.41	0.92	0.00	0.00	0.00	0.00	0.78	1.33	0.00
Nominal Variable	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Positional	0.00	7.41	0.00	2.13	0.00	4.55	0.00	0.78	1.33	0.00
Ratio	3.70	3.70	19.27	6.38	1.41	13.64	2.86	10.94	10.67	0.00
Repetition	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Single Structure Ratio	0.00	0.00	7.34	0.00	0.00	0.00	2.86	3.13	1.33	0.00
Unifying	3.70	0.00	1.83	8.51	5.63	4.55	0.00	2.34	4.00	5.71
Unspecified Homologue	0.00	100	0.00	0.00	0.00	0.00	0.00	0.00	1.33	0.00

	E31	E32	E33	E34	E35	E36	E37	E38	E39	E40
Total Number of Characters	**41**	**29**	**-**	**103**	**32**	**64**	**38**	**91**	**96**	**195**
Behavioural*	0.00	0.00	-	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Composite	7.32	0.00	-	2.91	0.00	0.00	0.00	0.00	7.29	0.00
Conjunction	0.00	0.00	-	0.00	0.00	0.00	0.00	2.20	0.00	0.51
Developmental*	0.00	0.00	-	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Extent	0.00	10.34	-	0.97	3.13	4.69	2.63	4.40	2.08	2.05
Inapplicable Data (missing)	0.00	3.45	-	0.97	0.00	0.00	2.63	3.30	2.08	8.21
Inapplicable Data (multistate)	26.83	3.45	-	14.56	15.63	10.94	7.89	4.40	0.00	9.23
Landmark	0.00	0.00	-	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Logically Related	0.00	0.00	-	0.00	0.00	0.00	0.00	0.00	1.04	0.00
Nominal Variable	0.00	0.00	-	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Positional	2.44	20.69	-	11.65	3.13	7.81	23.68	2.20	1.04	10.77
Ratio	0.00	6.90	-	6.80	6.25	4.69	2.63	0.00	1.04	8.72
Repetition	0.00	0.00	-	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Single Structure Ratio	2.44	3.45	-	0.00	6.25	1.56	2.63	0.00	1.04	2.05
Unifying	0.00	3.45	-	13.59	6.25	9.38	5.26	0.00	1.04	8.21
Unspecified Homologue	0.00	0.00	-	0.00	0.00	1.56	0.00	0.00	0.00	0.00

	F31	F32	F33	F34	F35	F36	F37	F38	F39	F40
Total Number of Characters	**213**	**56**	**20**	**46**	**88**	**38**	**51**	**32**	**20**	**20**
Behavioural*	-	-	-	-	-	-	-	-	-	-
Composite	1.41	10.71	0.00	2.17	0.00	0.00	1.96	3.13	0.00	5.00
Conjunction	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Developmental*	-	-	-	-	-	-	-	-	-	-
Extent	0.47	1.79	0.00	6.52	0.00	0.00	3.92	0.00	0.00	0.00
Inapplicable Data (missing)	7.04	1.79	0.00	0.00	11.36	0.00	0.00	0.00	0.00	0.00
Inapplicable Data (multistate)	0.47	37.50	0.00	0.00	1.14	10.53	33.33	15.63	5.00	0.00
Landmark	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Logically Related	0.47	0.00	5.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Nominal Variable	0.00	0.00	0.00	0.00	5.68	0.00	0.00	0.00	0.00	0.00
Positional	12.68	12.50	10.00	6.52	1.14	0.00	13.73	12.50	5.00	0.00
Ratio	0.94	5.36	15.00	17.39	0.00	5.26	1.96	15.63	0.00	0.00
Repetition	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Single Structure Ratio	0.00	3.57	0.00	2.17	0.00	2.63	1.96	0.00	0.00	0.00
Unifying	2.82	23.21	10.00	10.87	0.00	7.89	1.96	6.25	0.00	5.00
Unspecified Homologue	0.00	0.00	0.00	0.00	0.00	7.89	0.00	0.00	0.00	0.00

	E41	E42	E43	E44	E45	E46	E47	E48	E49	E50
Total Number of Characters	**154**	**51**	**36**	**40**	**59**	**88**	**29**	**44**	**45**	**37**
Behavioural*	1.30	0.00	0.00	0.00	1.69	0.00	0.00	0.00	6.67	0.00
Composite	2.60	3.92	2.78	7.50	5.08	0.00	0.00	0.00	2.22	2.70
Conjunction	0.00	0.00	0.00	0.00	1.69	0.00	3.45	0.00	2.22	0.00
Developmental*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.70
Extent	1.30	3.92	0.00	2.50	0.00	0.00	3.45	0.00	4.44	2.70
Inapplicable Data (missing)	2.60	0.00	2.78	2.50	1.69	11.36	0.00	2.27	2.22	0.00
Inapplicable Data (multistate)	4.55	0.00	13.89	20.00	18.64	0.00	0.00	0.00	2.22	8.11
Landmark	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Logically Related	6.49	3.92	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Nominal Variable	0.00	0.00	0.00	0.00	0.00	18.18	0.00	0.00	0.00	0.00
Positional	4.55	3.92	0.00	5.00	5.08	0.00	0.00	0.00	4.44	5.41
Ratio	3.25	21.57	2.78	0.00	10.17	0.00	6.90	6.82	13.33	2.70
Repetition	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Single Structure Ratio	1.30	0.00	0.00	0.00	0.00	0.00	0.00	2.27	0.00	0.00
Unifying	4.55	0.00	0.00	15.00	0.00	0.00	0.00	0.00	20.00	8.11
Unspecified Homologue	0.65	0.00	0.00	0.00	0.00	0.00	0.00	100	0.00	0.00

	F41	F42	F43	F44	F45	F46	F47	F48	F49	F50
Total Number of Characters	**95**	**77**	**29**	**173**	**40**	**32**	**39**	**125**	**60**	**54**
Behavioural*	-	-	-	-	-	-	-	-	-	-
Composite	7.37	0.00	0.00	0.58	0.00	6.25	2.56	3.20	1.67	1.85
Conjunction	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Developmental*	-	-	-	-	-	-	-	-	-	-
Extent	5.26	0.00	3.45	2.89	2.50	3.13	0.00	2.40	0.00	1.85
Inapplicable Data (missing)	1.05	2.60	3.45	0.00	0.00	0.00	2.56	8.80	3.33	5.56
Inapplicable Data (multistate)	0.00	9.09	13.79	3.47	0.00	6.25	35.90	11.20	0.00	0.00
Landmark	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Logically Related	0.00	0.00	10.34	1.16	0.00	0.00	0.00	0.00	1.67	0.00
Nominal Variable	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Positional	6.32	14.29	10.34	5.20	2.50	3.13	5.13	4.00	0.00	9.26
Ratio	7.37	2.60	3.45	3.47	2.50	3.13	0.00	9.60	6.67	1.85
Repetition	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Single Structure Ratio	3.16	0.00	0.00	0.58	0.00	6.25	0.00	0.80	3.33	5.56
Unifying	11.58	1.30	3.45	3.47	10.00	15.63	2.56	0.00	3.33	9.26
Unspecified Homologue	0.00	1.30	0.00	0.00	0.00	0.00	5.13	0.00	0.00	0.00

Average (E)
0.66
2.38
0.29
0.05
1.98
2.15
9.49
0.47
0.40
0.36
3.80
3.24
0.03
0.83
3.69
2.19

Average (F)	Average (All)
-	0.32
2.79	2.59
0.03	0.16
-	0.03
2.42	2.21
1.47	1.80
7.62	8.51
0.00	0.23
0.73	0.57
0.16	0.26
4.22	4.02
7.49	5.45
0.00	0.02
1.73	1.30
4.60	4.16
2.40	2.30

Appendix 3

The data matrix of amniotes used in Chapter 7. Character numbers are discussed in Chapter 7.

	12355 ab	67899 ab	11111 1234	11111 56778 ab	12222 90123	22222 45678	23333 91234	33333 56788 ab	34444 90122 ab	44444 34567
Seymouriidae	00000	00000	00000	00000	00010	?0000	?0000	0?000	00000	00000
Diadectomorpha	00000	00000	00000	00000	0001A	?0A00	?0A00	0?000	00000	00000
Caseidae	00000	00000	01000	00000	00000	10000	?0100	00100	00000	00000
Ophiacodontidae	00000	00000	10000	00000	10000	A1000	?0101	00100	00000	01000
Edaphosauridae	00000	00000	00001	00000	A?000	10000	?0A01	00100	00002	01000
Sphenacodontidae	00001	00000	11001	00010	10001	A0000	?0101	00100	00002	01000
Gorgonopsia	00001	00100	11001	0001?	1?000	A1000	?0101	0011?	???02	01000
Cynodontia	00001	10100	11001	00011	10010	?0000	?0A01	0011?	???02	0?011
Millerettidae	?0000	00000	00010	0000?	01000	0A000	?0000	00100	A0000	00000
Acleistorhinus	00000	10000	01010	0101?	0?001	10000	?1000	00100	00?00	00000
Lanthanosuchidae	00000	10000	01110	0A01?	1?001	10100	?0100	00100	00000	1???0
Nyctoleteridae	00000	10000	01010	0101?	0?000	01001	?0000	00101	00000	0???0
Procolophon	00000	10000	01011	00010	01100	01001	?0000	10101	00001	A0110
Owenetta	00000	10000	01011	0001?	0?100	01001	?0000	11101	00001	0???0
Bradysaurus	00010	10100	01011	01000	01010	?0100	?0000	00100	00000	1??10
Scutosaurus	000??	10100	01111	01000	01010	?0100	?0000	00100	00000	10010
Anthodon	00000	10100	01011	01000	01010	?0100	?0000	00100	00000	1??10
Captorhinidae	00000	00000	00001	00000	00000	00000	?0000	01100	00100	00000
Protorothyrididae	00000	00000	00000	0000?	00000	00000	?0000	01100	00100	00000
Araeoscelidia	00000	00000	00001	0000?	10000	01000	00000	01100	00100	00000
Claudiosaurus	00000	00000	0?001	1001?	1?00?	01000	01100	01100	00102	00??0
Younginiformes	00000	00000	00000	00010	10000	01000	11100	01101	10100	00110
Kuehneosauridae	0001?	11000	01001	1?01?	0?000	00010	10100	2??01	1111?	?0??0
Choristodera	10000	01?01	10001	00011	2?000	00000	00100	0??01	10?A2	00111
Trilophosaurus	1?11?	10?00	0?00?	?????	1?011	?0010	01??0	???01	10?01	00?11
Placodus	1001?	00101	01001	0?1??	21010	?1100	11000	11101	1011?	?0??0
Cyamodus	1001?	1010A	01000	0?1??	2?00A	0A000	01000	11101	101??	?0??0
Simosaurus	1001?	10?00	01000	101??	1?000	00010	11100	???0?	1010?	0???1
Proganochelys	0001?	00100	01000	10011	2?010	?0100	?0000	11101	10001	00010
Kayentachelys	0011?	01?00	01000	101??	2?000	00100	?0000	1????	1??01	????0
Euparkaria	1001?	10000	11000	0001?	1?000	00010	00110	11101	10102	00110
Proterosuchus	1001?	10000	11000	0001?	1?000	00010	00110	1?001	10101	0??10
Stagonolepis	1001?	00000	11001	0001?	1?000	00000	00100	11101	10101	00110
Mesosuchus	1101?	11000	01000	0001?	0?000	00000	00110	01001	10?02	0???1
Rhynchosaurus	1111?	11000	01000	00020	2?000	00000	00010	11001	10?01	00111
Clevosaurus	10000	00000	01000	101??	0?000	01010	11010	00101	11?02	00111
Huehuecuetzpalli	000??	10?00	01000	0?021	1?000	01010	00000	1??01	11?1?	?0111
Pamelaria	100??	11000	11000	0001?	1?000	00010	00100	11001	10101	00110
Prolacerta	100??	10000	11001	00011	1?000	00010	00100	11001	10101	00110

A=(01) B=(02) C=(12)

	44555 99011 ab ab	55555 12233 cabab	55555 45567 ab	55566 89901 ab	66666 23456	66677 78901	77777 23344 abab	77777 56788 ab	78888 90111 abc	88888 23456
Seymouriidae	0000?	?0000	00000	00?00	00000	00100	0000?	00000	0?00?	01000
Diadectomorpha	0000?	?0100	10A00	00?00	00000	10100	0000?	00000	0?00?	01A00
Caseidae	02010	00000	10000	11000	00000	00100	0000?	00001	0?00?	0000?
Ophiacodontidae	00010	10101	10000	11000	0?000	??100	0000?	00001	0?00?	0000?
Edaphosauridae	00010	10101	10100	11000	1?021	0?100	0000?	?1001	1?01?	0000?
Sphenacodontidae	00010	10101	10100	11000	1?021	10100	0000?	01001	0?00?	0000?
Gorgonopsia	02010	1011?	10100	11001	11121	?1112	1010?	1201?	0?00?	02100
Cynodontia	02010	1011?	10100	11001	11121	11112	1010?	1201?	0?00?	22101
Millerettidae	000A0	10000	10001	01100	0?000	?01??	00010	00001	0?00?	2010?
Acleistorhinus	02010	00100	10001	0??20	1?000	??1??	?000?	00001	0?00?	2?1??
Lanthanosuchidae	0B010	00100	10001	01020	00000	??1??	?000?	01001	0?00?	A11??
Nyctoleteridae	020??	?0000	11?0?	0110?	0?00?	??1??	000??	0?001	0?00?	211?1
Procolophon	0200?	?1?00	11?1?	01100	A0011	00100	00110	1111?	0000?	21111
Owenetta	0200?	?A000	11?1?	01100	0?010	??1??	00010	1111?	0000?	21111
Bradysaurus	0200?	?0100	11?1?	01101	01121	01000	11?11	11101	1100?	21111
Scutosaurus	0200?	?0100	11?1?	01101	01121	01000	10111	11101	1100?	21111
Anthodon	0200?	?0100	11?1?	01101	01121	01000	10011	11101	1100?	21111
Captorhinidae	0B00?	?0101	11?0?	01100	00000	00100	00010	00001	0?1?0	0010?
Protorothyrididae	0000?	?0001	10001	01100	1?000	??1??	000??	01001	0?00?	001??
Araeoscelidia	0010?	?0001	10001	01100	1?000	??1??	00010	00101	0000?	0010?
Claudiosaurus	00111	0??01	11?0?	01100	??100	??1??	00010	00101	1000?	001??
Younginiformes	00110	00001	1A001	01100	1?000	??10?	00010	00001	1?00?	011??
Kuehneosauridae	02111	?1?1?	11?0?	0110?	1?01?	??0??	00010	???00	?????	001??
Choristodera	1?110	01?1?	11???	01110	1?001	1?01?	?0010	11000	0?00?	1010?
Trilophosaurus	1?10?	?1?1?	11?1?	01111	?1011	1101?	00010	1011?	1100?	21110
Placodus	02111	?1?1?	11?0?	01100	1?110	1101?	11?0?	12?1?	0100?	12111
Cyamodus	02111	?1?1?	11?0?	01110	0??10	11010	11?0?	12?1?	011?0	11110
Simosaurus	01111	01?1?	11?0?	0???0	?????	?10??	11?0?	?101?	1?00?	1010?
Proganochelys	1?00?	?1?01	11?1?	01111	?1121	10011	00011	1111?	111?1	01110
Kayentachelys	1?00?	?1?1?	11???	01101	1????	?????	?0?10	1101?	1????	2????
Euparkaria	1?110	01?1?	11?0?	01101	?1121	1?011	?0010	1101?	0?00?	1011?
Proterosuchus	00110	01?01	11?0?	00?11	11021	10011	00010	11001	0?00?	1011?
Stagonolepis	1?110	01?1?	11?0?	01111	11121	1001?	00010	1111?	0100?	0C11?
Mesosuchus	02111	01?01	11?1?	01111	11011	??11?	00010	1001?	0?00?	1011?
Rhynchosaurus	1?110	01?01	11?1?	01111	11011	??011	00010	?001?	0?00?	11111
Clevosaurus	02110	11?01	11?1?	01110	11011	1?010	00010	1101?	1?00?	02101
Huehuecuetzpalli	1?111	?1?01	11?1?	01100	?????	?????	0????	?????	?????	01100
Pamelaria	1?111	01?1?	11?0?	01110	11110	1?11?	00010	0001?	1?00?	01100
Prolacerta	01111	01?01	11?0?	01110	11110	1?11?	00010	00001	1?00?	01100

	88889	99999	99991	11111	11111	11111	11111	11111	11111	11111
	78890	12345	67890	00000	00111	11111	12222	22222	23333	33333
	ab		1	23567	89012	34567	90123	56678	90123	45678
								ab		
Seymouriidae	00100	00000	00000	00000	00000	00000	00000	00000	00000	00000
Diadectomorpha	00000	00A00	00000	00000	00000	00000	00000	00000	00000	00000
Caseidae	00001	?0100	00000	00010	00000	00001	01000	00000	00000	00000
Ophiacodontidae	00001	?1100	10000	00010	00000	0000A	01000	00000	00000	00000
Edaphosauridae	00001	?1101	00000	01010	01000	00001	01000	00000	00??0	00000
Sphenacodontidae	00001	?1100	10000	01010	01000	00001	01000	00000	00000	00000
Gorgonopsia	00001	01100	10010	01011	01110	00111	01000	00120	00101	00011
Cynodontia	01?21	01100	10A11	01011	01110	11111	01?01	01?20	01111	01011
Millerettidae	00000	?1100	00000	?0010	00?0?	000?0	?00?0	101?0	0?0?0	10000
Acleistorhinus	00?00	?0100	0????	?????	?????	?????	?????	?????	?????	?????
Lanthanosuchidae	001?0	?1101	00?00	???20	0????	?01?0	????0	????0	?????	?????
Nyctoleteridae	0??00	11100	00000	00020	01?0?	00101	?1000	00100	00000	10001
Procolophon	10100	11101	00A00	01020	01100	10101	?100A	01?10	01000	10011
Owenetta	10100	1110A	00?00	?0020	01?0?	?01??	?1000	?0101	01??0	100?0
Bradysaurus	10000	11101	01000	?1020	01111	00101	?1000	01?20	01000	11011
Scutosaurus	10000	11101	01100	?1020	01111	10101	?1000	01?20	01000	11011
Anthodon	10000	11101	01?00	?1020	01111	10102	?1001	11?20	01000	10011
Captorhinidae	00000	?1100	00000	00000	00000	A0000	01000	01?00	00000	00000
Protorothyrididae	00000	?0100	00000	01010	00000	00000	01000	00100	00000	00000
Araeoscelidia	00100	?0100	00000	01010	00000	00010	01000	0A100	00000	00000
Claudiosaurus	00100	?0101	000?0	11110	00010	111?0	10010	11?10	01010	00010
Younginiformes	0??00	?1101	00000	11010	001?0	A1A1?	10000	101B0	A1010	10010
Kuehneosauridae	???10	?1101	00010	?1?11	10?10	????0	10000	00101	2?111	?0010
Choristodera	001?0	?1101	00010	11111	00110	111??	10000	101B1	21110	?0010
Trilophosaurus	0???0	11121	00?10	11110	00?10	111?1	10000	11?01	00000	00110
Placodus	00000	11101	00000	?1111	11?1?	11100	10111	10101	21??1	?0010
Cyamodus	10000	?1101	00?1?	?1111	11?10	11?00	10111	10101	21??1	?0010
Simosaurus	10110	?1101	00011	11111	01010	101?2	10111	11?01	21?11	10010
Proganochelys	10110	0112?	?1110	11???	?0111	101?2	1?101	11?21	01100	00101
Kayentachelys	?????	?????	?????	?????	?????	??1??	?????	?????	?????	?????
Euparkaria	00100	?1101	00010	11110	10110	11??1	10001	10?11	01110	00110
Proterosuchus	00100	?1101	00010	?1110	10110	111?0	10011	10?11	01100	00100
Stagonolepis	00000	?1101	00010	?111?	10110	11??1	10001	10111	01110	00110
Mesosuchus	?0110	?1101	00010	11110	00110	111?0	10?0?	11?01	01100	00100
Rhynchosaurus	?0000	11101	00010	11?1?	00110	111?0	10001	1010?	01?00	00110
Clevosaurus	01?10	11111	00001	?1010	00110	112?0	10000	10120	00111	00010
Huehuecuetzpalli	?0110	??111	00011	11010	00110	11210	1000?	10120	20111	000?0
Pamelaria	00100	01101	00010	11110	00110	1???0	10000	10101	21110	00110
Prolacerta	00100	01101	00010	11110	00110	111?0	10000	10101	21100	00110

	11111 11111 11111 11111 11111 11111 1
	34444 44444 45555 55555 56666 66666 6
	90123 45678 90123 45678 90123 45567 8
	ab
Seymouriidae	00000 00000 00000 00000 00000 00??? 0
Diadectomorpha	00000 00000 00000 00000 00000 00??? 0
Caseidae	00000 00001 00000 00000 00000 00??? 0
Ophiacodontidae	00000 00001 00000 00000 00000 00??? 0
Edaphosauridae	00000 00001 ?00?? ??0?0 ???0? ?0??? 0
Sphenacodontidae	00000 00001 00000 00000 00000 00??? 0
Gorgonopsia	01210 10111 22101 00001 00200 00??? ?
Cynodontia	11211 10111 21101 01001 10211 00??? 1
Millerettidae	0000? ?00?1 00001 000?1 0?001 00??? 0
Acleistorhinus	????? ????? ????? ????? ????? ?0??? ?
Lanthanosuchidae	????? ????? ????? ????? ????? ?0??? ?
Nyctoleteridae	00000 1?001 00001 00001 00010 00??? ?
Procolophon	00000 10001 11001 00002 01100 1A000 0
Owenetta	000?? ?0?01 ?0??? ??0?? ???1? ?0??? 0
Bradysaurus	01000 10001 11001 00002 11200 11000 1
Scutosaurus	01100 10001 11001 00002 11200 11021 1
Anthodon	01100 ?0001 1?001 00002 1?200 11111 1
Captorhinidae	00000 00001 00000 00001 00000 00??? 0
Protorothyrididae	00000 00001 00000 00001 00011 00??? 0
Araeoscelidia	00000 01001 00000 00011 01011 00??? 0
Claudiosaurus	1?011 10001 00000 00001 01001 00??? 0
Younginiformes	1?010 11001 00000 10011 01011 00??? 0
Kuehneosauridae	1?010 1???1 0???? ??0?1 0101? 00??? 1
Choristodera	10011 11??1 0?101 110?1 0?011 00??? 0
Trilophosaurus	11010 11001 20101 11011 01011 00??? 0
Placodus	11011 11?01 00011 0000? 01000 0110? 0
Cyamodus	110?1 11??1 ??0?? ????? ???00 01100 0
Simosaurus	01?01 10101 0?011 0000? ?1?1? ?0??? 0
Proganochelys	11100 11111 1A001 11012 11200 11101 ?
Kayentachelys	????? ????? ????? ????? ????? ?110? ?
Euparkaria	10??? 11?11 20111 11002 11011 11000 0
Proterosuchus	10?11 11011 20101 11012 01011 11000 0
Stagonolepis	10?10 11?11 20111 11002 11011 11101 0
Mesosuchus	1??11 11011 20101 11011 01011 10??? 0
Rhynchosaurus	1???? ?1011 20101 00011 01011 10??? 0
Clevosaurus	1?010 1?111 11011 11102 ?1011 ?0??? 0
Huehuecuetzpalli	1???? ?1111 11011 11?02 01011 00??? 0
Pamelaria	11010 11001 20?01 11001 01011 10??? 0
Prolacerta	11010 11001 20?01 11001 01011 00??? 0